

MODELLING THE SOCIAL DYNAMICS OF FEMALE GENITAL MUTILATION: METHODOLOGICAL DISCUSSION AND DEVELOPMENT

Thesis submitted for the degree of Doctor of Philosophy at the
University of Leicester

By

Laurence T Droy BSc (Leicester)

School of Media, Communication and Sociology

University of Leicester

December 2019

Modelling the Social Dynamics of Female Genital Mutilation: Methodological Discussion and Development

By Laurence T Droy

Abstract

Female Genital Mutilation (FGM) is a harmful traditional practice involving injury to female genitals for non-medical reasons. Despite intense international effort, the practice remains widespread. Academics and policymakers have shown considerable interest in understanding the social dynamics of FGM. The understanding of the practice as a ‘Social Convention’ or ‘Social Norm of Coordination’ has had a noticeable impact on efforts to eradicate it.

Analysis of the social dynamics of FGM has drawn on formal modelling, especially game-theory. This research approach is critically discussed. The use of formal modelling is characterised as essential to the problem, but fraught with challenges. Examples of these challenges are discussed and organised around the concept of uncertainty in model design. This provides insight into the status of existing models and theory of FGM.

A novel modelling strategy is developed and implemented to help address model uncertainty, based on techniques from agent-based social simulation. As part of this strategy, distinctive ‘core’ theories of FGM decision-making are identified and subjected to empirical testing. Possible de-idealisations and elaborations of currently used coordination models are explored through robustness analysis. Subsequently, an agent-based simulation model is developed which encompasses a range of these design possibilities. Using sensitivity analysis and secondary data, key uncertainties in the design of this model are identified and calibrated using empirical data. The model is then tested and deployed to identify ‘possible scenarios’ through which social dynamics could undermine anti-FGM interventions.

Alongside contributions to modelling the social dynamics of FGM, this activity contributes new empirical research on FGM decision-making. It also offers new methodological directions to the field and leads to a model-centric perspective on theories of the social dynamics of FGM.

Acknowledgements

I would like to express deep gratitude to my first supervisor, Dr. Edmund Chattoe-Brown, for his unwavering encouragement and thoughtful guidance throughout the development of this thesis. I owe a great deal to his generosity. I would like to thank Prof. Henrietta O'Connor, Prof. Jason Hughes and Prof. John Goodwin for their kind support of me in my professional work in the department. Their support has made my doctoral journey possible and it has been a pleasure to work with, and learn from, them. Many thanks also to my second supervisor Dr. Patrick White, whose guidance and commentary on draft work has been invaluable. I'm grateful to Prof. Charles Efferson and Prof. Gerry Mackie for taking the time to discuss their work with me. I'm indebted to all my friends and family for their patience and encouragement. I'd like to offer special thanks to my partner Jess Ramsdale, my parents Malcolm and Victoria Droy, and my friend Arran Stirton.

Table of Contents

Abstract	2
Acknowledgements	3
Table of Contents.....	4
List of Tables	9
List of Figures	10
List of Abbreviations	14
Early Outputs of the Thesis Project.....	14
Chapter 1: Social Dynamics, Formal Modelling and the Widespread Harm of Female Genital Mutilation	15
Introduction.....	15
Defining Female Genital Mutilation (FGM) and Adopting Appropriate Terminology	17
Female Genital Mutilation as a Harm to Health.....	18
Female Genital Mutilation as a Violation of Internationally Recognised Human Rights	19
FGM as a Prevalent and Growing Social Problem (Despite Opposition).....	20
The Role of Theories of the Social Dynamics of FGM.....	22
Formal Modelling and the Social Dynamics of FGM.....	26
Foundations of the Challenge Facing Modellers of Social Dynamics	29
Conclusions	33
Chapter 2: Models of the Social Dynamics of FGM, A Critical Discussion.....	35
Introduction.....	35
The Social Convention Model of FGM.....	37
Critiques of the Social Convention Model and the Problem of Idealisation.....	43
Idealisation and Diverging Model Designs	61
Empirical Validation of Formal Models of FGM	65
Conclusions	69
Chapter 3: Managing Uncertainty in Models of the Social Dynamics of FGM, Methodological Techniques	71
Introduction.....	71
Managing Uncertainty in Model Design.....	72

Combining Techniques for the Management of Model Uncertainty	81
From Traditional Analytical Methods to Agent-based modelling	82
The ‘Theoretical Core’: Separating Models of Social Dynamics from Theories of FGM Decision-Making	96
Conclusion: The Modelling Strategy	98
Chapter 4: The Theoretical Core – Hypotheses about FGM Decision-Making Processes in Senegal	101
Introduction.....	101
Hypothesis 1: Coordinated Decision-Making.....	103
Hypothesis 2: Competitive Decision Making.....	104
Hypothesis 3: Informational Influence.....	105
Avenues for Discriminating Between the Three Hypotheses	107
Beliefs about FGM in National Surveys.....	110
Testing Decision-Hypotheses Against the Relative Contribution of Beliefs to Decision-Making.....	111
Quantifying Variable (Belief) Importance	113
Methods, Context, and Data.....	114
Results	119
Discussion.....	123
Limitations of the Empirical Analysis.....	123
Conclusions	124
Chapter 5: Exploring the Robustness of Coordination Models of FGM to De-Idealisations and Elaborations	126
Introduction.....	126
The Theoretical Core: Social Coordination.....	127
The Standard Coordination Model	128
Key Dynamics of the Standard Coordination Model.....	133
Recognised Areas of Non-Robustness	137
‘Worked Example’: Representing the Intervention Process.....	140
Modifying the Decision-Making Process.....	145

Incorporating Heterogeneity of Social Influence	151
Introducing Social Structure	154
Incorporating Additional Intervention Features:.....	157
Results of the Robustness Analysis	160
Model Complexity and Interactions Between Features	163
Conclusions	168
Chapter 6: A General Model of FGM As a Social Norm of Coordination	169
Introduction.....	169
‘Generality’ in Modelling.....	170
Advantages of Building a General Model Prior to Calibration.....	171
Goals of the Analysis.....	172
Generalising Over Model Design Possibilities.....	173
Global Sensitivity Analysis	188
Implications of the Results.....	192
Chapter 7: Calibration and Validation of an ABM of the Social Dynamics of FGM.....	193
Introduction.....	193
Calibration and Validation of ABMs.....	193
Approach to Calibration and Validation Taken in this Project.....	203
Past Calibration of Models of the Social Dynamics of FGM	204
Targeting Parameters for Calibration.....	205
Strategy for Quantitative Estimation of μ_H and σ_H^2	207
Strategy for Quantitative Estimation of s^2	212
Strategy for Qualitative Estimation of $\rho_H y$	214
Empirical Calibration.....	216
Empirical Validation	223
Conclusions	237

Chapter 8: Possibilistic Failure Scenario Analysis.....	238
Introduction.....	238
Intervention Failure and Past Experiences.....	239
Intervention Failure in Past Coordination Models of FGM	240
Approach to PFSA.....	242
Failure Scenarios	245
Discussion and Conclusions	253
Chapter 9: Final Conclusions and Avenues for Further Development.....	256
Introduction.....	256
Methodological Contributions.....	257
Theoretical, Empirical and Applied Modelling Contributions.....	262
Limitations of the Analysis and Future Avenues for Progressive Modelling of FGM.....	267
Final Thought: A Model-Centric Perspective on Social Norm Theory as a Theory of Social Dynamics	272
Appendix A: Chapter 2 Appendices.....	273
Appendix A1: The Critical Mass in UNICEF (2007).....	273
Appendix A2: Coyne and Coyne’s Identity Economics Model of FGM.....	274
Appendix A3: The Diversity of Formal Models of FGM	280
Appendix A4: Survey of ‘Validation’ of Formal Models of FGM.....	285
Appendix B: Chapter 4 Appendices.....	289
Appendix B1: Analytical Strategy for Calculating SZOC and SSPC	289
Appendix B2: Robustness of Variable Importance Results across Regions of Senegal	291
Appendix B3: Robustness of Results to Sample Weighting.....	293
Appendix C: Chapter 5 Appendices.....	294
Appendix C1: Sources of Model Elaborations and De-Idealisations in Existing Literature.....	294
Appendix C2: Calculation of the Rate of ‘Final’ Decision-Making about the FGM Status of Girls within a Local Community	296
Appendix C3: A Simple Spatial Network Algorithm	298

Appendix C4: Creating Household Network Cliques	300
Appendix C5: Implementing an ‘Organised Diffusion’ Process	301
Appendix C6: Exploratory robustness simulation results	302
Appendix D: Chapter 6 Appendices	306
Appendix D1: Generalising Over Sources of Social Influence	306
Appendix D2: Breaking the Notion of ‘Thresholds’	309
Appendix D3: Re-Parameterising the Beta-Distribution in Terms of Mean and Variance	310
Appendix D4: Generalising over the Joint-Distribution of Continuous Agent Attributes in the ABM	314
Appendix D5: Results for the Global Sensitivity Analysis of Intervention Parameters	318
Appendix E: Chapter 7 Appendices	319
Appendix E1: Demonstration of the Latent Variable Estimation Approach	319
Appendix E2: Necessary Assumptions to Estimate H_i for participants in the 2005 Senegal DHS survey	320
Appendix E3: Arriving at the ‘Latent Variable’ Strategy through a Pragmatic Approach	323
Appendix E4: Description of Monte-Carlo Experiment	324
Appendix F: Chapter 8 Appendices	330
Appendix F1: Exploratory Analysis of Failure Scenarios	330
Appendix F2: Failure Scenario 2 (Out-of-Reach Effects)	331
Appendix F3: Failure Scenario 3 (Supporters Gambit Effects)	335
Appendix F4: Failure Scenario 4 (Localised Deviance Effects)	338
Appendix F5: Failure Scenario 5 (Social-Hump Effects)	346
Appendix F6: Failure Scenario 6 (Power Imbalance Effects)	353
Appendix F7: Failure Scenario 7 (Enforcement Imbalance Effects)	358
Appendix G: Early Outputs of the Thesis Project	362
Appendix H: Special Materials	364
Appendix H1: Domain Theory Versus ‘Off-The-Shelf’ Behavioural Frameworks	364
Bibliography	372

List of Tables

Table 1: Matrix Form Social Convention Model	39
Table 2: Variable Frequency and Importance (Beliefs about FGM, Senegal DHS 2005)	119
Table 3: Types of Threshold Distribution and their Effects on Dynamics	137
Table 4: Effects on Thresholds of Representing FGM as an Intrinsic Benefit	147
Table 5: Parameters Controlling the Marginal Distribution of Agent-Characteristics.....	181
Table 6: Parameters Controlling the Joint Distribution of Agent-Characteristics	182
Table 7: More-Order Rules Implemented in the Simulation	185
Table 8: Intervention Targeting Biases in the General Model.....	187
Table 9: Global Sensitivity Analysis of the General Model	190
Table 10: Logistic Regression of Support for FGM on Beliefs, Education and Wealth (Senegal 2005 DHS)	217
Table 11: Validation Performance of the ABM (Correlation Metrics)	228
Table 12: Summary of Failure Scenarios 2-7	252

Appendices:

Table A3.1: Formal Models of the Social Dynamics of FGM	280
Table A4.1: Validation of Formal Models of FGM.....	285
Table D5.1: Global Sensitivity Analysis Results (Intervention Parameters)	318
Table E4.1: MC Error rates for Mean(H) Estimator (by r-squared of the model)	329
Table E4.2: MC Error rates for s_2 Estimator (by r-squared of the model)	329
Table E4.3: MC Error rates for Var(H) Estimator (by r-squared of the model).....	329
Table F5.1: The Complex Relation Between Heterogeneity of Autonomy and Threshold Distributions	351

List of Figures

Figure 1: N-Person Social Convention Model of FGM.....	41
Figure 2: Threshold Model of Collective Behaviour: Normal Distribution of Thresholds.....	51
Figure 3: Threshold Model of Collective Behaviour: Uniform Distribution of Preferences	54
Figure 4: Threshold Model of Collective Behaviour: L-distributed Preferences	55
Figure 5: Threshold Model of Collective Behaviour: J-Distributed Preferences.....	56
Figure 6: Threshold Model of Collective Behaviour: U-Distributed Preferences	56
Figure 7: Causal Structure of FGM-related Decision-Making.....	112
Figure 8: Influence of Beliefs on Decisions about FGM (LPM)	120
Figure 9: Influence of Beliefs on Decisions about FGM (Logistic Model)	121
Figure 10: Persistence Dynamics of the Standard Coordination Model	135
Figure 11: Simulated Intervention Effects Under Simple Intervention Representation (Random Targeting).....	141
Figure 12: Simulated Intervention Effects Under Simple Intervention Representation (Targeting Those Most Reluctant to Abandon)	142
Figure 13: Simulated Intervention Effects Under Simple Intervention Representation (Targeting Those Most Willing to Abandon)	142
Figure 14: Simulated Intervention Effects Under Coalition-based Intervention (Random Targeting)	144
Figure 15: Coalition Formation Under Coalition-based Intervention (Random Targeting)	144
Figure 16: Simulated Intervention Effects Under Coalition-based Intervention (Targeting Reluctant Actors)	144
Figure 17: Coalition Formation Under Coalition-based Intervention (Targeting Reluctant Actors)	144
Figure 18: Simulated Intervention Effects Under Coalition-based Intervention (Targeting Willing Actors).....	144
Figure 19: Coalition Formation Under Coalition-based Intervention (Targeting Willing Actors)	144
Figure 20: Effects of Autonomy (α) on the Relation between Social Pressures and Social costs.....	150
Figure 21: Dynamics of Persistence with Explicit Norm Enforcement by a Subset of Actors	163
Figure 22: Dynamics of Persistence with a Social Reference Network	164
Figure 23: Dynamics of Persistence with Explicit Norm Enforcement by a Subset of Actors and a Social Reference Network	164
Figure 24: Simulation Output: Social Reference Network with Norm-Enforcement.....	165
Figure 25: Dynamics of Stability with a Uniform Distribution of Preferences	166
Figure 26: Dynamics of Stability with a Uniform Distribution of Preferences and Heterogeneity of Autonomy	166
Figure 27: Effect of Heterogeneous Autonomy on Threshold Distributions (PDF).....	167
Figure 28: Effect of Heterogeneous Autonomy on Threshold Distributions (CDF)	167
Figure 29: The Parameter Space of the Beta-Distribution	178
Figure 30: Beta-Distribution Density: ‘Normal’ Distribution	179
Figure 31: Beta-Distribution CDF: ‘Normal’ Distribution	179
Figure 32: Beta-Distribution Density: ‘Uniform’ Distribution.....	179
Figure 33: Beta-Distribution CDF: ‘Uniform’ Distribution.....	179
Figure 34: Beta-Distribution Density: ‘U-Modal’ Distribution	180
Figure 35: Beta-Distribution CDF: ‘U-Modal’ Distribution	180
Figure 36: Beta-Distribution Density: ‘J-Modal’ Distribution	180
Figure 37: Beta-Distribution CDF: ‘J-Modal’ Distribution.....	180
Figure 38: Beta-Distribution Density: ‘L-Modal’ Distribution.....	181
Figure 39: Beta-Distribution CDF: ‘L-Modal’ Distribution.....	181
Figure 40: qi Estimates Across Participants in the 2005 Senegalese DHS Survey	218

Figure 41: Marginal Empirical Distribution of μH Parameter	218
Figure 42: Marginal Empirical Distribution of $\sigma H2$ Parameter.....	219
Figure 43: Empirical Relationship Between μH and $\sigma H2$ in Communities in the 2005 Senegalese DHS Survey	220
Figure 44: Joint Relation: Avg. q_i of Others with a Different Ethnicity in Same Community by q_i of Self.....	221
Figure 45: Joint Relation: Avg. q_i of Others with the Same Ethnicity in Same Community, by q_i of Self.....	221
Figure 46: Bivariate Relation, Simulated Rates FGM vs Observed % Women Cut (by community) (ABM w. Random Initialisation)	227
Figure 47: Bivariate Relation, Simulated Rates FGM vs Observed % Mothers with cut daughters (by community) (ABM w. Random Initialisation)	227
Figure 48: Bivariate Relation, Simulated Rates FGM vs Observed % Women Cut (by community) (ABM w. Popularity-Based Initialisation)	227
Figure 49: Bivariate Relation, Simulated Rates FGM vs Observed % Mothers with cut daughters (by community) (ABM w. Popularity-Based Initialisation)	227
Figure 50: Marginal Distribution (Density) of % Women Cut (by Community)	230
Figure 51: Marginal Distribution (Density) of % Mothers Cutting (by Community).....	230
Figure 52: Simulated Marginal Rate FGM (ABM w. Random Initialisation)	232
Figure 53: Simulated Marginal Rate FGM (ABM w. Popularity-Based Initialisation)	232
Figure 54: Empirical Marginal Distribution % Women Cut.....	232
Figure 55: Empirical Marginal Distribution % Mother's Cutting.....	232
Figure 56: Empirical Marginal Distribution % Women Supporting FGM.....	232
Figure 57: Empirical Bivariate Relation: % Mothers Cutting by % Women Supporting FGM	234
Figure 58: Empirical Bivariate Relation: % Women Cut by % Women Supporting FGM.....	234
Figure 59: Simulated Bivariate Relation: % Agents Practicing FGM by % Agents Supporting FGM (Random Initialisation)	236
Figure 60: Simulated Bivariate Relation: % Agents Practicing FGM by % Agents Supporting FGM (Popularity-Based Initialisation)	236
Figure 61: Empirical Bivariate Relation: % Mothers Cutting by % Women Support	236
Figure 62: Empirical Bivariate Relation: % Women Cut by % Women Support.....	236
Figure 63: Bivariate Relation: % Support FGM by % Support FGM	236
Figure 64: Simulated Intervention with a Low Average Approval of FGM ($\mu H = 0.4$).....	247
Figure 65: Simulated Intervention with a High Average Approval of FGM ($\mu H = 0.8$).....	247
Figure 66: Simulated Intervention with Balanced Social Costs ($\delta 1=0$, $\delta 2=0$).....	249
Figure 67: Simulated Interventions with Unbalanced Costs ($\delta 1=1$, $\delta 2=-1$)	249
Appendices:	
Figure B2.1: Variable Importance Analysis by Region (logistic regression, key three belief variables colour-coded)	292
Figure B3.1: Variable Importance Analysis using Survey-Weighted Logistic Regression (key three belief variables colour-coded)	293
Figure E1.1: Demonstration of Recovering Latent Variables via Logit Regression	319
Figure F2.1: Simulated Intervention with Weak Educational Effect ($z_2 = 0.1$), Smaller Intervention Size ($z_1=0.35$) and a Minimum Probability that Opposers ($H_i < 0$) will Join the Initial Coalition of 0%.....	333
Figure F2.3: Simulated intervention with weak educational effect ($z_2 = 0.1$), smaller intervention size ($z_1=0.35$) and a minimum probability that opposers ($H_i < 0$) will join the initial coalition of 100%.....	334
Figure F2.4: Simulated Intervention with Weak Educational Effect ($z_2 = 0.1$), Larger Intervention Size ($z_1=0.70$) and a Minimum Probability that Opposers ($H_i < 0$) will Join the Initial Coalition of 100%.....	334
Figure F3.1: Simulated Intervention Targeting <u>Supporters</u> of FGM, with Weak Educational Strength ($z_2 = 0.25$)	336

Figure F3.2: Simulated Intervention Targeting <u>Opposers</u> of FGM, with Weak Educational Strength ($z_2 = 0.25$)	337
Figure F3.3: Simulated Intervention Targeting <u>Supporters</u> of FGM, with Strong Educational Strength ($z_2 = 0.75$)	337
Figure F3.4: Simulated Intervention Targeting <u>Opposers</u> of FGM, with Strong Educational Strength ($z_2 = 0.75$)	337
Figure F4.1: Simulated Interventions with High Localisation and Small Intervention Size ($z_1 = 0.1$)	340
Figure F4.2: Simulated Intervention with Low localisation and Small Intervention Size ($z_1 = 0.1$)	340
Figure F4.3: High Localisation with Larger Intervention Size ($z_1 = 0.4$)	341
Figure F4.4: Low Localisation with Larger Intervention Size ($z_1 = 0.4$)	341
Figure F4.5: Distribution of Expected Utility for FGM Practice at Simulation Setup (homogenous preferences $\sigma_H^2 = 0.0001$ and high network connectivity $\mu_{Rsocial} = 500$)	343
Figure F4.6: Distribution of Expected Utility for FGM Practice at Simulation Setup (homogenous preferences $\sigma_H^2 = 0.0001$ and low network connectivity $\mu_{Rsocial} = 20$)	344
Figure F4.8: Distribution of Expected utility for FGM Practice at Time 2 (heterogeneous preferences, low network connectivity, $\mu_{Rsocial} = 20$ and low homophily of preferences $\rho_{Hy} = 0.001$)	345
Figure F4.9: Distribution of Expected Utility for FGM Practice at Time 2 (heterogeneous preferences, low network connectivity, $\mu_{Rsocial} = 20$ and high homophily of preferences $\rho_{Hy} = 1$)	345
Figure F5.1: Simulated Intervention with High Social Hump ($\sigma_H^2 = 0.011$) and a Small Intervention Size ($z_1 = 0.2$)	347
Figure F5.2: Simulated Intervention with Low Social Hump ($\sigma_H^2 = 0.071$) and a Small Intervention Size ($z_1 = 0.2$)	348
Figure F5.3: Simulated Intervention with High Social Hump ($\sigma_H^2 = 0.011$) and a Larger Intervention Size ($z_1 = 0.35$)	348
Figure F5.4: Simulated Intervention with Low Social Hump ($\sigma_H^2 = 0.071$) and a Larger Intervention Size ($z_1 = 0.35$)	348
Figure F5.5: Simple Threshold Distribution: Positive correlation with H_i : Low ($\rho_{H\alpha} = 0.001$) and no social costs for Practicing FGM ($s_2 = 0$) and no actors preferring FGM intrinsically ($s_1 = 0$)	352
Figure F5.6: Complex Threshold Distribution: Positive Correlation with H_i : Low ($\rho_{H\alpha} = 0.001$) and No Social Costs for Practicing FGM ($s_2 = 0$) and no actors preferring FGM intrinsically ($s_1 = 0$)	352
Figure F5.7: Simple Threshold Distribution: Positive correlation with H_i : High ($\rho_{H\alpha} = 0.991$) and no social costs for Practicing FGM ($s_2 = 0$) and no actors preferring FGM intrinsically ($s_1 = 0$)	353
Figure F5.8: Complex Threshold Distribution: Positive correlation with H_i : High ($\rho_{H\alpha} = 0.991$) and no social costs for Practicing FGM ($s_2 = 0$) and no actors preferring FGM intrinsically ($s_1 = 0$)	353
Figure F5.9: Simple Threshold Distribution: Positive correlation with H_i : Low ($\rho_{H\alpha} = 0.001$) and symmetric social costs for Practicing FGM ($s_2 = 1$) and many actors preferring FGM intrinsically ($s_1 = 1$)	353
Figure F5.10: Complex Threshold Distribution: Positive correlation with H_i : Low ($\rho_{H\alpha} = 0.001$) and symmetric social costs for Practicing FGM ($s_2 = 1$) and many actors preferring FGM intrinsically ($s_1 = 1$)	353
Figure F5.11: Simple Threshold Distribution: Positive correlation with H_i : High ($\rho_{H\alpha} = 0.991$) and symmetric social costs for Practicing FGM ($s_2 = 1$) and many actors preferring FGM intrinsically ($s_1 = 1$)	353
Figure F5.12: Complex Threshold Distribution: Positive correlation with H_i : Low ($\rho_{H\alpha} = 0.991$) and symmetric social costs for Practicing FGM ($s_2 = 1$) and many actors preferring FGM intrinsically ($s_1 = 1$)	353
Figure F6.1: Simulated Intervention with 'Normally'-Distributed Authority ($\mu_w = 0.5, \sigma_w^2 = 0.031$), No Correlation with Approval of FGM ($\rho_{Hw} = 0.001$)	355
Figure F6.2: Simulated Intervention with 'Normally'-Distributed Authority ($\mu_w = 0.5, \sigma_w^2 = 0.031$), High Correlation with Approval of FGM ($\rho_{Hw} = 0.991$)	355
Figure F6.3: Simulated Intervention with High Variance of Authority ($\mu_w = 0.5, \sigma_w^2 = 0.111$) and High Correlation with Approval of FGM ($\rho_{Hw} = 0.991$)	356
Figure F6.4: L-Shaped Distribution of Authority ($\mu_w = 0.2, \sigma_w^2 = 0.111$), High Correlation with Approval of FGM ($\rho_{Hw} = 0.991$)	356

Figure F6.5: Simulated Intervention with 'Normal'-Distribution of Authority ($\mu_H = 0.5, \sigma_H^2 = 0.031$), High Correlation between Authority and Approval of FGM ($\rho_{H\alpha}$), and Preferential Targeting of Actors with the Highest Authority	357
Figure F7.1: Simulated Intervention with Explicit Norm Enforcement and a Balance of Willingness to Enforce	359
Figure F7.2: Simulated Intervention with Explicit Norm Enforcement, and Opposers Unwilling to Enforce Abandonment.....	359
Figure F7.3: Simulated Intervention Targeting Pro-FGM Enforcers and No Creation of Anti-FGM Enforcers	360
Figure F7.4: Simulated Intervention Targeting Actors at Random, with 50% of Coalition Participants becoming Pro-FGM Enforcers	361

List of Abbreviations

Agent-based Modelling (ABM)

Dependent Variable (DV)

Female Genital Mutilation (FGM)

Independent Variable (IV)

Linear Probability Model (LPM)

Multi-Criteria Model Validation (MCMV)

Office of the United Nations High Commissioner for Human Rights (OHCHR)

Ordinary Least Squares (OLS)

Pattern Orientated Modelling (POM)

United Nations (UN)

United Nations Children's Emergency Fund (UNICEF)

United Nations Fund for Population Activities (UNFPA)

World Health Organisation (WHO)

Early Outputs of the Thesis Project

There were a number of preliminary outputs of this thesis project. Early versions of some of the arguments and methods applied in the thesis appear in these early outputs (although no material has been copied directly into the thesis). See Appendix G for further details.

Chapter 1: Social Dynamics, Formal Modelling and the Widespread Harm of Female Genital Mutilation

Introduction

In this chapter, I discuss the social dynamics of Female Genital Mutilation (FGM), a harmful traditional practice present in much of the developing world (especially parts of Africa and Indonesia). I also discuss the use of formal models to understand these dynamics. I use the term ‘social dynamics’ to refer to the interactions and decision-making of members of a localised population (a.k.a. micro-processes), and the macro-level outcomes that arise from this. By ‘formal models’ of social dynamics, I mean the use of mathematical formalisms (e.g. equations) and computational formalisms (e.g. computer programs) to represent and reason about social dynamics. Game-theoretic modelling is the archetypal example of this sort of approach (c.f. Swedberg, 2001). Although agent-based social simulation is the primary tool used in this thesis (see Chapter 3).

I exclude statistical models (i.e. regression models) from the category of ‘formal models of social dynamics’. Researchers studying FGM using statistical analysis, have typically used generalised linear models or generalised linear mixed models, in combination with survey data. These methods are valuable. However, I argue in this chapter that they do not address certain questions that are of key interest when studying social dynamics.

Female Genital Mutilation (FGM) is a widespread traditional practice which harms the health and human rights of its victims. It represents an urgent and unsolved problem which has received considerable attention from international policy-makers. As I outline later in the chapter, theorising and policy development relating to FGM has been impacted noticeably by formal theoretical approaches. Development economists, social modellers involved in policy, or those already familiar with the range of existing approaches to theorising about FGM, will probably not find this remarkable. However, readers with other backgrounds might. Some consider the phenomenon of FGM, with its (typically) distant historical roots (though see, for example, Leonard, 1996, 2000), the heterogeneity of its cultural presentation across populations, its associations with mythology and symbolism (Boddy, 1982), and its important place in the cultural lives of many practicing families (Gruenbaum, 2001), to be the proper domain of the narrative anthropologist, rather than the modeller (with their putative tendency to over generalise).

Nevertheless, I argue in this chapter that *formal models are essential for reasoning about social dynamics* (i.e. the relation between micro-processes and macro-outcomes); a topic which cannot be ignored in theorising about FGM, or in efforts to promote its abandonment. That others share this view of the importance of understanding the dynamics of FGM is evident in the impact of certain theories and models noted in this chapter, especially the ‘social convention’ (a.k.a ‘social norm of coordination’) account of FGM.

Those who are familiar with modelling social processes will *also* recognise that this activity is fraught with challenges. In developing a formal model of a social system, the analyst is required to distil an open-ended complex phenomenon with innumerable features, into a simplified partial representation (the model) - typically without the aid of the kind of well-confirmed fundamental theories which guide model construction in the physical sciences (Weisberg, 2006). To act as a tool that can aid policy, such a model may be required to make predictions about the behaviour of social systems in relatively novel situations. Developing a model which is justified for such an application is recognised as a very hard problem (Edmonds, 2017: 43).

FGM is one of the fields of study in which formal models of social dynamics have been developed with the hope of assisting practitioners. Yet the most influential of these models, the social convention model, is heavily simplified. This fact, alongside the recent identification of problems created by this model’s simplifying assumptions (see Chapter 2), invites scrutiny of the way that formal models of FGM have been developed and been applied.

In this thesis, I offer an account of the challenges facing those wishing to model the social dynamics of FGM and assist practitioners. This account is organised around the concept of *model design uncertainty*. This is uncertainty about the design of formal models and their suitability for application in a policy context. I also develop and implement a modelling strategy which aims to help address such challenges. In this chapter, I introduce the substantive topic of FGM and discuss the role of theory and formal modelling of its dynamics. I also expand on the fundamental challenges facing modellers of social dynamics and outline the structure of the thesis.

Defining Female Genital Mutilation (FGM) and Adopting Appropriate Terminology

Female Genital Mutilation (FGM) is usually defined as the practice of cutting or otherwise physically altering female genitalia for non-medical reasons. OHCHR et al. (2008:1), representing an inter-agency consensus within the United Nations, offer the following definition:

“all procedures involving partial or total removal of the female external genitalia or other injury to the female genital organs for non-medical reasons”

Throughout this thesis, I subscribe to this definition and treat procedures which fall within it as instances of FGM. I focus on a developing-world context, especially Africa. I focus on Senegal in my own modelling efforts.

My use of the phrase ‘Female Genital Mutilation’ is deliberate but potentially contentious. In subscribing to this term, I am following the reasoning of international policy-makers. In 1991, the World Health Organisation (WHO) recommended that the term Female Genital Mutilation should be adopted by the United Nations (OHCHR et al., 2008: 22). This terminology is adopted in the 2008 inter-agency statement and represents a commonly accepted term among international policy-makers. An older name: “female circumcision” is the term (as translated) commonly used by those in FGM practicing populations. However, unlike the term ‘circumcision’, the term ‘mutilation’ has advantages for advocacy and does not imply a false equivalence with male circumcision. Arguably “a clear linguistic distinction” is needed between male circumcision and FGM, in order to stress the “gravity and harm of the act [of FGM]” (OHCHR et al., 2008: 22).

The extent of the procedure of FGM varies considerably. The World Health Organization (WHO), UNICEF and UNFPA have produced a typology of the different degrees of alteration and ablation of flesh which accompany FGM (OHCHR et al., 2008: 4).

The UN inter-agency statement classified FGM into four categories:

Type	Description (OHCHR et al., 2008: 4)
Type I:	“Partial or total removal of the clitoris and/or the prepuce (clitoridectomy).”
Type II:	“Partial or total removal of the clitoris and the labia minora, with or without excision of the labia majora (excision).”
Type III:	“Narrowing of the vaginal orifice with creation of a covering seal by cutting and appositioning the labia minora and/or the labia majora, with or without excision of the clitoris (infibulation)”
Type IV:	“All other harmful procedures to the female genitalia for non-medical purposes, for example, pricking, piercing, incising, scraping and cauterization.”

Around 90% of the currently living victims of FGM have been subjected to types I, II or IV mutilations. Around 10% of the victims of FGM have experienced Type III, the most severe form of FGM (OHCHR et al., 2008: 5).

Female Genital Mutilation as a Harm to Health

Since it is classified as the ablation of flesh without medical necessity, FGM is, in a certain sense, harmful by *definition*. However, the seriousness of FGM’s damage to the health of survivors may often extend well beyond a narrow description of the procedure itself.

There is a consensus among charities and campaigners (e.g. 28TooMany, 2016; FORWARD, 2016), international development agencies (OHCHR, 2008 – UN interagency statement) and health authorities (e.g. WHO, 2018; Royal College of Nursing, 2019) that FGM can lead to many types of physical and psychological harm. The World Health Organisation (2018) lists the immediate health risks of FGM as including haemorrhage, pain, shock, infection, and even death (due to bleeding or septicaemia). They also list a range of ongoing health risks related to child-birth, sexual functioning, psychological health and genital health. These potential harms include postpartum haemorrhage, pain during sexual intercourse, post-traumatic stress disorder (PTSD) and genital infection.

In the past, certain academics from within the medical anthropology tradition have questioned this consensus, and suggested that the harms of FGM may be exaggerated or lack an evidence-base (e.g. Obermeyer, 1999, but see Mackie’s, 2003, rebuttal). However, a recent systematic review published in the British Medical Journal (Berg et al., 2014) substantiates many existing concerns about the short-term and long-term health complications from FGM.

Female Genital Mutilation as a Violation of Internationally Recognised Human Rights

FGM is widely viewed as a violation of international human rights. It is beyond the scope of this chapter to reconstruct complicated past debates surrounding FGM and rights (c.f. Shell-Duncan, 2008 & Williams-Breault, 2018). However, the practice of FGM clearly contravenes a number of different authoritative principles of international human rights. These principles include the following articles, which have been directly associated with calls from UN agencies (OHCHR, 2008: 8) and others for the abandonment of the practice:

1. Article 25 of the Universal Declaration of Human Rights 1948 states that “everyone has the right to a standard of living adequate for the health and well-being of himself”. Yet FGM likely contravenes this right by violating the provisions of health and well-being (c.f. Shell-Duncan, 2008).
2. Article 3 of the Universal Declaration of Human Rights 1948 states that “everyone has the right to life, liberty and security of person”. Yet Williams-Breault (2018: 227) points out that in the context of FGM, women and girls “are not in full control of their lives, their liberty, or their bodies”.
3. As a form of violence which discriminates against women, FGM is often viewed as conflicting with the 1979 UN Convention on the Elimination of All Forms of Discrimination against Women (UNICEF, 2013)
4. FGM is also likely to be in conflict with the UN’s 1959 Declaration of the Rights of the Child. Principle 2 of the declaration states that every child has the right to: “to develop physically, mentally, morally, spiritually and socially in a healthy and normal manner and in conditions of freedom and dignity”. The majority of FGM survivors have been mutilated as infants or children (UNICEF, 2013).
5. As a harmful traditional practice, FGM is explicitly in contravention of Article 5 in ‘The Protocol to the African Charter on Human and Peoples’ Rights on the Rights of Women in Africa’, also called the Maputo Protocol, which has been signed by a union of African nations (GTZ, 2006).

In addition to these notable examples, the UN interagency statement (OHCHR, 2008: 31-32) lists a large number of international and regional human rights agreements which support the widely held view of FGM as a violation of fundamental human rights.

Much of this global consensus (i.e. associated with international principles and agreements that FGM contravenes), has been reflected in the enactment of domestic legislation criminalising FGM in countries where it is practiced. A recent report by the UK charity 28TooMany (2018b), in collaboration with legal professionals (and others), assessed the current legal status of FGM in 28 countries in Africa where the practice is widespread. They found that 22 countries (as of September 2018) had enacted legislation criminalising FGM, with 5 out of the remaining 6 countries having indicated an intention (or actually begun) to pass such legislation (UNICEF, 2013 reported similar information). Senegal (which is the focus of the empirical analysis in this thesis) is among those African countries that have criminalised FGM. It did so in 1999 (28TooMany, 2018: 2).

FGM is also widely criminalised in countries where the practice is not widespread. This indicates a condemnation of the practice by the legislatures of those countries and is intended to discourage the practice among immigrant diaspora originating in a practicing country. Included in this category are the UK¹, member states of the European Union², and other industrialised countries including Australia, Canada, New Zealand, and the United States (Center for Reproductive Rights, 2006, 2009).

FGM as a Prevalent and Growing Social Problem (Despite Opposition)

FGM is not a rare vestige of fading historical practices. It is a prevalent and *growing* (in absolute terms) social problem despite widespread opposition to the practice across the globe (see below).

UNICEF (2016) estimates that at least 200 million women and girls worldwide have been subjected to FGM. Studies of the prevalence of the practice have tended to focus on countries in Africa. However, FGM is an established tradition among populations in non-African countries, including Indonesia and India (UNICEF, 2016). Moreover, FGM occurs to some extent among the immigrant diaspora across the world whose origins are in FGM

¹ Outlawed by the Female Genital Mutilation Act 2003 in England, Wales and Northern Ireland, and by similar legislation in Scotland in 2005.

² At least nine members states have enacted specific criminal law provisions against FGM, and FGM is prosecutable under general criminal legislation in all members states (European Commission, 2013: 9).

practicing populations. This includes diasporic communities in Western Europe and North America (UNICEF, 2016).

There is active opposition to FGM from state actors across the globe. As noted previously (see above), opposition to FGM is widely codified at both the international and state levels, including widespread criminalisation of the practice. Furthermore, a number of countries where FGM is practiced have implemented national policies to eliminate it.

In Egypt, for example, (where FGM is both criminalised and close to universally practiced) a number of government (or government-sponsored) institutions have taken steps to discourage the practice. This includes efforts by the Ministry of Health and Population to prevent medicalised FGM³, national media campaigns discouraging FGM and the creation (and constitutional enshrinement) of a National Council for Childhood and Motherhood, which works on the elimination of FGM (Girl Generation Egypt). Ethiopia (a country where more than 20 million women and girls have been subjected to FGM, UNICEF, 2013), is another example. The Ethiopian government established a Women's Affairs Office in 2005, which included a focus on FGM. Later, in 2011, the government set it the goal of bringing levels of FGM down to 0.7% by 2014/2015, as part of its *Growth and Transformation Plan*.

Also, alongside (and in collaboration with) state actors, international development agencies have implemented intensive programs designed to reduce the practice of FGM in UN members states, including through the UNFPA-UNICEF Joint Program (UNICEF, 2013, UNICEF-UNFPA, 2014), which aimed to promote the abandonment of FGM in 17 countries where it is widespread.

Unfortunately, despite widespread opposition, FGM remains highly persistent, and there is even evidence of growth in the incidence of the practice in some areas. This highlights the *urgency* of the problem currently posed by FGM. A report produced by the UNFPA notes that: "in 2012, in 17 countries implementing intensive [interventions to end] FGM... it was performed on about 12 million girls aged 15-19." (UNFPA, 2015: 42). The authors forecast that if current trends continue, the *absolute* number of girls born between 2000 and 2005 forecast to be mutilated in 2020 will *rise* to 15 million (UNFPA, 2015: 42). This is related to *population growth outstripping the declining rate of practice* (UNICEF, 2016). Also, citing population growth, a more recent UNFPA report forecast that 50 million girls were at risk of FGM between 2018 and 2030 (UNFPA & UNICEF, 2019). Equally concerning is evidence of an

³ This is when FGM is performed by a medical professional, rather than a traditional birth attendant, or other traditional figure

increase in the rate (proportion mutilated) of FGM amongst girls aged 15-19 in at least 4 countries between 2007 and 2012 (UNFPA, 2015: 32). Historically, where the rate of FGM within countries has decreased across successive birth-cohorts over time (Koski and Haymann, 2017: 7), this change has ranged from very substantial (-26.7% between the 1965-1967 and 1995-1997 birth cohorts in Kenya) to trivially small (-0.9% between 1965-1967 and 1995-1997 birth cohorts in Mali).

The Role of Theories of the Social Dynamics of FGM

“Understanding the forces underpinning female genital mutilation/cutting (FGM/C) is a necessary first step to prevent the continuation of... [the] practice” (Berg and Denison, 2013: 837)

“Recent reviews of intervention efforts aimed at ending female genital cutting (FGC) have concluded that progress to date has been slow, and call for more efficient programs informed by theories on behavior change.” (Shell-Duncan et al., 2011: 1)

“Efforts to reduce the prevalence of female genital cutting have had mixed success, and, despite the long history of such efforts, the factors supporting or impeding cessation are not well understood.” (Hayford and Trinitapoli, 2011: 2)

“...we know that in order to bring about widespread change, initiatives must take into account the complex social dynamics surrounding FGM/C.” (UNICEF, 2010: vii)

A great deal of development work aimed at the encouragement of the abandonment of FGM has depended on working directly with individuals, families and communities practicing FGM to persuade them to abandon the procedure (Droy et al., 2018; Johansen et al., 2013). In some cases, this development work has drawn heavily on insights about social dynamics within FGM practicing communities (Efferson et al., 2015; UNICEF, 2007; UNICEF, 2010). Applications of such ideas, which are discussed further below, include the organisation of coordinated collective declarations, as well as the expectation of achieving collective ‘tipping-points’ in the abandonment of the practice. In fact, the notion of FGM as a harmful social convention (or social norm of coordination) has, at times, formed a core part of the theory of international strategies to eradicate the practice (Mackie, 2017).

The idea that FGM is a social convention, or, social-norm of coordination (now the preferred term), began in 1996 with Gerry Mackie who published an article in the *American Sociological Review* (one of the world’s premier sociology journals), titled: *Ending Footbinding and Infibulation: A Convention Account* (Mackie, 1996). The article dealt simultaneously with two phenomena: *footbinding* in China (from its ancient origins until its demise in the twentieth

century) and *infibulation* (a severe form of FGM concentrated in certain parts of Africa). The article considered both infibulation's (putative) origins in ancient empires and its persistence in contemporary populations. At the core of the article was the claim that both the phenomena of infibulation and footbinding share a common underlying social dynamic and that this dynamic can be understood in terms of a formal game-theoretic representation of a social convention arising from a coordination game. The formal model is analysed in later chapters; for now, a description is sufficient, as follows.

Within the framework of game-theory, a coordination game is a scenario in which players (here, actors in FGM practicing communities) face incentives (benefits, or reductions in costs, broadly understood) for practicing FGM if others do so, or for abandoning FGM if others abandon as well. Mackie (1996) claimed that in communities in which FGM is practiced, there is an established expectation that others will continue with the practice and that the practice is a pre-requisite for marriage⁴. He claimed that this motivates individuals to continue with the practice - because of the costs associated with unilaterally abandoning it (including an inability of a daughter to find a husband), based on an expectation that others will continue with the practice. Where this expectation has a strong historical precedent, and coordinated collective abandonment would be difficult (e.g. because there are many families involved and the practice is not discussed openly), the coordination game is said to represent a *self-enforcing convention*. It is self-enforcing in the sense that families continue with the practice on the expectation that others will, and this decision itself confirms and perpetuates the shared expectations which motivate the practice.

Using his game-theoretic representation, Mackie established a number of claims which have been significant in their impact on theory and policy surrounding FGM. He predicted that the conventional behaviour (i.e. FGM) could continue even in a scenario in which the majority of individuals in the relevant community privately thought the practice should end. This was predicted to occur because the decisions of members of the community were considered to be contingent on *social* costs and benefits of practicing FGM (e.g. in terms of marriage) rather than merely *intrinsic* costs and benefits (e.g. private attitudes to FGM). These social costs, in turn, depend on what others in the community are expected to do. Consistent with this notion, Mackie (1996: 1014) cited evidence that some individuals planned to mutilate their daughters despite believing the practice should be abolished. Mackie also

⁴ Although the idea that social pressures surrounding FGM are always about marriage has been subsequently relaxed (Mackie and LeJeune, 2009).

predicted that the prevalence of FGM would be driven (over time), by coordination incentives, to complete abandonment or to universal practice, within communities. Consistent with this prediction, Mackie (1996: 1010) argued that a pattern of ‘local universality’ (everyone practices FGM or no-one does) could be observed across communities in FGM practicing countries. Both of these empirical regularities have found some support in nationally representative survey data analysed since Mackie’s article was published (e.g. UNICEF, 2013).

Two major policy implications were derived from Mackie’s analysis. The first was that *collective action* was a necessary and efficient component of intervention efforts aimed at preventing FGM. Put simply, Mackie predicted that if a sufficient number of families in a practicing community were organised (in a publicly visible way) to simultaneously commit to abandoning the practice of FGM, then this abandonment would be stable and successful. Paraphrasing Mackie, if enough families could see that enough other families were abandoning the practice, then the perceived costs of miscoordination (i.e. abandoning FGM whilst others continue) could be alleviated enough to allow sustainable abandonment of the practice to occur.

We can contrast this with the dominant perspective regarding the abandonment of FGM at the time at which Mackie (1996: 1015) was writing. This was that the abandonment of FGM should be led by health education, i.e. providing information about the negative health effects of FGM. By contrast, Mackie’s formal model predicted that in many cases, increasing the perceived costs of FGM (i.e. in terms of health) would not *on its own* result in any meaningful reduction in the practice. Mackie argued that such ‘intrinsic’ costs of FGM are dwarfed by the social costs of unilateral abandonment.

The second major policy implication was related to what can be labelled *spillovers* (Efferson et al., 2015 – Supplement: 14) and what Mackie (1996: 1011) called the *tipping point*. This was the idea that if a sufficient number of families could be organised to simultaneously (and visibly) commit to abandoning the practice, then this commitment would ‘tip’ the remainder of the population (including those who were not involved in organised abandonment) into also abandoning the practice. This prediction comes straight from the formal properties of Mackie’s model and is based on the idea that if a sufficient number of families in the community commit to abandoning FGM, this will change the relevant social incentives for the remaining families (i.e. by making the abandonment of FGM less costly than the continuation of the practice, see Chapter 2).

Mackie (1996) had a considerable impact on subsequent theoretical and policy-related thinking about FGM. The *social convention* approach, or *social norm or coordination* approach, as it is now usually called (Efferson et al., 2015; Mackie, 2017), claims to help explain the success of the popular Tostan community development program in promoting the abandonment of FGM in villages across Senegal (Mackie, 2000: 253, though see critical discussion of this claim in Chapter 2). In 2007, the social convention (or social norm of coordination) approach was taken up by UNICEF in a publication which placed it as a core part of a proposed strategy to end FGM “in one generation” (UNICEF, 2007: 26). In 2008, it was endorsed as the “... ‘common approach’ by 11 United Nations agencies, 9 foreign ministries, the International Organization for Migration, the European Commission, and the World Bank” (Mackie, 2017: 3).

The convention approach was played an important role in the United Nations Population Fund’s (UNFPA) and United Nations Children’s Emergency Fund’s (UNICEF)’s joint program (Mackie, 2017: 3). This program began in 2008 (and continued at least until 2015). It aimed to promote the abandonment of FGM. An important part of its design principle was “strategically leveraging social dynamics in favour of abandonment [of FGM]” (p.vii), through a “social norms perspective”⁵. This perspective was said to be “at the core of the programme framework” (UNICEF-UNFPA, 2014: vii).

In 2009, the Secretary-General of the United Nations noted that “It is now widely acknowledged that it [FGM] functions as a self-enforcing social convention or social norm” (cited in Mackie, 2017: 3). Also, in 2009, UNICEF’s Innocenti Research Centre published a working-paper refining and reaffirming a version of the social convention approach (Mackie and LeJeune, 2009). UNICEF (2013) published an extensive review of a large number of nationally representative household surveys which collected data about FGM in the preceding decades. This review had the social norm approach as its theoretical framework, and its results were claimed to reaffirm the relevance of the approach.

Mackie (1996) remains one of the most highly cited academic articles on FGM (at the time of writing). The social norm/convention account has been referred to extensively in studies developing statistical models of FGM, typically as part of the motivation for model selection and for interpretation of the estimated parameters of such models (Freymeyer and Johnson,

⁵ The social norm of coordination account of FGM is treated as synonymous with the social convention account in most cases (Mackie, 2017: 2), and any philosophical distinction between a social norm and a social convention is said to be of little practical interest (Mackie, 2017: 10; Mackie, 2018).

2007; Hayford, 2005; Hayford and Trinitapoli, 2011; Kandala et al., 2009). It has also been a key part of the theoretical motivation for subsequent in-depth qualitative (Shell-Duncan et al., 2011) and other primary research (Efferson et al., 2015) in FGM practicing parts of Africa.

Formal Modelling and the Social Dynamics of FGM

As noted above, formal game-theoretic modelling has played an important role in the genesis and application of the social convention/social norm of coordination account of FGM. However, the theory itself is a collection of different elements, including the game-theoretic model (Mackie, 1996; UNICEF, 2007), various theoretical ‘narratives’ about social change (Mackie, 1996; Mackie and LeJeune, 2009; UNICEF, 2007), and empirical claims about decision-making (Mackie, 1996, 2017; Mackie and LeJeune, 2009; Mackie et al., 2015; UNICEF, 2007, 2013). This distinction is clear, for example, in Mackie et al. (2015: 28), which defines a social norm of coordination in terms of claims about interdependent decision-making.

Distinguishing between these different elements is important. As I illustrate in Chapter 2, failure to distinguish between social norm theory as a *theory of decision-making*, versus as the motivation for a *particular model* of social dynamics, can create conceptual confusion when attempting to evaluate or develop the theory⁶.

Mackie (1996) was the first of a number of studies that have used formal models in their analysis of the social dynamics of FGM. These subsequent studies (e.g. Chesnokova and Vaithianathan, 2010; Coyne and Coyne, 2014; Efferson et al., 2015 - Supplement, 2019; Novak, 2016; Platteau et al., 2017; Ross et al., 2016; UNICEF, 2007), which are discussed further in the next chapter, have generally followed Mackie’s approach of adopting a (broadly) rational choice perspective. However, they have used a variety of different model designs, some of which are closely related to Mackie’s coordination model, and some of which are more distinctive. These studies have in common that they use formal tools to investigate how decision-making and social interaction within FGM-practicing communities could lead to different macro-level outcomes, especially the persistence or abandonment of FGM within a population. These studies typically include attempts to assist practitioners by pointing to actions that might help them to eradicate the practice.

⁶ I also revisit these distinctions in Chapter 4, as part of an attempt to empirically test ‘core’ theories of FGM decision-making.

The principal rationale for using formal models is that they provide capacity for reasoning about the social dynamics that connect micro-processes (decisions, actions and interactions within communities) and macro-level outcomes. Importantly, formal models can facilitate this kind of reasoning in a way that other methods cannot. The main alternatives to formal modelling for theory construction available to theoreticians interested in FGM are data-driven variable-based approaches (i.e. statistical models, see Shoemaker et al., 2003 as an exemplar of this paradigm, and see Macy and Willer, 2002 for a contrast with a formal ‘dynamics’ perspective) and informal verbally expressed theories (i.e. narratives). Yet there are reasons to be sceptical that *in general* either of these approaches can reliably handle the relationship between micro-processes and macro-outcomes (Chattoe-Brown, 2013).

It is possible to construct simple and comprehensive verbal descriptions of hypothetical micro-processes whose macro-level dynamics are highly counter-intuitive (as shown through formal simulation). This is sometimes called the problem of *emergence* in social dynamics (but see Epstein, 2006: 31-38). A popular and illustrative example of this is Thomas Schelling’s model of residential segregation (Chattoe-Brown, 2013). This is a very simple model whose micro-processes can be described quickly and comprehensively, but whose macro-dynamics (strong residential segregation despite universally low levels of separatist preference) cannot reliably be anticipated without formal deductive tools: typically, a computer simulation. It will become clear to the reader in later chapters that analogous problems exist in trying to reason informally from descriptions of the social processes underlying FGM to the macro-level outcomes that these descriptions entail. An important example discussed further in Chapter 2, is the potential impact of “heterogeneity of preferences” on dynamics of social coordination. The existence of examples such as these, in which it is hard to reason verbally (i.e. informally) about the relationship between simple micro-processes and their macro-level outcomes, demonstrate that informal reasoning is not a *reliable*⁷ tool for such a task.

Variable-based data-driven statistical modelling typically faces something of a reversal of this problem. Although multiple generalised linear models, for example, can effectively make ‘visible’ (Goldthorpe, 2015: 13-14) macro-level regularities arising from social processes (such as associations between wealth and the practice of FGM after controlling for other demographic variables), reasoning back to the underlying micro-level social processes is much less straightforward. It may be common practice, for instance, to offer a causal or

⁷ It is important to note that this is not the same as the claim that informal reasoning about micro-macro relations is always wrong. Often it will be correct. But, it’s hard to be sure without a model!

narrative interpretation of the relationship between each of the statistically significant variables in a regression equation and the response variable (typically the decision of a parent to practice FGM, or not). Indeed, the model fitting will estimate a distinct ‘effect’ for each term in the regression equation. Yet, even if we take this process of interpreting casual coefficients as unproblematic, a social theoretician will recognise that social outcomes are not simply the result of an *accumulation of independent causal factors (or narratives)*, but rather the result of a complex process involving different interrelated features of social life, especially constant interaction between the ‘cases’ (e.g. individuals) from whom such models are estimated.

Quantitative analysts of FGM are well aware of this, of course, as shown by their close engagement with the social convention account of FGM (e.g. Hayford, 2005). However, social dynamics like those identified in the social convention model cannot be ‘read from’ the coefficients estimated in statistical models. The support (even if informal) of the formal theoretical model is required so that the statistical regularities can be interpreted and their compatibility with the theorised social dynamics can be assessed (See Goldthorpe’s, 2015: 13-14 discussion of the difference between making social regularities ‘visible’ and ‘transparent’). In the next chapter, for example, I point out that a number of statistical analyses support the idea that decisions about FGM are correlated within communities even after controlling for local socio-economic commonalities. This certainly implies that some form of positive social influence exists, yet analysts have relied on the social norm perspective to provide a model of *how* that positive social influence plays out at the micro-level⁸.

Unlike regression-based statistical analyses, formal models have the potential to address the complex aggregate consequences of multiple social actors making decisions and interacting over time. They do this by actually representing (in some form) these aspects of social processes as part of the model. Direct formal engagement with the complexities of social interaction has allowed modellers to derive impactful conclusions about FGM, relating to

⁸ Another distinctive use case (although not one which has seen much application to studying FGM) is the use of statistical analysis to model decision-making processes at the individual level. I use just such an approach in Chapter 4. However, while this is clearly an important contribution to the understanding of micro-level processes, problems run in the opposite direction. In this case, we lack a model of how these decision-making processes aggregate up to macro-outcomes. In either case, the regression method struggles to ‘bridge’ the different levels in a way which is informative about social dynamics.

dynamics such as positive spillovers, which are out of reach of either traditional statistical analysis, or informal narrative theory.

An important issue underlies the need for this consideration of the relative merits of different approaches to theorising about social dynamics. Reasoning about the social dynamics of FGM cannot be safely neglected from either a theoretical or policy perspective. The evidence discussed in Chapter 2 and elsewhere (Mackie, 2017; UNICEF, 2013) shows that FGM is a ‘relational’ phenomenon, in the sense that social dynamics play a central role in the perpetuation or abandonment of the practice. This is shown convincingly in both localised qualitative research (which highlights a range of substantive interactions and social inter-relations in the practice of FGM), and large-scale statistical research (which demonstrates local social context to be an important and robust source of variation in actors’ decisions and attitudes about FGM).

Furthermore, the social convention model of FGM, which represents the currently dominant understanding of the social dynamics of the practice, provides an example of how failing to consider such dynamics *could* be disastrous for efforts by policymakers to discourage the practice. In particular, the model suggests that under conditions of strong coordination, interventions focusing on attitude change (a common mode of operation for anti-FGM programs, Johansen et al., 2013: 2) may fail to generate *any* substantive or stable behavioural change (because of the social costs of unilateral abandonment).

Foundations of the Challenge Facing Modellers of Social Dynamics

As I’ve argued above, formal modelling is essential if we are to study the social dynamics of FGM. Yet, as I discuss here, researchers face fundamental challenges in building and applying formal models of social dynamics. These challenges, and their philosophical foundations have seen little discussion in the literature on FGM. Yet in chapters 2 and 3 I show that concepts introduced by such a discussion can help to clarify the status of existing models and theory, and can help to identify new methodological approaches which address existing challenges.

Real social processes are complex and multifaceted. As sociologists will recognize, ‘real’ social processes involve a constellation of actions and interactions, over time, within diverse populations of social actors, each of whom is unique and has multifaceted social relations with other social actors in the population (and so on). This general statement holds, of

course, for the persistence of FGM in the many thousands of territorial communities in which it occurs. With finite capacity, knowledge and resources, modellers typically have no hope of producing a complete, literal or comprehensive representation of social processes in their models. The open-ended challenge facing modellers then is that of producing a *limited* formal representation of social dynamics which is useful *despite* the fact that it necessarily omits, simplifies and abstracts a great many features of the ‘real’ social processes that it is being used to study.

Imperfections, abstractions, idealisations and so on are readily acknowledged as ubiquitous (and potentially problematic, see below) in the fields of microeconomics (a field from which a number of the formal models of FGM originate, Kuorikoski et al., 2010) and social simulation. Railsback and Grimm (2012) refer to the challenge facing modellers as that of identifying the *important* elements and processes in the social systems under study. They construe the fundamental challenge facing modellers as that of ensuring that a model can be considered a ‘structurally realistic’ (p.227) representation of the way that the system produces the dynamics of interest, despite being heavily simplified relative to the real system. More generically, we might say that a model needs to be *sufficiently similar* to its target, in ways relevant to the problem that it is being used to address (Weisberg, 2012: 135-156).

Yet, it is rarely possible to know in advance what all the important elements and processes of a social system are, or what the consequent distortion produced by omissions or simplifications in a model of that system will be (Edmonds and Moss, 2005; Kuorikoski et al., 2010; Railsback and Grimm, 2012). To paraphrase an argument found in Weisberg (2006: 730-731), in the physical sciences, modellers usually have a stock of well-confirmed background theory that they can rely on to estimate the consequences of omissions and simplifications in their models. In the social sciences, few well-confirmed background theories exist to reassure us about the consequences of our model design choices. Kuorikoski et al. (2010: 546) echo this argument, pointing out that “in economics there is no fundamental theory that tells the modeller which assumptions give cause for alarm and which do not and how one should go about making the models more realistic”. In their discussion of a similar kind of uncertainty in modelling ecological systems (a topic with an analogous kind of complexity to social systems), Haag and Kaupenjohann (2001) note that modellers are required to find some means of transforming an open system with innumerable parameters into a solvable closed system with a limited set of parameters; in so doing “[w]hat seems a meaningful parameter to one observer may be irrelevant to another” (p.51).

This situation is further complicated when there are multiple modelling efforts within a single domain. Modellers may begin with different ideas about how to represent the target social phenomenon, with the possible result being that they make different decisions about which aspects of the phenomenon to represent in detail, and which to omit, simplify or idealize. When, as is the case for models of FGM (see Chapter 2), interested observers (or indeed the modellers themselves) are then presented with alternative models with contradictory dynamics, they are left with a formidable intellectual challenge. This challenge is that to resolve these differences and make progress, they will need to disentangle the contributions of genuine empirical disagreements (which may be adjudicated empirically) and the contributions of *divergent idealisations* (and other ‘extra-empirical’ commitments - Weisberg, 2012: 155), to these discrepancies.

These challenges are highlighted not only by an abstract consideration of the methodological issues involved but also by evidence from applied methodological research (particularly in the field of social simulation where there is less ‘orthodoxy’ about how formal models should be designed). This evidence comes from studies indicating the *fragility* of models of social dynamics to particular model design choices. Fragility, as the term is used here, can be understood as the *sensitivity* of key model outputs to assumptions which are *uncertain* (Leamer, 1983; Saltelli et al., 2008). Where sensitivity of this kind exists, it implies that the conclusions and insights derived from a model may depend on aspects of the model design which we have only limited confidence are accurate, correct or appropriate (etc.) (Kuorikoski et al., 2010).

In, for example, a re-implementation of Robert Axelrod’s well-known formal model of social norm (of cooperation) evolution, Galán and Izquierdo (2005) demonstrated that the influential conclusions of Axelrod’s analysis were profoundly fragile. Modifying arbitrary assumptions and parameters, or even running the model for a longer period, was enough to produce fundamentally contradictory dynamics. Analyses performed by Muelder and Filatova (2018) and Schindler (2013) reveal similar problems. Muelder and Filatova (2018) took a single social science theory and implemented it in a range of different model specifications. Each specification was broadly plausible. However, the authors systematically varied particular aspects of the design. Areas of variation included: specific equations, the decision models of actors, and the distribution-function of stochastic parameters in the model. The authors showed that these relatively arbitrary variations in model design, although individually defensible, resulted in qualitative and quantitative differences in predicted dynamics. Schindler (2013) performed a similar exercise. The authors took a single

‘family’ of models of land-use based on a single model design strategy. They showed that this single strategy could accommodate a range of broadly plausible model designs. Moreover, these different models predicted substantively different kinds of land-use dynamic.

These issues of *model design uncertainty* should be of considerable concern to researchers interested in studying the dynamics of FGM, especially if they wish to assist practitioners. The issues raise questions about how formal models can be established as adequate policy tools (Aodha and Edmonds, 2017; Edmonds and Aodha, 2019).

In this thesis, I organise these issues around the concept of *model design uncertainty*. The need for researchers, who have limited knowledge to begin with, to rely on a model which is partial, idealised, and so on, creates uncertainty about the adequacy of the model and its design. It makes it hard to know whether a model should be used as a policy tool, and it makes it hard to compare different models to assess progress in the field.

In many cases (see Chapter 2) modellers have sought to assist practitioners by making model-based recommendations about the kinds of policy-actions and interventions that might end FGM. We can characterise this kind of application of formal models as part of the ‘engineering’ paradigm (Edmonds and Aodha, 2019). It is within this ‘paradigm’ that *model design uncertainty* becomes most concerning. In some physical science scenarios, formal models can be expected to provide reliable predictions that fulfil an engineering function. In social science scenarios, the fundamental *model uncertainty* noted here calls into question the notion that formal models can be expected to make precise or reliable predictions about how policy actions will affect the social systems that they target (Aodha and Edmonds, 2017).

There are clear signs that these issues have created problems for attempts to study the social dynamics of FGM through formal models. The social convention model of FGM discussed above has been subjected to sustained criticism in recent studies (Efferson et al., 2015; Novak, 2016; Platteau et al., 2017). As I discuss in further detail in Chapter 2, these critiques can be usefully understood as being about the presence of a particular *idealisation* in the design of the model. More importantly, I show that this issue of idealisation (which I use as an umbrella term for simplifications, abstractions, omission and so on) is general across existing models of FGM, and so represents an ongoing and open-ended challenge for the field. Recognising these issues helps to clarify the status of existing models and theory. Furthermore, as I show in Chapter 3, this recognition may help to direct the field toward methodological approaches that better address and manage model design uncertainty.

Conclusions

The social dynamics of FGM are an important topic for researchers and practitioners wishing to understand the practice and promote its abandonment. Ignoring social dynamics would leave crucial aspects of the phenomenon un-theorised. It would also risk leaving policy-makers with no way to anticipate how social dynamics could disrupt their efforts to create positive social change.

Formal models are an essential tool in the study of these dynamics. However, in building and deploying these models, researchers face fundamental challenges. The inevitably partial nature of formal representations of social systems creates significant uncertainty about the adequacy of particular model designs. Researchers understandably want to employ their models to assist practitioners. Yet, using models of unknown adequacy within a social engineering paradigm risks misleading or harming policy efforts (see examples in Chapter 2).

As I've emphasized in this chapter, FGM is an urgent human rights problem of global significance. It is also an area of policy in which formal models, and ideas derived from them, have had an undeniable impact. This more than justifies the critical examination and development of formal modelling in this context. That said, it is worth noting the wider relevance of any discussion of these issues. Burgeoning technological development in the social sciences has led to a growth in the application of modelling techniques to policy problems which are outside of the stereotypical concerns of modellers of complex systems (e.g. financial economics, transport systems, etc.). The potential advantages that formal modelling brings are easy to state. However, their limitations and potential harms – and the way these can be mitigated – are a more subtle issue, one that practitioners addressing problems like FGM or other social issues into which formal modelling is encroaching, may have had limited exposure to. Therefore, although the discussion in this thesis focuses on the phenomenon of FGM, it can also be seen as a particular case study of the use of formal modelling to study, and address, complex social problems.

The research project reported in this thesis has both an applied and methodological element. The important methodological aims of the project are to critically discuss the challenge of modelling the social dynamics of FGM, to identify methodological approaches which may help to address these challenges and to demonstrate their applicability to the topic. The applied aim of the project is to employ these methodological approaches as part of a coherent and fruitful modelling strategy. This serves the dual goals of contributing to the

task of modelling the social dynamics of FGM and contributing to discussions of the methodology needed to do this in a useful and responsible manner.

I outline the modelling strategy used in the applied component of the thesis in Chapter 3, with Chapters 4-8 detailing these modelling efforts (and auxiliary activities). In Chapter 2 (the next chapter), I critically discuss existing formal theoretical models of FGM. I examine a number of key models in the existing literature. I highlight problems of idealisation in existing model designs and the challenges this creates for adjudicating between them. Subsequently, I review the extent of empirical validation of existing models, highlighting the limited extent to which current efforts have been able to ‘confirm’, or discriminate between, particular model designs. Chapter 3 begins the constructive project of discussing methodological approaches that may help to deal with these issues.

Chapter 2: Models of the Social Dynamics of FGM, A Critical Discussion

Introduction

To the best of my knowledge, this chapter constitutes the first general review of formal modelling of the social dynamics of FGM of its kind. There has been no general discussion of the methodological issues facing the field⁹. Platteau et al. (2017) provided a detailed discussion of models of FGM related to coordination, but they did not address the connection to underlying philosophical and methodological issues (see Chapter 1). Moreover, they did not address *other* established formal models of FGM in any detail. Also, as I discuss further below, parts of their critique appear to conflate issues of model uncertainty with questions about the empirical content of Social Norm of Coordination theory.

In this chapter, I review formal modelling of FGM, focusing on the fundamental issues discussed in the previous chapter. I begin by outlining and discussing the game-theoretical social convention model of FGM (Mackie, 1996; UNICEF, 2007). This model has been the subject of critique in a number of recent studies. I argue that these critiques are best understood through the concept of *idealisation* and that this helps to clarify the current status of the model and associated theory.

Existing critiques of the social convention model have centred around a single simplifying assumption/idealisation (the assumption of *homogeneity of preferences*, see below). However, the problems created by this simplification clearly highlight the need to consider *general implications of simplifying assumptions and idealisations* (and related aspects of model development practice) for the use of formal models of FGM as policy tools. In other words, the fragility of the social convention account to *one of* its simplifying assumptions invites critical enquiry about *other* potentially problematic design features; in the social convention model and in other models.

As part of recent critiques of the social convention model (see below), there has been a discussion of empirical evidence which might support or undermine it. Yet there has been

⁹ The abstract issues themselves, such as idealisations in formal modelling, are discussed in a number of articles (see Chapter 1), but no general discussion of the implications of these ideas for the field of modelling FGM has taken place.

no general review of the extent of empirical validation of the design of existing models of FGM.

Empirical validation is an important activity that might be hoped to allay concerns about the adequacy of the design of existing models of FGM. The term *empirical validation* typically refers to testing the predictions of a model about social dynamics (e.g. predicted rate of a behaviour over time) against empirical data germane to those predictions (e.g. real rates of a behaviour in some target population, Chattoe-Brown, 2014; Polhill and Salt, 2017, Railsback and Grimm, 2012) – although I use the term more generously in this chapter to refer to any deployment of empirical evidence to support the ‘correctness’ of a model design.

In principle, validation provides a demonstration of the instrumental value of a model (e.g. by showing that it reproduces some observable regularity) and may imply that it is, in some sense, ‘good enough’. Yet, in practice, concerns about the trustworthiness of a model may remain unless empirical validation is substantial and highly relevant to the purpose of the model. If, for example, uncertainty about a model design is relatively high (Polhill and Salt, 2017), and the model is validated on relatively weak empirical patterns (especially those that are only partially related to the intended function of the model, c.f. Railsback and Grimm, 2012, on the selection of relevant empirical tests) then the extent to which this ‘confirms’ the suitability of the model for application in novel settings is open to question.

Issues related to the adequacy of empirical validation can be understood more formally through the concept of *equifinality* (Evans et al., 2017; Poile and Safayeni, 2016). Equifinality occurs when multiple otherwise contradictory models are able (in principle) to reproduce the same empirical patterns - leaving the ‘correct’ model design *under-determined* by the validation data. Unfortunately, this is a common problem in modelling social dynamics (Polhill and Salt, 2017). Formal models of social dynamics, given their commitment to representing complex social processes directly, often have a large number of ‘free parameters’ that may allow them to erroneously fit data. The open-ended nature of the modelling process also provides considerable flexibility to modellers (i.e. in the way they construct and refine their model designs, Poile and Safayeni, 2016). Such a situation may allow otherwise erroneous models to reproduce pre-specified empirical patterns. A contributory dimension to this problem is that empirical data on social phenomena are often limited and subject to their own uncertainties (e.g. as part of data collection and interpretation).

In the final part of this chapter, I review the extent of empirical validation of existing models of FGM. Existing evidence provides support for some of the general motivating assumptions of existing models. However, empirical research has had limited success thus far in confirming the adequacy of any particular model designs, or in discriminating between alternative designs. This is due to the relative weakness of macro-level patterns found in empirical research that have been said to support particular models. It is also due to the problem of idealisation itself, which creates ambiguity about how the ‘failure’ of empirical validation is to be interpreted.

The Social Convention Model of FGM

The formal social convention model of FGM is outlined in Mackie (1996) and UNICEF (2007). The social convention model is the paradigmatic formal model of the dynamics of FGM. It has had the most substantive impact on policy and theory. As I discuss further below, the model has also been subject to sustained critique.

I focus primarily on the model as presented in UNICEF (2007) since this provides the more elaborate and up-to-date version of the specification. The model comes in two forms: a *matrix-form* and *n-person* versions. These versions use different analytical techniques from game theory to model social dynamics. The matrix-form model considers only two representative actors, who are the players in the ‘game’. The *n-person* model applies to two or more players, and it can be considered an elaboration of the *matrix-form* model, with additional dynamics. It is useful to start with the matrix-form model as a simple version of the n-person model.

The foundations of both models are ‘utility functions’ for individual social actors. These functions literally give the total costs and benefits to individual players as a function of their decision to practice FGM or not (a.k.a their ‘payoff’). Crucially, these utility functions are interdependent, which means that the ‘payoff’ to each actor for deciding to practice or abandon FGM also depends on the choices of other actors. These payoffs are made up of different kinds of costs and benefits, including social costs for abandoning FGM and ‘intrinsic’ costs (see below) for practicing it. These are different ‘components’ of actors’ utility functions, and they are represented algebraically (i.e. using variables).

In UNICEF (2007), the n-person model relies partly on a graphical device called a Schelling coordination diagram. Also, unlike some more recent modelling of FGM (e.g. Platteau et al,

2017) neither model explicitly defines a continuous utility function for individual actors. Instead, both models rely on *ranking* the magnitude of different costs and benefits for actors.

However, the continuous utility functions in both models are easy to deduce from the Schelling diagram and description, and I present them here to strengthen the exposition of the models and to establish areas of continuity (and discontinuity) with subsequent modelling efforts which do make these functions explicit.

The matrix-form game-theoretic model of FGM in UNICEF (2007) assumes two actors, both of whom are families in FGM practicing communities. These families are assumed to make decisions on behalf of their daughters. The families act as unitary agents (as if they were one person) and can decide whether their daughters are cut or not. Both families must decide whether to practice FGM or not (their only available actions). The decision of both families depends on their assessment of the expected utility for practicing FGM or not. The expected utility here is simply whatever subjective costs and benefits families expect for practicing FGM, and is defined by their utility functions. Families are explicitly assumed to be (subjective) utility maximisers. They will make whatever decision they expect to give them the highest utility (payoff).

The utility that each family receives from practicing FGM depends *partly* on intrinsic costs associated with FGM. Since practicing FGM harms the health and human-rights of daughters, families are assumed to factor this cost into their decision-making. Moreover, this cost is intrinsic to the practice of FGM and doesn't depend on what others are doing. This is represented in the model by each family expecting a fixed cost for practicing FGM (denoted here as H).

Each family also faces costs and benefits which *depend on what the other family decides to do*. If one family decides to cut, whilst the other family does not, the non-cutting family will face a social cost (denoted M), associated with loss of social status and/or marriageability. The other (non-practicing) family receives a benefit (denoted S), which represents the advantage that their daughter(s) will now have in the marriage market.

If neither family cuts, then neither are at a disadvantage, and neither pays a health cost. The costs and benefits for these different combinations of choices can be represented in a game-matrix (below). Each cell of the matrix (Table 1) represents the total costs and benefits that each player receives, given both players' choices (their utility, or 'payoff'). So, in the first cell,

both players receive $-H$ (row player's payoff is before the comma, column player's payoff is after the comma).

Table 1: Matrix Form Social Convention Model

Family 1 (row chooser), Family 2 (column chooser)	Cut	Don't Cut
Cut	-H, -H	S - H, -M
Don't Cut	-M, S - H	0, 0

The model assumes that the payoff $S - H$, which is the payoff for unilaterally practicing FGM, is less than 0, which is the payoff if neither player cuts. This means that if one family is not expected to cut, then the other family will also prefer not to cut. It is also assumed that $-M < S - H < 0$, so that if one family is cutting, then the other family will prefer to cut.

Even in this highly restricted form (relative to the *n-person* version below), a number of important applications of the model emerge. The model has two pure Nash-equilibria (situations from which neither family will wish to unilaterally depart). These equilibria are: both families cut, or both families do not cut. On the expectation that the other player will practice FGM, a given family prefers to practice it. This preference exists despite each family attributing a negative intrinsic utility to FGM.

FGM is therefore expected to be universal among coordinating families and to persist among (and be practiced by) families who would prefer that it ended. The key policy applications of this matrix-form game are related to these features. The game suggests that if families' expectations can be changed, such that they expect each other to not practice FGM (e.g. through a public declaration or another coordination device), then both players will prefer not to practice FGM and will move to a stable equilibrium of not cutting. Similarly, the game suggests conditions under which intervention approaches which *only* change preferences may fail to generate behavioural change. If, for instance, interventions only serve to make families aware that they are paying the cost: H (or to increase that perceived cost), then they will not produce behavioural change (except in the case that $H > M$, where the health costs are seen to outweigh costs to marriageability or social status, which is assumed to be unlikely).

The *n-person* elaboration of this game retains most of the features of the matrix game (including the dynamics responsible for its explanatory and policy applications), whilst adding others. The *n-person* version allows there to be a population of n identical families who

face similar costs and benefits to the families in the two-player game. The key difference is that these costs and benefits become a function of *the proportion of* families who make a given a choice (or are expected to do so). Essentially, the expected utilities faced by each of the n families match the pay-offs faced by the row-choosing family in the matrix game at the points where everyone else practices FGM, or no-one else does.

For instance, if a given family expects *everyone* else to practice FGM, then their expected-pay-off for practicing FGM is $-H$, and their pay-off for abandoning FGM is $-M$ (we will label this $-\hat{M}$ representing the *maximum* social cost for FGM practice). Similarly, if a family expects all other families to abandon FGM, then their expected pay-off for practicing FGM is $S - H$ (now referred to as $\hat{S} - H$ since this is the maximum social benefit for unilateral practice) and their expected pay-off for abandoning FGM is 0 . At levels of FGM practice ($p \in [0,1]$) *between* the abandonment of FGM by all other families and practice of FGM by all other families (i.e where $0 < p < 1$), pay-offs to each family are represented in the diagram as a linear function of the proportion of families making a choice to abandon FGM.

Let p represent the proportion of families practicing FGM at time t ($1 - p$ being the proportion abandoning), the expected pay-off to a given family at time $t + 1$, for practicing (or abandoning) FGM can be represented through the following utility-functions:

$$U(\text{practice}) = \hat{S} \cdot (1 - p) - H$$

$$U(\text{abandon}) = -\hat{M} \cdot p$$

Readers can examine these functions with values of p at 0 and 1 to see the resemblance to the matrix-form game. As mentioned, the dynamics of the n -player game are that we imagine family n deciding to practice FGM or not, based on the expected utilities for practicing FGM or not (which, in turn, depend on the choices of the other n -players). Since we retain the assumptions of the matrix game about the magnitude of the different payoffs (including that all families have identical utility functions), we know that if enough families have abandoned FGM (or are expected to), then family n will also abandon the practice (since $S - H < 0$).

The relationship between the expected utilities for practicing FGM and the number of families abandoning the practice (p), is visualized graphically in the UNICEF policy-paper. This helps with the exposition of families' decision-processes in the model. It also helps to make clear the dynamics of the model. I have produced a version of this graphic here (Figure

1), which assumes (in accordance with assumptions about the size of different costs and benefits), that $M = 10$, $S = 4$ and $H = 7$.

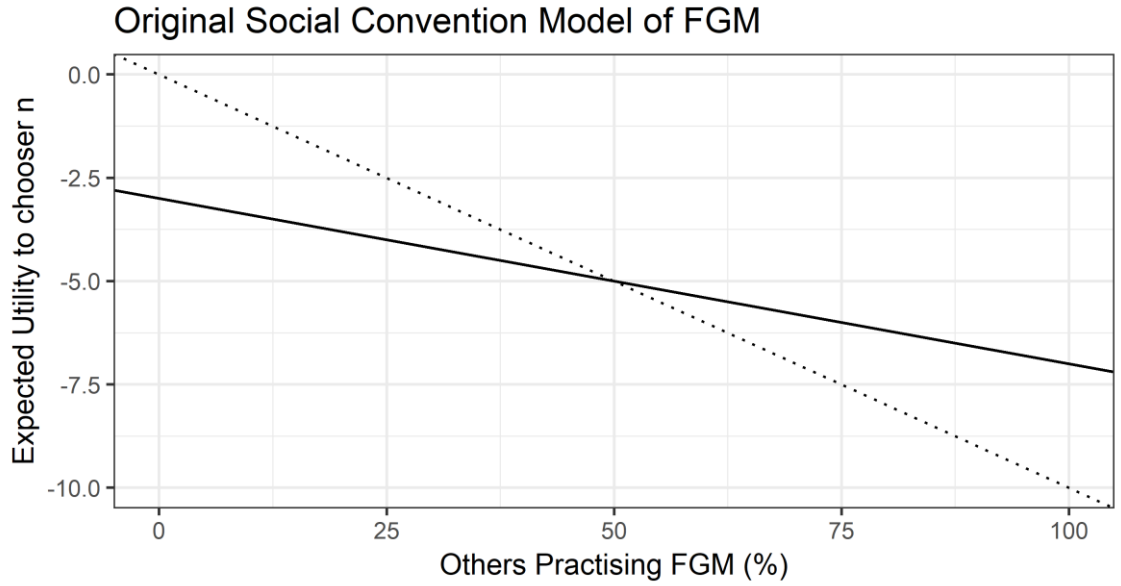


Figure 1: N-Person Social Convention Model of FGM (graphical form, the dotted line represents payoff for practicing FGM. The solid line represents payoff for abandoning FGM, and both represent the expected utility of the choice as a function of the % of others practicing FGM)

Note that there is a point (here 50%) at which the pay-off for abandoning FGM exceeds the pay-off for continuing with the practice. Since all actors are assumed to share identical utility-functions, we know that if the proportion of actors abandoning FGM (or expected to do so) exceeds this point then all actors will decide to abandon FGM. Conversely, until this point is reached, all actors will decide *not* to abandon FGM. This point is called the *tipping point*¹⁰. We can see that this model shares some of the dynamics of the matrix-form game: that there are only two stable equilibria in the population - everyone practices FGM, or no-one does. Indeed, the relevant population of n coordinating social actors will always be drawn to one of these equilibria according to the model.

Whilst this model does not lack (or contradict) any of the dynamics of the matrix-form game, it has additional dynamics that are absent from the matrix game. These dynamics become visible because the model is able to represent a large number of interacting families. The new dynamics centre around the tipping-point feature. Given the existence of a tipping point, the

¹⁰ UNICEF (2007) do discuss another dynamic arising from the social convention model, called the 'critical mass'. However, this is not a properly formalised feature of the model, and has largely been ignored in subsequent modelling. I discuss this feature, and the problems with its specification, in Appendix A1.

recognised implication of the model for policy, therefore, is that if development workers can alter the expectations of actors in the population (e.g. by helping/encouraging them to publicly commit to abandoning FGM - or to actually do so), such that they expect a number greater than the tipping point, of others, to abandon the practice (or if this many actors can be made to actually abandon the practice) - then *all* actors in the population will choose to abandon the practice and the population will reach the stable equilibrium of zero cutting (Efferson et al., 2015; Mackie, 1996; UNICEF, 2007). The key difference between this model and the matrix-form game, then, is that rather than abandonment needing to be coordinated among all the players, if the tipping point can be reached (i.e. involving a subset of the population), then this abandonment will spill over into the rest of the population (including those not involved in the coordinated abandonment itself). Naturally, the location of the tipping point (in this model) depends on the magnitude of the various costs and benefits, but it always exists somewhere between 0% and 100% if the relative magnitude of costs and benefits (e.g. $\hat{M} > H$) are maintained.

Tipping-points (a.k.a positive spillovers) are a key prediction of the model. Efferson et al. (2015) argued that such predictions have motivated considerable international development efforts that focus on the use of collective action – often in the form of public declarations, in which actors simultaneously and visibly commit to abandoning FGM, with the intention being that if a sufficient number of actors can be arranged to participate in these declarations, this will ‘tip’ the remaining population into abandonment and/or ensure stable abandonment among those families participating in the declarations. Mackie (2000)’s interpretation of the (now famous) Tostan NGO’s community education program in Senegal, was that its alleged (see below) success was related to a tipping-point dynamic of this kind.

Tostan facilitates community-led education programs in villages in the developing world (originally Senegal) involving a curriculum of educational seminars and community-building activities. One feature of these programs is that the participants, and others, can be encouraged or facilitated (at the end of the program) to organize (and take part in) public declarations, in which they publicly commit to no longer practice FGM (Diop et al., 2008). Mackie (2000) characterizes this as the village achieving a convention shift (i.e. tipping-point) in which sufficient numbers of individuals are seen to abandon FGM, such that this abandonment becomes stable and potentially spreads to other parts of the local community (e.g. non-participants).

A further important and distinctive policy implication introduced by the *n-person* convention model concerns the effectiveness of changing individual preferences as part of anti-FGM campaigns. In the matrix-form game, making families devalue FGM (e.g. through health and human-rights education) is not expected to produce any behavioural change (except in the unusual $H > M$ scenario). Indeed, it has no clear function beyond ensuring that families prefer the non-cutting equilibrium. Instead, the intervention needs to change families' *expectations* about what others are doing.

However, in the *n-person* game, we can see that increasing the perceived cost of FGM: H , can be useful even at values of H less than M . In terms of the coordination diagram, increasing H reduces the y -intercept of the utility function for practicing FGM, which, in turn, lowers the tipping-point. Therefore, an additional policy-relevant feature of the *n-person* model is that preference changes related to FGM are expected to reduce the number of actors who need to act collectively to bring about widespread changes in behaviour (UNICEF, 2007). However, it is important to note that preference change is considered something of an adjunct activity – it is not considered to be sufficient to generate change on its own in most cases.

In further comparisons with other formal models of FGM, I focus primarily on the *n-person game*. This is because the matrix-form game is really an interim step in the exposition and justification of the *n-person* game, which supersedes it. I refer to this (the *n-person* model in UNICEF, 2007) as the original social convention model of FGM from now on.

Critiques of the Social Convention Model and the Problem of Idealisation

The social norm of coordination theory of FGM, and the associated formal convention model outlined above has recently been subjected to sustained criticism and scrutiny (Efferson et al., 2015; Novak, 2016; Platteau et al., 2017). Efferson et al. (2015)'s empirical study of FGM appears to have been the impetus for these critiques. Efferson et al. (2015) conducted an empirical study in 45 villages in Gezira, Sudan – a country (and region) in which FGM is widely practiced. They measured the rate of FGM practice within these villages. The 'rate of FGM' was operationalised as the proportion of girls in each village who were cut before or during their summer vacation, prior to entering primary school. The researchers examined the *distribution* of rates of FGM across the 45 communities, as a test of whether FGM “exhibits the characteristics of a social norm based on coordination” (p.1446).

They argued that if FGM is a social norm of coordination, then there should be ‘discontinuity’ (p.1446) in the distribution of rates of FGM across villages, with the rate of FGM close to 100% or close to 0% in most villages. Instead, they found that most villages exhibited ‘interior’ rates of FGM. This means that, in most villages, the rate of FGM was *more than* 0% and *less than* 100%. In fact, their measured distribution of rates was roughly uniform. There was no clustering of rates near to 0% or 100%.

The significance of this finding is that it starkly contradicts the predictions of the game-theoretic social convention model discussed above. The social convention model predicts that only rates close to 0% or close to 100% will be stable. Any ‘interior’ rates should be driven down to 0% or up to 100%, depending on the location of the ‘tipping point’. Efferson et al. (2015: 1447) refer to this as a ‘signature’ of social norm of coordination theory.

The authors’ findings led them to suggest that coordination may not be an ‘appropriate theoretical framework’ (p.1446). Others have reached similar interpretations of the study’s findings stating that social convention theory of FGM has been ‘disproved’ (De Cao and Lutz, 2018: 3), that the theory that “FG[M] is upheld by social norms” has been “questioned” (Wahlberg, 2017: 16), that Efferson et al. (2015) have presented evidence not consistent with the ‘social-norm hypothesis’ (Poyker, 2016: 10), and so on. In these studies, and elsewhere (see below), Efferson et al.’s (2015) findings have been interpreted through the lens of scientific falsification.

The logic being employed is straightforward. If social norm of coordination theory predicts that rates of FGM should cluster around 0% and 100%, and Efferson et al. (2015) found that they don’t, then the theory is falsified. One of the main progenitors of social norm of coordination theory also appeared to accept this framing¹¹. Mackie (2017: 1) wrote that “The authors [Efferson et al.] seek to falsify an influential theory of the practice of female genital cutting that understands it as a social norm”. He then offered a detailed rebuttal of Efferson et al (2015). A major part of this rebuttal focused on potential flaws in their empirical findings¹². Efferson et al.’s (2015) study (which was published in *Science*) was widely publicised

¹¹ Mackie (2017:38) acknowledges parts of Efferson et al.’s (2015) supplementary materials as a critique of homogeneity of preferences assumption in the social convention theory, but doesn’t draw the fundamental connection between this and the status of social norm theory (see discussion below).

¹² Mackie’s (2017) critique is quite compelling. The most troubling element of the critique is the evidence presented that many girls who were uncut at the time of Efferson et al.’s (2015) measurement are likely to be cut later in their lives, thus undermining

and a similar interpretation to the above was also given in popular press (e.g. ‘*Female Genital Cutting is Based on Private Values Rather than Social Norms*’ Medical News Today, 25th September 2015).

While this framing of the situation appears quite straightforward, it actually ignores important aspects of the relationships between the social convention model, social norm theory and Efferson et al.’s (2015) empirical findings. By clarifying these relationships, one arrives at a different interpretation of Efferson et al.’s (2015) findings.

The neglected issues are the distinction between social norm of coordination theory and the game-theoretic social convention model, and the presence of idealisations in the design of the game-theoretic model. As I noted in Chapter 1, social norm of coordination theory is a collection of ideas. The game-theoretic model is part of this collection. However, the theory also includes elements which are separate from the model. Most important among these is that the theory makes substantive empirical claims about how actors make decisions. Specifically, the theory says that decisions about FGM are interdependent and coordinated. Actors will abandon FGM if sufficient others do so. They will also continue if sufficient others do so.

The game-theoretic model incorporates these ideas into its design and attempts to represent the social dynamics that might result from them. Inevitably, however, the model also includes a number of simplifying assumptions (i.e. idealisations). The significance of this cannot be overstated. Put simply; if the failure of the game-theoretic model to match empirical observations can be attributed to these simplifying assumptions *rather than problems with the underlying theory of decision-making* then this failure does not falsify the theory itself.

As it turns out, the prediction of the social convention model that ‘interior’ rates of FGM will be unstable, may just be an *artefact* of one of its simplifying assumptions. Specifically, the assumption that all actors attribute the same value to FGM (H in the model described above). This is called the ‘*homogeneity of preferences*’ assumption. When this assumption is relaxed (e.g. made more realistic by allowing H to vary between individuals, i.e. ‘*heterogeneity of preferences*’) the model *does* allow for stability of interior rates of FGM. In supplementary material for Efferson et al. (2015 – Supplement), as well as subsequent modelling studies which included criticism of social norm of coordination theory (Novak, 2016; Platteau et al., 2017) this

their claim to have measured rates of FGM practice in villages. I do not discuss Mackie’s rebuttal in detail since, as I argue, the correctness of Efferson et al.’s (2015) findings are essentially a moot point.

technical distinction is acknowledged. Yet the issue has still been communicated in a manner which implies theoretical falsification. Novak (2016: 2) stated that “empirical findings show that FG[M] is not a social coordination problem in Sudan” and Platteau et al. (2017: 33-34) stated that “there is little empirical evidence supporting a social norm story based on coordination incentives”.

Distinguishing clearly between potential problems with the underlying theory (social norm of coordination) and idealisations in the design of a particular formal model *based on* that theory (the n-person social convention model in UNICEF, 2007) provides important clarification. Efferson et al.’s (2015) empirical findings *don’t* show that FGM not a social norm of coordination (nor is Mackie’s, 2017, critique of their empirical findings required to defend the theory). Efferson et al.’s (2015) findings instead highlight that there may be a problem with one of the *idealisations* in the design of the social convention model. This is a problem with the model design, not with the substantive theory.

In turn, this suggests the need for a different kind of analysis. What is required is an assessment of how the *idealisation* (i.e. the ‘homogeneity of preferences’ assumption) in the social convention model may be distorting the dynamics of the model. This is a problem that requires model-based theoretical analysis rather than empirical analysis.

In fact, the model-based analyses included in Novak (2016) and Platteau et al. (2017) fulfil exactly this function. Novak’s (2016) ‘new theory’ based on ‘heterogeneous thresholds’ is identical to the social convention model once the homogeneity of preferences assumption is relaxed (see below). Both Novak (2016) and Platteau et al. (2017)¹³ presented analyses of essentially similar models, which should be understood as ‘de-idealised’ versions of the social convention model with heterogeneous preferences.

Recognising this doesn’t denude Novak’s (2016) or Platteau et al.’s (2017) analysis of its importance. But it suggests something important about the issue at stake in Efferson et al.’s (2015) findings. This issue is the presence of idealisations in the social convention model, and the potential for these idealisations to distort our understanding of coordination dynamics. The issue is *not* whether or not decision-making around FGM is coordinated. As I detail below, the predictions and policy implications of the social convention model can be shown to be profoundly sensitive to the assumption of homogeneity. As such, the social

¹³ Platteau et al. (2017) also discuss some other formal models not connected to coordination dynamics – although it is unclear whether they always intend these to apply to FGM.

convention model is strongly undermined by idealisations in its design. This, however, is a problem of model design uncertainty, not the falsification of the theory that FGM as a social norm of coordination.

In the following sections, I outline Novak’s (2016) ‘heterogeneous threshold’ model of FGM. I also show that the social convention model becomes a heterogeneous threshold model as soon as the assumption of homogeneity of preferences is relaxed. I then provide an overview of the dynamics of heterogeneous threshold models. Crucially, these are fundamentally different from the dynamics of the original social convention model. In particular, they support stable interior rates of FGM, and in some cases they do *not* support the existence of a tipping point that could be reached through coordination. Thus, the idealisation in the original model is shown to have potentially distorted its dynamics in significant ways.

Novak’s Heterogeneous Threshold Model of FGM

So, Novak’s (2016: 1) ‘new theory’ of the social dynamics of FGM is a heterogeneous threshold model. Heterogeneous threshold models themselves are an old idea that significantly predates formal theoretical analyses of FGM (Granovetter, 1978). However, I provide an overview of the construction of Novak’s model here as an illustration of how she arrives at a threshold model through assumptions about the dynamics of FGM. I adapt the notation somewhat for comparability.

One starts by defining a number of elements of the model. There is a *rate* of FGM in a local population of n households. This rate is the proportion of households practicing FGM. It is denoted p and is a real number in the interval $[0,1]$. Each household has an FGM status denoted g . g is 1 if the household practices FGM, 0 otherwise. Households choose between practicing FGM or not. The model is loosely premised on ‘rational choice’ assumptions. Each household has a utility function associated with each choice. Households make the choice that maximises expected utility. For each choice, the expected utilities are a function of p (the rate of FGM in the population). These utilities have a ‘social’ component denoted as follows:

$$s_{g=0_i}(p) \geq 0$$

$$s_{g=1_i}(p) \geq 0$$

These represent the social costs to actor i for not practicing $s_{g=0_i}(\cdot)$, or practicing $s_{g=1_i}(\cdot)$, FGM (respectively). In both cases, the cost is zero if all others in the community make the same choice, and the cost is assumed to be strictly increasing in the proportion of others who have made a *different* choice to the choosing household (e.g. because of social costs associated with marriageability and/or reduced social acceptance in the community, Novak, 2016: 6).

Novak assumes an intrinsic cost for practicing FGM ($H_i \geq 0$), with no intrinsic utility for abandoning it. This results in the following two utility functions associated with practicing or not practice FGM:

$$U_{g=1_i} = -s_{g=1_i}(p) - H_i$$

$$U_{g=0_i} = -s_{g=0_i}(p)$$

Households are assumed to make whichever choice has the greatest expected utility.

Novak's analysis is general with respect to the particular functional form of the relationship between p and $s_{g=1_i}(p)/s_{g=0_i}(p)$. She just assumes that $s_{g=1_i}(p)$ decreases with p , whilst $s_{g=0_i}(p)$ increases with p . Since one utility function is monotonically increasing in p whilst the other is monotonically decreasing, it follows that there exists a value of p at which households are indifferent between practicing FGM or not, with practicing preferred at any point above this value (and vice versa).

Following Novak's terminology, we will call this point p^* . p^* is called a 'threshold'. Novak assumes in general that the maximum social costs for abandoning FGM ($-s_{g=0_i}(1)$) is less than the cost for practicing it ($-H_i$), such that the social costs of unilateral abandonment will outweigh the intrinsic costs of FGM, and conversely, that $-s_{g=1_i}(1) - H_i$ is less than zero, such that the costs of unilaterally practicing FGM will outweigh the costs of abandoning it. This ensures that the threshold: p^* is between $p = 0$ and $p = 1$ for all households.

In other words, as the number of others in the community expected to practice FGM decreases to zero (e.g. through coordinated intervention efforts), all households will reach a point at which they too prefer to abandon the practice rather than continue (and vice-versa).

Novak assumes that the threshold p^* is heterogeneous across the population. Each household has its own threshold. This heterogeneity could arise either from differences in

the intrinsic costs that social actors attribute to FGM or from differences in their social cost functions. The dynamics of Novak's model depend on the *distribution* of thresholds in the population. These dynamics can be established with a graphical technique, which I outline below.

The key points of interest are the potential differences between the dynamics of heterogeneous threshold models, and the dynamics of the original social convention model (see above). These differences are of particular interest because we can arrive at a model which is equivalent to Novak's by de-idealising the social convention model, as I will now show.

First, consider that the tipping point in the original social convention model is itself a 'threshold'. It's simply a threshold that is identical for all actors in the population (Novak, 2016: 8). Formally it is the point at which:

$$U(\textit{practice}) = U(\textit{abandon})$$

Which occurs when:

$$\hat{S} \cdot (1 - p) - H = -\hat{M} \cdot p$$

It follows that the threshold (p^*) for *all* actors in the original social convention model is:

$$p^* = \frac{H - \hat{S}}{\hat{M} - \hat{S}}$$

Where H is the intrinsic cost that each social actor places on FGM. Under the original social convention model, this is a constant. However, now let us 'de-idealise' the assumption by allowing H to vary across the population such that actors' can have different views of the intrinsic benefits of FGM. This follows quite reasonably from no more than the acknowledgement (which is not contested) that actors vary in the extent to which they value (or devalue) the practice of FGM within a given community. One can replace H by H_i to acknowledge that H can vary between actors.

One can model variation in H_i using a random variable which has a 'realised' value for each actor in the population. In this case, we would treat H_i as a random variable with a minimum value of \hat{S} (the benefit of unilaterally practicing FGM) and a maximum value of \hat{M} , the maximum cost for unilaterally abandoning FGM. Intuitively, we can then define H_i as:

$$H_i = \hat{S} + q \cdot (\hat{M} - \hat{S})$$

Where q is a continuous random variable in the interval $[0,1]$. Substituting this back into the equation for the tipping-point, we find that:

$$p_i^* = \frac{S + q \cdot (M - S) - S}{M - S}$$

This reduces to¹⁴:

$$p_i^* = q$$

In other words, without needing other adjustments, an acknowledgement that the social actors vary in the intrinsic value they attribute to FGM, leads from the original convention model to a heterogeneous threshold model similar to the kind espoused in Novak (2016).

The key point here is conceptual *not* technical. The heterogeneous threshold model is *entirely compatible* with coordination theory. We arrive at such a model based on the same ‘substantive’ (Kuorikoski et al., 2010) assumptions about decision making that motivated the original social convention model. The only thing that needs to change is one of the model’s simplifying assumptions. The important thing about Novak’s (2016) and Platteau et al.’s (2017) analysis is that it shows that after this simplifying assumption is relaxed, the dynamics of the social convention model change – revealing the problematic influence of the simplification itself, as I will now describe.

Heterogeneous Threshold Models of Collective Behaviour

I start by providing an overview of how the dynamics of heterogeneous threshold models work, before showing how those dynamics can deviate dramatically from the original social convention model. I draw on Granovetter’s (1978), who provided the seminal introduction to models of this kind.

Granovetter’s (1978) canonical example is of a riot. Some actors may be willing to join a riot if at least **10%** of others (their threshold) in the crowd are doing so, but not otherwise. Some other actors may not be willing to join the riot until at least **90%** of the crowd is doing so. And so on. Under the threshold model, the *distribution* of thresholds in the population of potential rioters, along with which (and how many) actors are engaged in the behaviour at some initial point, will determine what proportion of the population ends up engaging in the

¹⁴ This also shows that if relevant payoff inequalities are maintained, the S term in the original social convention model is purely decorative, it has no effect on the formal dynamics of the model (since thresholds are purely a function of q)

practice. Granovetter (1978) presents a powerful graphical technique to help illustrate why this is the case, and to establish the social dynamics that will occur under different kinds of distributions

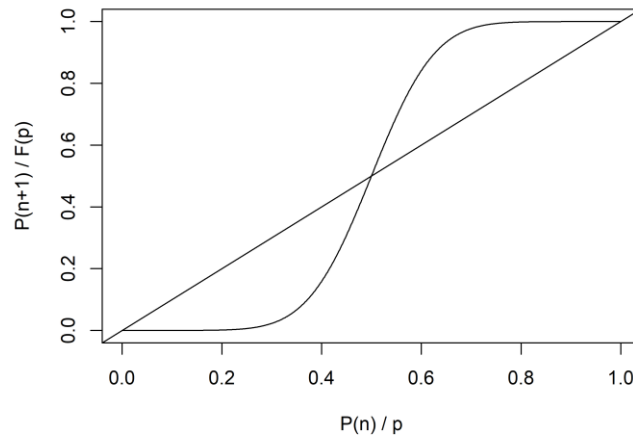


Figure 2: Threshold Model of Collective Behaviour: Normal Distribution of Thresholds (For the curved line: the x-axis represents thresholds; the y-axis represents the proportion of thresholds less than x in the local population. The straight line represents the identity function ($x=y$))

The technique relies on examining the *cumulative* distribution of thresholds in the population. Figure 2 presents an example of such a distribution, assuming thresholds are normally distributed around a mean of 0.5, with a standard deviation of 0.1.

There are two valid ways to think about the curved line on this graph. The first is that the x-axis represents threshold values p^* and the y-axis represents the function $F(p^*)$, where $F(p^*)$ represents the proportion of threshold values in the population which are less-than-or-equal-to p^* . The second is that the x-axis represents $P(n)$ which is the proportion of actors joining the riot (or practicing FGM, etc.) at time n , while the y-axis represents $P(n+1)$ which is the proportion of actors joining the riot at time $n+1$. This second interpretation allows us to establish, given a particular distribution of thresholds and particular starting point (on the x-axis) what the long-run outcome will be. We start with a value on the x-axis and move to the corresponding y-axis value, this is our prediction for the proportion of actors joining the riot in the next time step. This y-value becomes our next x-value, and so on (this is called ‘cobwebbing’).

Granovetter (1978: 1425) presents a clear explanation of why this relationship between the cumulative distribution of thresholds, and the dynamics of the threshold model exists. I paraphrase his explanation here. Imagine some arbitrary initial state for the population. Let’s

say 30% of the population is rioting, so $P(1) = 0.3$. What proportion of the population will be rioting at time 2? It will be the proportion of the population whose threshold is less than or equal to 0.3. To find this value, we simply look-up $F(0.3)$ on the graph (the y-value when the x-axis value is 0.3). Since $F(0.3)$ is below the diagonal, we know that $F(0.3) < 0.3$, so the number of actors rioting at time 2 will be less than 30%, and so on.

This approach to exploring threshold dynamics has some general features. Where $F(p)$ is below the diagonal, $P(n+1)$ will be lower than $P(n)$, so the proportion of rioters will decline over time. Where $F(p)$ is above the diagonal, the proportion of rioters will increase over time. Crucially, when $F(p) = P(n)$, a stable point has been reached.

We are, therefore, in a position to evaluate the dynamics of the population from the graph. Levels of participation in the riot below 0.5 will decline to the stable point of $F(t) = P(n) = 0$ and levels above 0.5 will increase until the stable point of $F(t) = P(n) = 1$ is reached. Interestingly, although the point $F(p) = P(n) = 0.5$ is also stable, this stability is relatively weak. In the case of stability at $P(n) = 0$, the population will be drawn back to this point even if the rate of participation in the riot fluctuates slightly. Conversely, when $P(n) = 0.5$, small fluctuations may *tip* the population toward rates of 1 or 0.

Homogeneity of Preferences as a Costly Idealisation

Novak (2016) and Platteau et al. (2017) analysed the implications of heterogeneity of preferences. Their analyses clearly demonstrate that these dynamics can diverge dramatically from those of the social convention model. Certain distributions of thresholds, such as a truncated normal distribution (see above) imply broadly similar dynamics to those of the convention model. However, in other distributions, dynamics such as *local universality* (interior rates will be driven to 0 or 1) and *positive tipping points* disappear. In some cases, heterogeneous threshold models predict stable interior rates of FGM, an absence of tipping points, and even negative tipping points (i.e. recidivism at almost all levels of abandonment). Recidivism would occur if a proportion of a large initial coalition of actors were willing to abandon FGM, but then ‘reverted’ back into the practice because of remaining social incentives. This can also happen under the original convention model if change is below the tipping point. However, Platteau et al. (2017) point to distributions of thresholds in which full recidivism is predicted to occur for almost any level of (temporary) positive behavioural change despite the population of actors all having individual thresholds in the 0%-100% interval. I illustrate a number of relevant scenarios, including this one, in this section.

We have already discussed the normal distribution function. We can see that under normally distributed preferences with a mean of 0.5 and a standard deviation of 0.1, the dynamics of the population can be similar to those of the original social convention model. The population is typically driven to the stable states of everyone practicing FGM, or no-one doing so. Also, there is a critical ‘tipping point’ in the model ($P(n) = 0.5$). If the rate of FGM in the population can be lowered beyond this point, then the rate of FGM would be expected to be driven to 0. This implies strong potential for positive spillovers, as predicted by the original social convention model. As such, this distribution is compatible with the policy and explanatory applications of the original social convention model. It continues to recommend, for instance, that coordinated abandonment efforts that can reach a central tipping point will be very effective, driving the rest of the population to full abandonment.

Now, let us consider a radically different sort of distribution, in particular, the one proposed by Efferson et al. (2015, Supplement - using the typo-corrected version – see Mackie, 2017: 37) as an example of a possible distribution of thresholds under which positive ‘tipping-points’ from coordinated abandonment are not expected to be possible:

$$i \in \{1 \dots N\}, \frac{i-1}{N} < p_i \leq \frac{i}{N}$$

Where i is the index for the i th actor in the population and p_i is their threshold ($p \in [0,1]$). A valid case of Efferson et al. (2015)’s distribution can be simplified to:

$$i \in \{1 \dots N\}, p = \frac{i}{N}$$

This is simply a discrete statement of the continuous uniform distribution. Recognition of this, in the context of the graphical approach discussed above, makes it clear why this distribution (Figure 3) implies a population in which *no tipping-points* will occur.

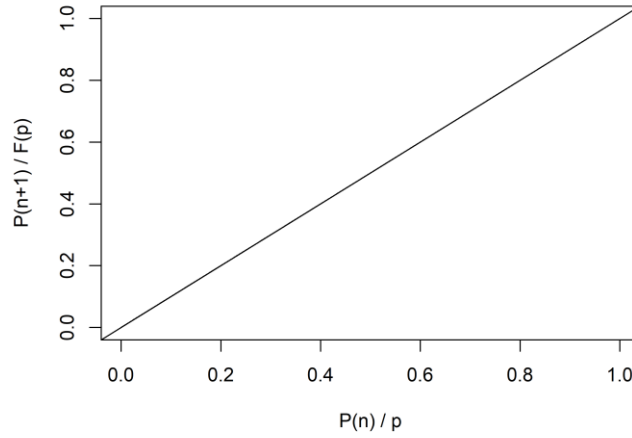


Figure 3: Threshold Model of Collective Behaviour: Uniform Distribution of Preferences

The cumulative distribution of the uniform probability density function *is* the line for which $P(n) = P(n + 1)$. Therefore, no positive spillovers can occur through coordination alone in this population because each rate of FGM is stable. Mackie (2017: 37) points out that a strict uniform distribution is arbitrary and small deviations from this distribution could allow for tipping. However, even in the more plausible scenario of stochastic variation around a uniform distribution in a finite population, spillovers still tend to be highly restricted (see Chapter 5).

Platteau et al. (2017) note kinds of distribution which have even more drastic implications for the policy predictions of the social convention model. One such case is the *L-shaped* distribution. This is a distribution where p_i is low for most actors, with a declining number of actors having moderate or high thresholds. Figure 4 illustrates such a distribution:

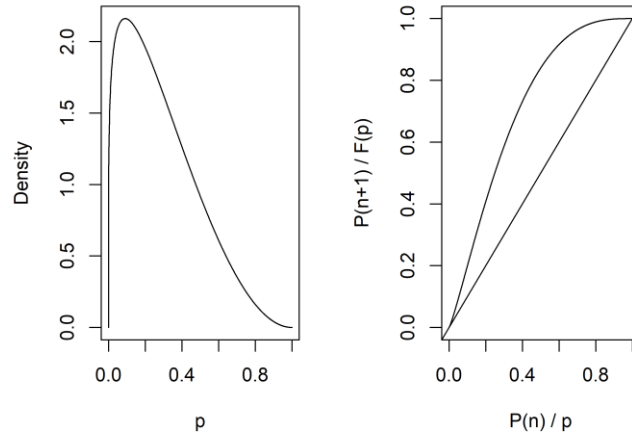


Figure 4: Threshold Model of Collective Behaviour: L-distributed Preferences

Notice that under this threshold distribution, the cumulative distribution of thresholds is above the diagonal line for almost all points above zero and below 1. We know, therefore, that rates of FGM in the population above zero are predicted to be driven to the equilibrium near to 1. This has the startling implication for policy that almost complete recidivism is predicted to occur under coordinated interventions that do not (successfully) include the entire local population.

Threshold models predict that if preferences in the population were distributed in this way (L-modal), then change would be extremely difficult to orchestrate through collective action alone - instead, extensive preference change would be required before stable abandonment could be brought about.

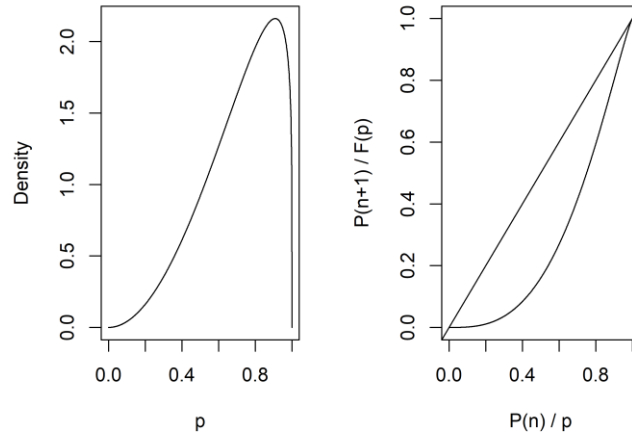


Figure 5: Threshold Model of Collective Behaviour: J-Distributed Preferences

A related distribution of interest is the J-shaped distribution (Figure 5). This distribution mirrors that of the L-shaped distribution and has opposite dynamics. In this (J-shaped distribution) scenario, rates of FGM below 1 will tend to be driven to the equilibrium near 0. Under this distribution, threshold analysis predicts that even small coordinated interventions in an FGM practicing community (assuming rates are currently stable at $P(1) \approx 1$), will generate large positive spillovers, driving the community to abandonment or near abandonment.

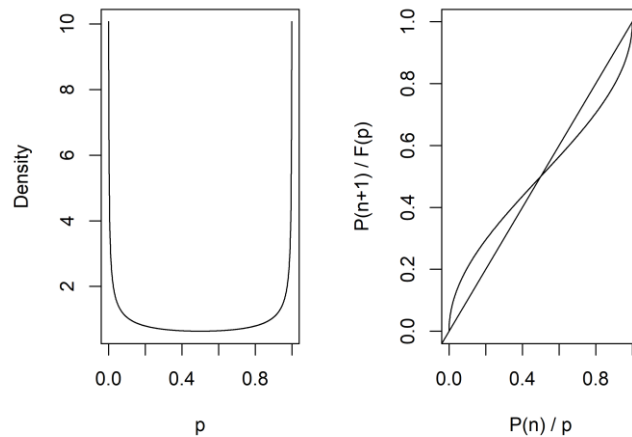


Figure 6: Threshold Model of Collective Behaviour: U-Distributed Preferences

Yet another distribution of considerable interest is the U-shaped distribution (Figure 6). This is a bi-modal distribution in which a large subset of actors attribute very low costs to FGM, and a large subset attribute very high costs to the practice, but few actors have intermediate

perceived costs. This distribution presents an interesting contrast to the normal distribution. Here we see that the point $P(n) = 0.5$ becomes stable in a robust sense, while rates close to 0 or 1 are liable to be driven to this central equilibrium.

This has important policy implications; it is a scenario where coordinated interventions to move a population away from near-universal prevalence are predicted to be successful but to lead to an intermediate rate of the practice. Getting beyond this intermediate rate to complete abandonment is expected to be much more difficult.

The above analyses show that the ‘homogeneity of preferences’ simplification in the social convention model is problematic. It leads the model to make important predictions that turn out to be highly fragile once the assumption is relaxed (i.e. de-idealised). This a problem of model uncertainty, not a problem of the underlying theory. UNICEF (2007) actually explicitly recognised that the assumption of homogeneity of preferences is false:

“Our diagram assumes that a typical chooser exists. In actual populations, one would expect to find a spectrum of attitudes, with a small number eager to abandon on the one side, a small number reluctant to abandon on the other side, and a larger number of typical choosers in the middle.” (UNICEF, 2007: 49)

Significantly, they appear not to have anticipated how problematic this assumption is. UNICEF (2007: 49) acknowledged that, in reality, there might be “stragglers” (those who value FGM more) who do not follow others once the tipping point is reached. However, the existence of a tipping point, which, if reached, will drive all-but these stragglers to complete abandonment, was treated as if it were relatively robust. As such, the idealisation was treated as relatively inconsequential for the policy and explanatory implications of the model. Close-to universal prevalence and close-to universal abandonment were still considered to be the only stable states in coordinating populations and reaching this point through collective action remained a core strategy recommended by the model. This is an instance of the fundamental challenges presented by model design uncertainty, as discussed in the previous chapter. This challenge is that it is often difficult to know a-priori the way in which particular idealisations (simplifications, omissions, and so on) might disrupt the insights gained from a formal model.

Efferson (2015), Novak (2016) and Platteau et al. (2017) all acknowledge that key differences in dynamics arise through the introduction of heterogeneity-of-preferences into the representation of social dynamics. However, much of the discussion have been organised around empirical adjudication of social norm of coordination theory and the presentation of

‘new theory’, rather than around the issue of idealisation in model design and the need to address model design uncertainty. There is enormous potential continuity between the social convention model and the models presented in Platteau et al. (2017) and Novak (2016). Differences between their designs are trivial in terms of social norm of coordination theory. In fact, arguably, Novak’s model (2016) is *more, not less*, like a ‘true’ representation of a social norm of coordination theory. Certainly, it is entirely compatible with the ‘core’ claims of the theory about decision-making.

The recognition of model design uncertainty as the core issue has important implications. It suggests that the empirical components of Efferson et al.’s (2015) analysis, and Mackie’s (2017) rebuttal of them, are somewhat moot. After all, there is no ‘empirical’ disagreement that the homogeneity of preferences assumption is an idealisation. Re-orientating the attention of the field from promulgating a ‘new’ theory to addressing questions of model design uncertainty has the potential to reconcile the apparently contradictory positions of Mackie (2017) and critics of social norm of coordination theory. In fact, it shows differences between their theoretical positions to be, in essence, illusory.

It also has important *progressive* implications for the field. Treating de-idealisations of the social convention model as ‘new theory’ runs the risk of prematurely ‘closing’ the problem. It suggests that a solution has been found which replaces the old theory with a new one. Yet a recognition that the social convention model has been undermined by idealisations in its design actually *opens* the problem. As I discuss further in a later section, it raises questions about *other* idealisations presented in the social convention model, and in other models in the field, including those ‘replacing’ the social convention model.

Understood through this lens, analyses by Novak (2016) and others highlight the significant value of systematically exploring the impact of ‘idealisations’ more widely in the field. I discuss this further in Chapter 3. In the following sections, I provide further illustration of the potential ‘costs’ associated with the idealisation in the social convention model, to further emphasise the practical importance of this issue.

The ‘Cost’ of Idealisation

De-idealisations of the social convention model show that its prediction that a ‘tipping point’ will always exist in coordinating populations may be an artefact of its simplifying assumption. This, in turn, raises important questions about policy ideas that have been inspired by the model. Novak (2016) has argued that policy-makers may wish to focus on individual-level interventions that aim to change preferences (i.e. rather than community-level coordination).

Efferson et al. (2015) and Platteau et al. (2017) make a similar point; furthermore, they present relatively sophisticated arguments about the implications of heterogeneity for targeting social actors within local communities. In the literature describing the original social convention model, UNICEF (2007) suggested that anti-FGM interventions will, and should, recruit those individuals in the community who are most eager to take part, in order to quickly establish the necessary support to reach the tipping point. In the context of the original model, the logic of this strategy is clear: those most willing to participate will be the easiest to recruit, and once recruited, will help development workers to reach the ‘tipping point’ of actors willing to abandon FGM (UNICEF, 2007: 49).

However, recognising heterogeneity of preferences means recognising that these ‘willing’ participants may differ from others in the ‘threshold’ at which they willing to abandon FGM. Recent formal analyses of scenarios of this kind (which I discuss further in Chapter 5) have suggested that interventions which deliberately (or by accident) recruit social actors who are the most willing to abandon FGM, actually risk *minimising* their chances of producing positive spillovers (Efferson et al., 2015 - Supplement, 2019). Instead, Efferson et al. (2015 - Supplement) argue that interventions which deliberately target actors who are in favour of FGM (changing their preferences and behaviour) may be better able to generate positive spillovers. By contrast, targeting those most willing to abandon FGM may fail to generate positive spillovers *even* under large interventions (under some distributions of preferences).

It is concerning that the problematic idealisations present in the social convention model could have resulted in policymakers being inadvertently misled, both about the general reliability of coordinated action as a tool to create widespread and stable abandonment of FGM, and about how this coordination should be arranged. Efferson et al. (2015) suggest that significant development funds have been allocated to programs that adopt a coordination-focused approach, a claim which is supported by publications produced by international policymakers. These documents make extensive reference to the social norm perspective and public declarations as important parts of anti-FGM program strategy (e.g. UNICEF, 2007; UNICEF, 2010; UNICEF-UNFPA, 2014).

An equally concerning, but more subtle, issue arising from the potentially misleading idealisation of coordination dynamics in the convention model is its potential impact on the *interpretation* of the outcomes of intervention efforts. In the late 1990s, the NGO Tostan was ascribed substantive success in promoting the abandonment of FGM after a group of villages participating in their ‘community empowerment program’ arranged a collective declaration

of the abandonment of FGM in a number of villages in Senegal. The interpretation of these declarations as a coordination device - reportedly resulting in the abandonment of FGM in the respective villages - led to this event being heralded as a validation of the predictions of the social convention model. It was argued that a local ‘convention shift’ – a tipping point - had been reached successfully (Mackie, 2000). Putting aside whether episodes of this kind *can* actually validate the social convention model (see discussion below), the combination of the putatively successful outcomes of Tostan and the interpretation of these outcomes within the social convention/social norm of coordination framework appears to have affected the way in which anti-FGM programs themselves are evaluated.

In evaluating the success of the wider efforts of Tostan, as well as those of UNICEF’s international program to encourage the abandonment of FGM (which was explicitly predicated on social convention/social convention theory), frequent reference has been made to the ‘achievement’ of public declarations as a result of intervention efforts (UNICEF-UNFPA, 2014: 4), treating these declarations as ‘an indicator and factor for social change’ (p.29), and implicitly interpreting this as an indication that a tipping point had been reached. Yet, the fragility of the social convention model’s predictions about tipping-points, calls into question whether a reliable correspondence: between public declarations (and other coordination devices) and widespread social change, can be safely assumed (see also Powell, 2017). In other words, large public declarations might have occurred, involving a greater part of the community, but if no potential for a ‘tipping-point’ exists, then this may not lead to widespread change in the remainder of the community.

In the aftermath of critiques of the social convention model, particularly those found in Efferson et al. (2015), studies have called into question the correspondence between public declarations and actual behavioural change. Available evidence suggests limited widespread abandonment of FGM in Kolda region of Senegal, an area where Tostan’s early activity was concentrated (Platteau et al., 2017: 6). Tostan’s activities have become widespread across Senegal (and elsewhere), prompting public declarations in thousands of communities. Yet, analysis of nationally representative surveys from Senegal show relatively weak evidence of changes in the national level of FGM - despite the considerable duration of the Tostan program (Kandala and Shell-Duncan, 2019).

Of course, these apparent discrepancies might have nothing to do with the social convention model or social norm of coordination theory. Interventions to end FGM are much more complex and multifaceted than their representation in the formal social convention model.

Indeed, the Tostan program is a holistic education and empowerment program, within which FGM is one among many issues (Cheikh Seydil Moctar Mbacke, 2018). As such, neither intervention success nor failure clearly justifies or condemns the model or the theory.

However, these episodes clearly highlight another *potential* danger arising from over-confidence in limited formal models of social dynamics: the distortion of the way in which empirical data about policy outcomes, such as reports that large public declarations have taken place, are interpreted. Aodha and Edmonds (2017) identify this as a common risk in the (mis) application of models of complex policy issues: over-confidence in the model can promote a myopic view of the data relevant to policy evaluation.

From One Problematic Idealisation to Many?

Analyses found in Efferson et al. (2015 – Supplement, 2019), Novak (2016) and Platteau et al. (2017) convincingly demonstrate the importance of representing actor heterogeneity in models of FGM. They show that the dynamics of coordination models can be sensitive to the idealisation of this feature - distorting their dynamics and recommendation for policy. Yet, it is not obvious from casual inspection of the original social convention model that this particular idealisation/simplification (i.e. homogeneity) is so problematic. As I illustrated above, discussions of the issue of heterogeneity in the context of the social convention model *prior* to the contemporary critiques of this assumption did not acknowledge the extent to which heterogeneity could upset the predictions of the model. The author(s) of the original model may not have been fully aware of the ramifications of this assumption.

In recognising this as a part of a fundamental problem of model design uncertainty, we arrive at a general implication for the field: *there is no reason to think that heterogeneity is the only problematic idealisation in the social convention model of FGM, or in other models of FGM*. Idealisation in models of the dynamics of FGM can thus be characterised as an *open problem* that requires ongoing investigation.

Idealisation and Diverging Model Designs

Some of the models of the social dynamics of FGM discussed in existing literature can be seen more-or-less as extensions or de-idealisations of the original social convention model. However, other models of FGM have been developed which depart more clearly from a social convention framework. Here I focus on Chesnokova and Vaithianathan (2010) and Coyne and Coyne (2014) because they present the most distinctive stand-alone models of FGM with their own dynamics and policy implications. I outline each of the models (the

latter in Appendix A2, due to space constraints) and discuss the way in which their designs and dynamics differ from the social convention account (and one another). In particular, I highlight the difficulties that *model design uncertainty* creates when trying to adjudicate between them.

Chesnokova and Vaithianathan's Model of FGM

Chesnokova and Vaithianathan (2010) conceptualized the dynamics of the practice of FGM explicitly in terms of the operation of a competitive market, in which FGM acts as an investment by parents. The (potential) return from this investment comes from securing an advantageous marriage. Chesnokova and Vaithianathan (2010) conceive of the marriage market as operating over two 'rounds'. In the first round, potential brides and grooms are matched at random. In the second, they are matched according to their value as spouses. For grooms, this value depends on wealth; for wives, it depends on their FGM status. All parties can decide whether to accept their first random match (if both agree) or to force both parties to go through to the second round, in which all parties have to accept their allotted partner.

Brides (or their families on their behalf) may 'choose' to be subjected to FGM as a costly investment, in the hope of achieving a better outcome in the marriage market. Grooms have no such prior investment to consider (wealth is treated as exogenous). Chesnokova and Vaithianathan (2010) take this to be a representation of the structure of the marriage market in FGM practicing communities. They ask, given such a structure, what strategies brides will be incentivised to adopt with respect to FGM¹⁵, and how they might be encouraged to adopt the strategy of *not cutting*. As such, their model is meant to both explain the persistence of FGM and to make recommendations for policies to promote its eradication. A full exposition of the model would be rather involved (depending on the enumeration and testing of mixed-strategy equilibria in a stochastic game). Instead, I focus on the elements of the model (i.e. its design and dynamics), rather than the formal derivation of those dynamics.

Of particular interest is the way that the model deviates in design and dynamics from the social convention model. The designs of the two models diverge in various ways. Whereas the *n-person* version of the social convention model is a game being played iteratively and globally amongst a community of coordinating actors, Chesnokova and Vaithianathan's (2010) model involves the stochastic (round 1) and deterministic/stochastic (round 2) *pairing*

¹⁵ These available 'strategies' with respect to FGM are to always practice FGM or not ('pure' strategies) or to practice with a given probability ('mixed' strategies).

of actors. The decision-making in the social convention model is that of a (repeated) binary choice between practicing FGM or not, based on the choices of the previous n players. In Chesnokova and Vaithianathan (2010), social actors have complete information, and they adopt a utility-maximizing mixed-strategy (i.e. a probabilistic strategy) involving decisions about cutting (Women) and acceptable marriage partners in round 1 (Men and Women). As in the original social convention model, actors associate FGM with a homogeneous fixed cost. However, actors are also differentiated by gender (men versus women), wealth (rich men or poor men) and cutting status (cut women and uncut women). In Chesnokova and Vaithianathan (2010) social incentives to practice FGM depend on the possibility of obtaining an advantageous spouse, which in turn depends on *both* the number of other cutting families *and* the availability of advantageous spouses in the local population. By contrast, in the social convention model, social utilities depend *only* on the proportion of other cutting families.

With these various differences in model design comes a divergence of predictions about social dynamics. Both Chesnokova and Vaithianathan's (2010) model and the original social convention model have in common that they predict that a global threshold may exist in the population and that if the rate of FGM can be driven below that point, then the population will tip into universal abandonment. However, they make fundamentally contradictory predictions about the stability of rates of FGM within communities. As noted, in the original social convention model, only the universal practice or abandonment of FGM within communities is expected to be stable. Under Chesnokova and Vaithianathan's (2010) model, universal abandonment is stable. However, universal practice may *not* be¹⁶, and a range of possible *interior* $([0,1))$ rates of FGM are potentially stable.

Adjudication between these two models presents a considerable intellectual challenge. The immediate thought might be to test their contradictory predictions (e.g. stable interior rates versus unstable interior rates) in order to discriminate between them 'instrumentally'. Unfortunately, this is unlikely to be as informative as one might hope. The problem is that the outputs of both models reflect a *combination* of empirical and 'extra-empirical' commitments (i.e. idealisations, omissions, simplifications, abstractions, etc.) by their authors (Weisberg, 2012). As such, comparing these outputs to data may not help us to discriminate between the important theoretical commitments of the two models. We know, for example,

¹⁶ Except in the implausible scenario in which FGM is costless

that the predictions of the original social convention model regarding the *necessary* instability of interior rates of FGM may be *artefacts* of the model's simplifying assumptions. Conversely, we do *not* know if the predictions of the Chesnokova and Vaithianathan's (2010) model are robust consequences of its core theoretical assumptions or artefacts of its own distinctive simplifying assumptions¹⁷. In both cases, this creates significant challenges in using tests of the models' outputs to adjudicate them.

A better strategy might be to try to empirically adjudicate the assumptions of the two models directly. One could examine, for instance, whether decisions about FGM are more strongly motivated by competition incentives (e.g. to get a better spouse than others) or coordination incentives (e.g. to avoid exclusion and ostracization from the marriage market, or local society) in real populations. This is exactly the strategy adopted in Chapter 4. However, it is important to recognize that this addresses only part of the problem. The presence of clear idealisation in both models means that such a test would only convey limited credibility to the *predictions* of either model since we do not know the extent to which these predictions are dependent on their substantive empirical assumptions, as opposed to idealisations in the model design.

As I emphasize in the next chapter, the uncertainties created by the inevitability of idealisations and simplifications in models of social dynamics mean that adjudicating between them is challenging. To do so convincingly, one needs to deploy a *combination* of empirical knowledge and contextual model-based knowledge (e.g. of the role that idealisations play in the model).

This comparison of Chesnokova and Vaithianathan (2010) with Mackie (1996) provides a 'worked example' of the challenges that model design uncertainty poses for adjudicating different models of the social dynamics of FGM (especially from different studies). Space constraints prevent this kind of analysis being repeated for all models of FGM (let alone all pairs). However, in Appendix A2 I present an exposition and commentary of another distinctive model of the dynamics of FGM: Coyne and Coyne's (2014) Identity Game Model, which has policy implications that diverge from, and may even contradict, the social convention model. In Appendix A3, I also provide a more general survey of the diversity

¹⁷ They are apparently *not* artefacts of certain aspects of the structure of their marriage market (Chesnokova and Vaithianathan, 2010: 22), but this far from covers the full set of simplifying assumptions involved.

present in existing models of FGM (in terms of model structure, decisions-contingencies, dynamics and derived policy implications).

Empirical Validation of Formal Models of FGM

Given concerns about uncertainty in existing model designs, questions about the extent to which the adequacy of these models has been demonstrated through rigorous engagement with empirical evidence become especially salient. In this section, I review attempts to actually validate formal models of FGM. Consistent with the previous discussion, I find that a certain degree of validation of important motivating assumptions (about the presence of positive social influence in FGM practicing communities) has occurred. However, I find that there is little evidence of adequate validation of any *specific* formal model designs. I present a survey of attempts to validate formal models of the dynamics of FGM identified in the literature in tabular form in Appendix A4. In this section I focus on discussing the key areas of validation (or alleged validation) highlighted by this survey.

With respect to the confirmation of motivating theoretical assumptions (which we can think of typically as substantive assumptions about decision-making), empirical research provides some support for current efforts. Research cited in a number of studies supports the contention that *positive social influence* exists in FGM practicing communities – meaning a tendency of actors to influence one another (in some way) to make similar choices about FGM. This is evidenced by a range of concerns expressed by actors in FGM practicing communities related to the necessity of FGM for marriageability and other negative social consequences (i.e. exclusion and criticism) for failing to practice it (Abdalla, 1982; Almroth et al., 2001; El Dareer, 1982; Gruenbaum, 2001; Hernlund and Shell-Duncan, 2007; Hicks, 1996; Lightfoot-Klein, 1989; Shell-Duncan et al., 2011; Shell-Duncan and Hernlund, 2006; UNICEF 2013). Similarly, a range of statistical analyses has found relatively robust associations between the decisions of individuals to practice FGM and local community characteristics (Hayford, 2005; Hayford and Trinitapoli, 2011; Kandala et al., 2009; Kandala and Shell-Duncan, 2019; Modrek and Liu, 2013). Furthermore, descriptive regularities in national survey data (UNICEF, 2013, Mackie, 2017) related to FGM, including

- Widespread practice by those who oppose the tradition
- Polarization of rates of practice between local communities
- Geographic and ethnic heterogeneity in rates of practice

all support the notion of positive social influence in FGM practicing populations.

The main contradictory finding arises from Bellemare et al.'s (2015) multi-level modelling analysis of FGM. The researchers found in their study that the *net* variation in the practice of FGM was relatively low at the local village level, relative to the household and individual level. This has been interpreted, by some, as evidence contrary to social norm of coordination theory (Platteau et al., 2017: 5). However, this may be a misinterpretation. Crucially, the dependent variable in Bellemare et al.'s (2015) analysis was the *approval* of FGM, rather than the actual decision of individuals to practice FGM. This distinction is crucial because the coordination paradigm allows for individual variation in attitudes to FGM. It is *decisions* which are predicted to face strong social coordination pressures (Mackie, 2017).

Nevertheless, despite empirical findings that are encouraging in their support of the notion of positive social influence in FGM practicing communities, the contribution of these findings to the justification and adjudication of *particular model designs* appears much more limited. The original social convention model is predicated on a particular notion of positive social influence (social coordination), but then so are identity-economics models of FGM (Coyne and Coyne, 2014), heterogeneous threshold models of FGM (Novak, 2016), models based on marriage competition (Chesnokova and Vaithianathan, 2010), and so on.

Other empirical observations treated as confirmatory of particular models face similar problems. Notably, the finding that FGM is either widespread or absent within practicing communities (so-called local universality) has been treated as confirmator of the social convention model (Mackie, 1996, 2017). Yet it is not clear that it actually has this implication. The de-idealisation of the assumption of homogeneity in the social convention model (discussed earlier in this chapter) shows that this prediction is fragile and doesn't necessarily depend on the model's substantive assumptions.

In other cases, the social convention model has either been challenged, or argued to be supported by, the outcomes of policy efforts. Mackie (2000) likened the outcomes of the Tostan intervention in Senegal (involving public declarations of abandonment) to his theoretical account. Conversely, Platteau et al. (2017) criticised the social convention model, partly on the basis of the claim that Tostan's interventions were not successful in substantively reducing the occurrence of FGM in a region where their activity has been concentrated.

Yet, there are reasons to be sceptical that either case has a significant bearing on the validity of the social convention model. Interventions to combat FGM have been complex and multifaceted, making it very difficult to attribute their success (or lack thereof) to individual intervention components (Droy et al., 2018). The Tostan intervention involves a range of other activities that might account for any success they achieved. Conversely, the failure to achieve abandonment through collective declarations *might* indicate some problem with the social convention model, or it might simply indicate that the tipping point was not reached successfully. As noted in Diop et al.'s (2008) long-term evaluation of the Tostan program, some participation in their public declarations was relatively transient (some participants only found out about the declaration the night before, p.18). Moreover, Berg and Denison (2012) and Johansen et al. (2013) note that anti-FGM interventions face a range of implementational issues (which are incidental to the social convention account), such as hostility to the interference of outside actors in village affairs.

Other notable 'confirmatory' regularities cited in favour of various formal models have included: associations between FGM and advantageous marriage outcomes (Chesnokova and Vaithianathan, 2010), associations between FGM and education (Ouedraogo and Koissy-Kpein, 2012), and the concentration of support for FGM among older age-groups (Coyne and Coyne, 2014). These regularities are certainly compatible with their respective models. However, it is not clear that they convey substantive confirmation of the models themselves. Chesnokova and Vaithianathan (2010), for example, do find evidence that women in Burkina Faso who are cut tend to marry earlier and tend to be richer. However, this does not show that advantageous marriage is the *main concern* which motivates parents to 'invest' in the practice (as their model assumes), nor does it tell us anything about the many other particular modelling assumptions that the authors make. Ouedraogo and Koissy-Kpein's (2012) and Coyne and Coyne's (2014) 'confirmatory' findings face entirely analogous problems.

It is useful to revisit the implications of Efferson et al.'s (2015) main empirical findings (rates of FGM in local communities) in the wider context of attempts to validate models of FGM. As I discussed earlier in this chapter, their findings have led to a recognition of problems in the design of the social convention model. Yet, this recognition is somewhat incidental. The homogeneity of preferences assumption was already known to be false. It is the model-based analysis that shows that this particular assumption is problematic. As I show in Chapter 5, relaxing the assumption of global interaction can *also* allow coordination models to display stability of interior rates of FGM (in situations where they otherwise wouldn't). Similarly,

Chesnokova and Vaithianathan's (2010) model *does* assume that all actors attribute the same cost to FGM. Yet, their model *also* predicts that interior rates of FGM can persist. The lesson here is simple but important. In the presence of idealisations (and other uncertainties about a model's design) empirical 'falsification' of a model's predictions shows us that there is a problem. But, it doesn't *in and of itself* tell us what the problem is.

Conclusions

This chapter represents the broadest review of established formal models of the social dynamics of FGM to-date and is also the first to address the fundamental issue of model design uncertainty facing the field. Model uncertainty represents a clear challenge for the field, with controversy surrounding the dominate model of the social dynamics of FGM: the social convention model, attributable to a problematic *idealisation* in its design.

Idealisations and the problems that these create also present difficulties for the adjudication of distinctive model designs. This is because divergences between the dynamics of existing models are attributable to an uncertain combination of their substantive empirical assumptions and subjective extra-empirical design choices made during their design. Idealisation also creates challenges for model validation, since validation ‘failure’ doesn’t necessarily establish which aspects of a model design are problematic.

Substantial empirical evidence related to the social dynamics of FGM has accumulated. However, the ability of this evidence to discriminate between particular model designs appears limited. There is considerable support for the basic motivating assumption of existing models, that strong positive social influence exists in FGM practicing communities. However, many of the macro-level patterns that have been associated with the social dynamics of FGM are potentially compatible with a range of model designs that include positive social influence. The field would likely benefit from the accumulation of more detailed macro-level data that might better discriminate between model designs. However, it is also likely that much of the uncertainty about existing (and future) models will need to be addressed at the design stage, through the further use of micro-level data and model-based exploration of design assumptions. Relying on macro-level data to discriminate between different model designs seems unlikely to be a viable strategy on its own.

Recognising the fundamental challenge of design uncertainty when modelling the dynamics of FGM helps to clarify the status of existing models and theories. It also sheds light on the practical challenges that the field faces, in particular, the need to address this uncertainty as part of model development and application. Recent controversies surrounding the social convention model highlight the potential costs of using uncertain models whose dynamics may depend on arbitrary aspects of their design.

In the next chapter I discuss methodological techniques from the field of social simulation which have not been widely or systematically applied in the study of FGM, but which may

help the field to meet the challenges of model design uncertainty. I propose a model development *strategy* which combines a number of these techniques and aims to capitalise on potential synergies between them.

The remaining chapters implement this strategy in order to develop and deploy an agent-based social simulation model of the social dynamics of FGM (see Chapter 3). Through implementing the strategy, I aimed to contribute to a number of the specific issues noted in this chapter, including

- a) Testing underlying ‘core’ theories of decision-making employed in models of the dynamics of FGM (independently of the model themselves)
- b) Exploring other simplifying assumptions in models of FGM dynamics
- c) Identifying detailed macro-level patterns to provide a stronger model validation test *and actually testing a model against these patterns*
- d) Developing an application for modelling which can contribute to FGM policy *despite* continued uncertainty about model adequacy.

Chapter 3: Managing Uncertainty in Models of the Social Dynamics of FGM, Methodological Techniques

Introduction

In this chapter, I discuss possible methodological responses to the problem of uncertainty in modelling the social dynamics of FGM (see Chapters 1 and 2). Model design uncertainty does not, of course, only occur for the problem of modelling the social dynamics of FGM. It is endemic in modelling complex systems (Saltelli et al., 2008; Saltelli and Funtowicz, 2014), spanning the concerns of sociologists (Chattoe-Brown, 2014), economists (Kuorikoski et al., 2010), ecologists (Grimm et al., 2005), biologists (Levins, 1993), and others. Moreover, it is of ongoing interest to a subfield of philosophy of science concerned with the development, justification and application of scientific models (Weisberg, 2006, 2012; Magnani and Bertolotti, 2017).

The methodological aspects of model design uncertainty are particularly well articulated in the discipline of social simulation (Poile and Safayeni, 2016; Polhill and Salt, 2017; Schindler, 2013) as are the implications of this uncertainty for policy (Aodha and Edmonds, 2017; Edmonds and Aodha, 2019). Social simulation is a discipline in which modellers have considerable flexibility in the design of their models, and have considerable uncertainty about their subject matter (i.e. social systems). As such, it's an area where the problems created by model design uncertainty are widely discussed and in which various methodological techniques have been articulated and which may help to address such issues (see below).

I begin this chapter by discussing these techniques and highlighting the ways their application might benefit researchers studying the social dynamics of FGM. I organise the discussion around 'model-based' and 'empirically-driven' techniques for model development, as well as 'application-focused' techniques which address the usage of uncertain formal models for policy purposes. I also discuss the potential benefits of combining these techniques. Some of the techniques discussed have analogues in the work of past modellers of FGM. However, none of them has been widely and systematically applied in the field. I argue that to apply the techniques discussed, the field may need to transition from analytical methods like game-theory to social simulation methods like agent-based modelling, which are more flexible in their design. Later in the chapter I propose a modelling strategy which combines these techniques. In remaining chapters of the thesis, I report an implementation of this strategy.

Managing Uncertainty in Model Design

Model-based Approaches

Popularized by the pioneering arguments of population biologist Levins (1966), model practitioners have increasingly recognised that they can *manipulate* their models to better understand key sources of uncertainty in their design. Key among these ‘model-based’ approaches are sensitivity analysis (SA) and robustness analysis (RA). Both involve model design manipulation, and both are concerned with the effects of these manipulations on model dynamics.

Models often contain parameters and other design elements (e.g. structures in the model) which are uncertain (Thiele et al., 2014). Parameters are literally numerical quantities (i.e. Real numbers) that play a role in a model. The distribution of thresholds represented in a heterogeneous threshold model might be controlled by parameters determining the mean and variance of the distribution, for example (see Chapter 6). ‘Structures’ in a model are more substantive aspects of its design, such as the representation of social interaction as occurring through random pairing (Chesnokova and Vaithianathan, 2010) or the structure of the decision-function of social actors (see Novak, 2016 for an example in the FGM literature). Elements, as the term is used here, encompass both of these aspects of a model’s design (see also ‘elements’ in Chattoe-Brown, 2017a, 2019; Chattoe-Brown et al., 2019).

One could use other terms to describe the parts of a model: sub-models (Railsback and Grimm, 2012), assumptions (Kuorikoski et al., 2010), features (Weisberg, 2012), design choices, aspects, and so on. The key point is that there are recognisable ‘parts’ to a model. Some parts one may be confident of, other parts one may be uncertain about.

Elements of a model might be uncertain because some aspect of a phenomenon is unknown, not measured, or would be difficult to measure (such as the distribution of thresholds in a threshold model). They might be merely hypothesised, or they might contain obvious simplifications (e.g. the idealised assumption of homogeneity of preferences in the social convention model). Elements of a model may also be uncertain because some parts of its design are relatively arbitrary, and included only to facilitate the analysis. An example of the latter might be representing social interaction as occurring through random pairing, in order to facilitate certain kinds of mathematical analysis (see also ‘tractability assumptions’ in Kuorikoski et al., 2010).

If these uncertain aspects of the model design *and possible alternatives to them* (e.g. alternative parameters or structures) are clearly defined, then sensitivity analysis can be used to *apportion* uncertainty about the dynamics of the model, into uncertainty about different elements of the model's design (Saltelli et al., 2008; Thiele et al., 2014). Sensitivity analysis might show, for instance, that the dynamics of a model are particularly *sensitive* to the value of one of its parameters. This would suggest that much of the uncertainty about these dynamics can be attributed to uncertainty about *that* parameter in particular. The basic procedure is straightforward. Uncertain elements in the model are varied systematically, then the contribution of each element to changes in the output of the model is measured statistically (see Saltelli et al., 2008 for an authoritative overview of relevant techniques).

Like sensitivity analysis, robustness analysis involves manipulating uncertain elements of a model's design and observing the effect of these manipulations. However, the aim of robustness analysis is to draw more general conclusions about the adequacy of a model. In essence, the aim of robustness analysis is to establish whether the key dynamics of a model remain *consistent* after manipulating uncertain aspects of its design (Kuorikoski et al., 2010; Levins, 1966; Orzack and Sober, 1993; Railsback and Grimm, 2012; Weisberg, 2006). This might involve adding new elements to a model, trying alternative designs, de-idealising elements of a model, and so on. If a model's dynamics remain consistent across these manipulations then they might be said to be robust (Levins, 1993). In essence, this 'robustness' would show that conclusions drawn from the model (that are based on the dynamics in question) don't depend on the choice between the different elements that were considered in the robustness analysis (Kuorikoski et al., 2010). Conversely, if the dynamics of the model are *not* consistent after manipulations of its design, then this shows that it is 'non-robust'. This means that the key dynamics of the model *depend on* at least one of the uncertain choices made in its design. In evaluating, or deploying, the model, particular attention will need to be paid to such choices.

These model-based forms of analysis have general applicability and could be usefully deployed in order to identify key areas of uncertainty in current or future models of the social dynamics of FGM. The formal analyses found in Novak (2016) and Platteau et al. (2017), and discussed in the previous chapter, can be considered a limited kind of robustness analysis, and they illustrate the potential benefits of the technique. Their analysis can be construed as an analysis of the robustness of the dynamics of the social convention model to the de-idealisation of one its simplifying assumptions: the assumption of homogeneity of preferences. The demonstration that the dynamics of the model are *not* robust to this

manipulation represents important progress for the field. It reveals the inadequacy of the model, and it identifies the assumption of homogeneity of preferences as a critical and costly idealisation in the model.

This technique could be applied much more generally and systematically to explore a whole range of possible de-idealisation and elaborations (e.g. new design elements) of existing models. This would mean manipulating the design of existing models by adding design elements, or relaxing simplifying assumptions (i.e. making elements of the design less simplified) and then assessing whether there is a substantial change in the dynamics of the model. If there is, then this would suggest a potential inadequacy of the existing model design. It would also reveal an important area of uncertainty in the design of the model. If the manipulation had involved adding an element that had previously been omitted (or abstracted away, such as a social network), this might suggest that the element cannot be safely left out of the model¹⁸. If it had involved de-idealising an element of the model (like allowing for heterogeneity of preferences), then this would suggest that the original idealisation was unsafe and potentially inappropriate.

There is no guarantee that all important and uncertain elements in a model design can be identified in this way. However, there is at least the possibility of pre-emptively identifying *some* of the areas of fragility in the design of the model (such as costly idealisations) with the expectation that this will make the model more reliable. Saltelli et al. (2013: 225) summarise this argument with a single phrase: “find sensitive assumptions before they find you”!

Although not framed as a robustness analysis per-se, Efferson et al. (2019) have recently explored the effects of various modifications of heterogeneous coordination models of FGM on predictions about *positive spillovers*. This activity is an exemplar of the kind of model-based analysis that I’m advocating be adopted more widely in the field. Chesnokova and Vaithianathan (2010) also included an explicit ‘robustness analysis’ element in their study. However, this was limited to only a single alternative representation, rather than a more general exploration of different possible model designs.

While there are examples of robustness analysis (or activities somewhat analogous to this) in the field, I am not aware of any examples of formal *sensitivity* analysis. Sensitivity analysis becomes particularly important when a wide range of uncertain design elements (e.g. many

¹⁸ Chattoe-Brown (2017a) includes a demonstration of exactly this kind, showing that the inclusion of a social network ‘matters’ in a model of infectious disease transmission.

parameters) are considered for the single model. If the modeller intends to empirically calibrate (see below) some of the uncertain elements in their model, then sensitivity analysis can help them to *prioritise* among these elements, by focusing on those areas of uncertainty that have the biggest impact on the dynamics of the model (Saltelli et al., 2008). Current models of FGM are arguably too simple for this kind of analysis to be relevant. However, as models in the field become more sophisticated (e.g. if results of robustness analysis reveal problems with existing simplifications), then this task may become increasingly relevant. As part of the development of my model (Chapter 6) I used sensitivity analysis to help establish a set of *priorities* for empirical calibration.

Both sensitivity and robustness analysis could be used in the field to better address the issue of model design uncertainty. However, it is important to note that while these techniques can help to *manage* uncertainty, they can't eliminate it. Put simply; even if a model's dynamics are robust across many alternative model designs, this doesn't establish that those dynamics are correct (Orzack and Sober, 1993). Instead, it shows that the dynamics of the model are not sensitive to the different possibilities considered in the analysis (Kuorikoski et al., 2010; Weisberg, 2006). The model might be 'non-robust' to other plausible modifications that weren't considered. Or, parts of the model design not treated as 'uncertain' might be incorrect (i.e. the underlying theory of decision-making). Equally important, if the model *is* found to be sensitive to a particular decision about its design, robustness analysis will not necessarily tell us which choice to make¹⁹. To actually *resolve* uncertainty about a model design, we need empirical data.

Empirically Driven Approaches

At least three important methodological approaches to engaging with empirical data in order to reduce uncertainty about a model design are discussed in the social simulation literature. These are independent calibration and independent validation (Chattoe-Brown, 2014, 2017a), and multi-criteria (a.k.a pattern-orientated) validation (Grimm et al., 2005; Railsback and Grimm, 2012). The principles underlying these techniques have general applicability and could be of significant value if applied widely and systematically in attempts to model the dynamics of FGM.

¹⁹ One exception might be if the element in question is recognized as an idealisation, in which case one already knows it is a 'false' part of the model.

Independent calibration is the use of ‘independent’ empirical data to reduce uncertainty about one of the individual elements of the design of a model. This is different from an overall assessment of the adequacy of a model’s design. Instead, independent calibration is the use of empirical data to improve the accuracy of a *particular* element of a model, considered in isolation. If, for example, Chesnokova and Vaithianathan (2010, c.f. Chapter 2) had used empirical data to show that potential wives and grooms *really are* matched randomly and then assortativity in the marriage market, then this would be an independent empirical calibration of this aspect of the model.

The distinctive feature of independent empirical calibration, is that it involves the use of an independent source of empirical data to justify parts of the design of a model (this is sometimes also called sub-model parameterisation, where a ‘sub-model’ is an element of the model design, Railsback and Grimm, 2012). This is distinct from justifying the design of the model *instrumentally*, i.e. because it helps the overall model to reproduce some dynamic that has been observed empirically. The use of independent empirical data to justify the theory of decision-making in models of FGM (see discussion of ‘Theoretical Core’ below), does occur in existing literature on models of FGM. This is an example of empirical calibration. However, I am only aware of one example of the use of detailed empirical data to calibrate *other* elements of model design (i.e. the elements that extrapolate from a theory of decision-making to a model of social dynamics). This example is found in Novak (2016), who used survey data to *independently* calibrate the representation of threshold-distributions in her model (see Chapter 7 for discussion of her methods).

The potential contribution of independent calibration to managing uncertainty associated with model designs is conceptually straightforward. Using empirical data to improve the fidelity of one aspect of a model reduces uncertainty about that aspect of the model and so indirectly contributes to the credibility of the overall model. As I discuss further below, independent calibration can be particularly valuable if used *in combination with* the model-based analysis techniques discussed above.

Nevertheless, independent calibration has its own challenges and limitations. It may be tempting to adopt the viewpoint that calibration can be used to confirm all of the sub-components of a model, and thus directly remove uncertainty about the model design. However, it is important to recognise that as a practical activity, calibration is dependent on other aspects of the model development process, which are themselves subject to uncertainty.

First, calibration is dependent on the elements actually included in a model. If an element is not included in a model, it cannot be calibrated - thus its omission remains a source of uncertainty not reduced by calibration (Chattoe-Brown, 2019, makes this point in distinguishing between calibration activity and prior model specification, as do Polhill and Salt, 2017, in distinguishing between ontology and calibration). Consider, for instance, that the representation of social interaction in a population in terms of a social network structure, can only be calibrated if a network is actually represented (which it is not in existing models of FGM, except recently in Efferson et al.'s (2019), work and early outputs of this project, see Appendix G).

Second, the activity of empirical calibration does not necessarily eliminate ambiguities associated with idealisations and simplifications. We might say, for example, that a network representation in a model is 'calibrated' if the density of the network is correctly represented. However, clearly, there is still scope for the representation to be further calibrated, such as through the correct representation of clustering in the network. And so on. Even if, in an extreme scenario, complete network data is available for the whole population of interest, this representation will still be idealised by virtue of the fact that real social networks are constantly changing (dynamic) and that real relationships between individual pairs of social actors are subject to their own unique idiosyncrasies.

As such, idealisation still remains present, and calibration itself does not tell us about how consequential a given idealisation will be for the insights gained from a model. Instead, empirical calibration only helps the modeller to reduce uncertainty about aspects of the model design whose uncertainty is recognised and well-defined.

Independent empirical validation, by contrast to calibration, involves comparing the overall dynamics of a model to patterns of macro-outcome observed empirically. In this sense, it offers a kind of holistic evaluation of the adequacy of a model. If a model reproduces macro-level empirical outcomes accurately, then this may afford it some credibility as a representation of the underlying social dynamics. As noted in Chapter 2, activity that is analogous to empirical validation has occurred in some part of the literature. The social convention model, for instance, has been associated with macro-level observations that FGM tends to be locally universal when practiced (although, of course, see empirical disagreements between Efferson et al., 2015, and Mackie, 2017).

The key concept emphasised in parts of the social simulation literature is *independence* (Chattoe-Brown, 2017b). This is the idea that data used to validate a model should be

independent of data used to design it. The problem that independence addresses is as follows. Models (especially if large and complex) may be flexible enough to arbitrarily reproduce macro-level patterns, given enough effort by researchers to tweak their designs. This might occur *even if* the model is actually a poor representation of the underlying dynamics and its predictions will not hold in novel contexts (see discussion of *equivifinality* in Chapter 2). The classic example of this kind of ‘tweaking’ of a model to arbitrarily fit data would be to take a model with many unknown parameters and to use an algorithm to find arbitrary values of these parameters that let the model reproduce the macro-level outcomes of interest (Thiele et al., 2014).

By contrast, the idea of independent validation is that the model is *not* built to deliberately fit macro-level data. Instead, the design of the model is based on independent sources of theory and data (often at the micro-level), and then it is ‘validated’ on other independent data (often at the macro-level). When this methodology is followed (and validation succeeds) it lends greater credibility to the idea that the model reproduced observed macro-level outcomes *because of* similarities between its design and real social dynamics. These principles have not been articulated in past studies modelling FGM, and some existing studies do not explicitly distinguish between data used to design and then test their models²⁰ (where this occurs at all). Explicit adoption of the principle of independent validation within the field has the potential to add credibility to future modelling work, and in so doing to directly reduce uncertainty about the adequacy of model designs.

The multi-criteria method (Railsback and Grimm, 2012) of assessing models provide a way of strengthening the empirical validation test. It can be used whether validation is independent or not. In essence, it is the use of *multiple* distinctive macro-level empirical patterns as tests of a model. To maximise the strength of the test, these patterns should be as distinctive and different from one another as possible. The more patterns used (and the more distinctive these patterns), the stronger the ability of the test to discriminate between adequate and inadequate models.

As noted in Chapter 2, the social convention model has been associated with more than one macro-level empirical pattern. These include ‘local-universality’, discrepancies between the popularity and prevalence of FGM, and the concentration of FGM practice in particular

²⁰ It is, for instance, far from clear whether the regularity that older women are often defenders of FGM practice is an independent prediction of Coyne and Coyne’s (2014) model of FGM, or one of the ingredients in its design.

localities and ethnic-groups. While this is clearly a step in the right direction, these patterns are quite vague. As I note in Chapter 2, they are potentially compatible with a range of models of FGM. Strengthening the patterns with detailed empirical data, or finding new *detailed* macro-level patterns, could help in the construction of a stronger validation ‘benchmark’ for modelling in the field as a whole. Recent efforts in Efferson et al. (2015) and Mackie (2017) to measure the actual distribution of rates of FGM across local communities, can be seen as an extension of the ‘local universality’ regularity which strengthens the validity test provided by this empirical pattern. In Chapter 7, I strengthen the macro-level pattern of discrepancies between the popularity and prevalence of FGM, by measuring the joint distribution of popularity and prevalence of FGM across local communities in Senegal.

Application-focused Approaches

Considerations of model-based and empirically-driven approaches to addressing uncertainty in model design suggest that the field could benefit from widespread and systematic adoption of such techniques. However, it may be too much to hope that this adoption can, in the short term, entirely eliminate uncertainty about the design of models in this domain. Model-based approaches can help us to locate sources of uncertainty in a model’s design, but are limited by the modeller’s capacity to conceive of, represent and analyse many different model design possibilities. Empirically-driven approaches are central to justifying the design of models, but are constrained by prior model design decisions (see discussion of empirical calibration) and limitations to the data available for calibration and validation.

This raises an important question. When there is residual uncertainty about the adequacy of a model of social dynamics, how should we use this model to contribute to complex policy-problems (such as efforts to eradicate FGM)?

Where uncertainty about a model design exists, the ‘engineering’ paradigm of model application, in which models are used to select the ‘optimal’ policy choice may not be appropriate (Edmonds and Aodha, 2019). The chief concern is that an uncertain model will be used to *drive* decision-making, with good policy thus hostage to the (unknown) inadequacies of the model itself. In Chapter 2, for example, I discussed the possibility that inadequacies in the social convention model have negatively affected policy efforts (e.g. by inflating expectations about the positive impact of public declarations of FGM abandonment). In this section, I consider proposals which depart from this paradigm of model application. These proposals involve the use of models of social dynamics to help practitioners anticipate different *possible* scenarios in which those dynamics might affect their

efforts. The aim of such proposals is to help practitioners to recognise and plan for such scenarios, as distinct from making recommendations for policy action. I refer to this as the ‘possibilistic’ approach to model application (Edmonds and Aodha, 2019)

The idea of a ‘possibilistic’ approach to model applications is articulated most clearly in Edmonds and Aodha (2019). The core rationale for the approach is that there is often considerable uncertainty about models of complex systems. Moreover, the misapplication of uncertain models in decision-making can lead to undesirable consequences, including subverting policy goals. Instead, by using models to illustrate *possible* outcomes of policy actions, modellers can avoid this risk, while still being able to contribute to policy. In particular, the emphasis is on widening practitioners’ awareness of possible outcomes (and the dynamic processes that could produce them). Possible scenarios in which policies could fail as a result of dynamic processes may be of particular interest in this regard. This is as opposed to narrowing practitioner’s awareness by making specific predictions about what *will* occur if particular actions are taken. This has the added benefit of avoiding competition between the model and other kinds of informal knowledge held by practitioners.

Like the social engineering paradigm, a ‘possibilistic’ approach still depends on making model-based inferences which have some degree of justification. If the model used to anticipate the scenarios in question has no credibility, then the approach becomes pure speculation. However, the required degree of justification is somewhat different than under a social engineering paradigm. For a possibilistic analysis of policy-failure to be useful only requires that the relevant predicted failure scenarios have *some* credibility. On the other hand, under the social engineering paradigm, there is a much greater requirement for the model to make reliable and precise predictions which are born out in future observations.

This idea of exploring ‘possible’ policy outcomes can also be found in an independent intellectual tradition. This tradition, which is often referred to as *scenario planning* is a direct response to the recognition that pervasive uncertainty exists when designing social policies or making strategic decisions (Amer et al., 2013; Nair and Howlett, 2017; Schoemaker, 1995; Stirling, 2010; Volkery and Ribeiro, 2009). Policymakers and other decision-makers often do not know what affects their actions will have, nor how their goals might be subverted by social processes which they didn’t anticipate. Yet, decisions must be made, and to the extent possible, these decisions should be driven by expertise and careful reasoning. A solution proposed and used by both academics and practitioners is to engage in *scenario planning*. Scenario planning is an activity in which different credible possible outcomes are identified

through careful reasoning about social (and other) processes, and the consideration of different kinds of uncertainty about how these processes might unfold. The purpose of scenario planning is not to identify precisely which outcome will occur, but rather to capture a range of credible possibilities. These credible scenarios then act as ‘early warnings’ of particular outcomes and as sources of information for contingency planning. Famously, the Shell Corporation used scenario planning in order to anticipate the *possible* decline of oil prices prior to 1973, allowing them to put into place contingency plans prior to the market crash (Wack, 1985).

A re-orientation of the goals of researchers studying the dynamics of FGM from ‘social engineering’ to identifying ‘possible’ scenarios of intervention failure might have a considerable advantage. It could help researchers to meet their goal of helping practitioners to end FGM *despite* uncertainty about the accuracy of the predictions that their models generate.

Combining Techniques for the Management of Model Uncertainty

Model-based, empirically-driven and application-focused approaches to the management of model uncertainty shouldn’t be seen as competing alternatives, but rather as techniques which work best in combination. Consider for example, that the value of sensitivity analysis and robustness analysis will be constrained if no attempt is made to improve (e.g. through independent empirical calibration) elements of the model found to drive uncertainty about its dynamics. Conversely, without sensitivity analysis, the modeller may not know which elements of their model design would benefit most from empirical calibration (see also Chapter 6). Furthermore, the success of model development efforts may be difficult to assess without empirical validation.

A model which has been developed in an ad-hoc fashion (i.e. without any attempt to manage design uncertainty) *could* be used in a ‘possibilistic’ context, however the value of the ‘possible’ scenarios generated by the model may be seriously limited, relative to a model for which some credibility has been established through model-based and empirically-driven activities.

From Traditional Analytical Methods to Agent-based modelling

In subsequent chapters of this thesis, I employ the methodological techniques discussed in this chapter to extend and develop existing efforts to model the dynamics of FGM (see ‘Modelling Strategy’ below). I use *agent-based modelling* (ABM) as my technical framework for model development. ABM is a modelling technique that simulates populations of autonomous social actors, in order to represent and analyse social dynamics (see details below).

Here, I argue that ABM is a recommendable approach not only for my own efforts but as a general framework for modelling the dynamics of FGM.

My primary argument is based on a consideration of the strong *representational capacity* of agent-based modelling. Put simply, ABMs can represent almost any feature of a social system that can be clearly articulated. Consequently, ABM doesn't require modellers to make arbitrary assumptions or idealisations purely for technical reasons. As I detail below, this special capacity means that modellers can apply the kinds of methodological techniques that I have advocated for in this chapter in a *principled* way (i.e. without being constrained arbitrarily by their modelling technique).

Nevertheless, while this argument emphasises the *general flexibility* that ABM affords to modellers, other proponents of ABM have focused on the methods' *distinctive* capacities – especially its ability to accommodate particular kinds of low-level details associated with complex social systems (see below).

In the following subsections, I provide an overview of ABM as an approach. I then present my primary argument for the use of ABM in this field, based on its representational capacity and consequent advantages for applying the methodological techniques described in this chapter. Subsequently, I consider alternative arguments for the use of ABM and discuss versions of them which strengthen the case for its application to the study of FGM.

What is Agent-based Modelling?

In the social sciences, ABM is generally understood as the practice of using object orientated computer programming (Squazzoni, 2012) to represent a population of social actors as autonomous agents within a computer program, and then simulating the action and interaction of these agents over time (according to rules of behaviour defined for each agent) (Epstein, 2006; Epstein and Axtell, 1996; Gilbert and Troitzsch, 2005; Railsback and Grimm, 2012; Squazzoni, 2012; Wilensky and Rand, 2015). Often, these agents represent individual

people within a population, although in other cases they might represent ‘autonomous’ elements of a social system, like businesses and institutions (Railsback and Grimm, 2012).

In a typical ABM, agents within the simulation are equipped with decision-processes, information-sensing (things they observe or can learn), attributes (often heterogeneous across agents), actions/interactions they can undertake, and so on (Gilbert and Troitzsch, 2005; Railsback and Grimm, 2012; Wilensky and Rand, 2015). The famous Sugarscape model (Epstein and Axtell, 1996; Gilbert and Troitzsch, 2005), for instance, represents a population of agents who can decide which direction to move on a 2D grid of cells (a decision-process), can observe resources on nearby cells (information sensing), are endowed with different rates of internal resource consumption (a ‘metabolic rate’ attribute) and are able to move to different locations and collect local resources (actions). They interact indirectly, through consumption of shared local resources (the ‘sugar’) and (in some mode variants) directly – by trading resources with one another.

The simulation of these agents within an ABM is then ‘run’ by repeatedly having the individual agents make decisions, processes information, act, interact, and so on, in a decentralised autonomous fashion within an explicit environment (such as 2D space, or a network) (Epstein, 2006; Gilbert and Troitzsch, 2005; Squazzoni, 2012; Wilensky and Rand, 2015). In Sugarscape, for instance, agents repeatedly decide to move in various directions, consume resources, interact, and so on over time within a shared 2D 50-by-50 spatial grid. Agents who cannot collect enough sugar to satisfy their metabolism die. They are then replaced by new agents with randomised attributes. The *aggregate* dynamics of the population within the simulation then ‘emerge’ from the individual actions and interactions of these agents. One of the well-known dynamics of Sugarscape is that from an initially symmetrical distribution of resources (‘sugar’) across the agents, a highly unequal (right-tailed) distribution quickly emerges, with a small number of agents usurping a disproportionate fraction of environmental resources.

Just like more traditional techniques (e.g. *n*-person game-theory) for modelling social dynamics (see below), ABMs attempt to represent key features of social actors and their interactions, and then to understand how these might aggregate to create different dynamics at the population level. In the case of the famous Schelling ABM of spatial segregation (Chattoe et al., 2000; Wilensky and Rand, 2015), the key features of social actors are asserted to be their identity-group, their relative spatial position to one another and their tolerance for immediate spatial neighbours (other agents) who have a different identity-group

membership to them. The ABM simulation is run to understand how these agent-features (the model's 'micro specification', Epstein, 2006: 7) will aggregate (under various conditions) to create different patterns of spatial segregation in a 2D environment.

In a similar way, in Granovetter's, (1978) analytical (*non*-ABM) threshold model of collective behaviour, the key features of social actors are asserted to be their binary choice of actions (e.g. whether they participate in a riot or not), and their threshold (defined as the proportion of others who need to riot before they will). Analytical techniques (e.g. such as the graphical method detailed in Chapter 2) are used to understand how these features will aggregate together (under various conditions) to create different levels of participation in collective behaviour.

However, while analytical methods and ABMs can have these broad aims in common, ABMs differ in that individuals (as well as their attributes, relations, and interactions) are explicitly separated into autonomous objects in a computer-simulated environment. In Watts & Dodds' (2007) threshold ABM of collective behaviour, for instance, agents are explicitly represented as objects connected by a social network. In the original analytical version of the threshold model and in similar analytical models used to study FGM (e.g. Novak, 2016) different social actors are *not* represented explicitly. Instead, the method derives results from a functional description of the cumulative distribution of thresholds within the overall population (see Chapter 2 for details).

In general, there is no requirement that the aggregate dynamics of an ABM be discoverable analytically (i.e. as the solution to a set of equations). Instead, analysts use observations of the behaviour of the model under different simulated experiments²¹ (known as 'simulated data', Gilbert and Troitzsch, 2005: 17) to document and understand its dynamics.

ABM and Representational Capacity

The different models of the dynamics of FGM discussed in the previous chapter depend on a range of different analytical techniques. These techniques have included matrix-game representations (Mackie, 1996; UNICEF, 2007), geometric representations such as the coordination diagram (Mackie, 1996; UNICEF, 2007), sequential game trees (Coyne and Coyne, 2014), iterated mathematical functions (i.e. the graphical approach to heterogeneous threshold models, Novak, 2016) and differential equations (Platteau et al., 2017). Each of

²¹ Repeated simulations using different starting conditions, stochastic seeds, and parameter values, for example, the Schelling segregation model under different level of agent tolerance for neighborhood diversity (Chattoe-Brown, 2013).

these techniques has different representational capacities. Matrix games, for instance, can model the interactions between two representative actors, but cannot represent interactions among heterogeneous populations involving n actors. Differential equation modelling can handle *n-person* situations where the future aggregate state of a system is a relatively simple function of its previous state (as occurs under simple versions of the heterogeneous threshold model, Platteau, et al., 2017). It can't reliably do so when this future state emerges in hard-to-predict ways from the complex low-level interactions of the system's constituent elements (Wilensky and Rand, 2015). These are all examples of limited representational capacity. Their consequence is that they arbitrarily limit the features of a social system that the analyst can account for when building a model.

ABM and Managing Model Design Uncertainty

The limited representational capacity of the traditional analytical approaches applied to the study of FGM (so far) would present a significant barrier to applying the methodological techniques described in this chapter. If certain aspects of a model's design are *imposed* by technical limitations, rather than chosen by the modeller, then it will not be possible to either test the robustness of the model to modifications of these elements or to improve them through empirical calibration (see earlier discussion). In other words, to fully apply the model-based and empirically-driven techniques discussed in this chapter, modellers need to be able to make a principled choice about the elements that they include in the design of their model.

Here then, the increased representational capacity of ABM makes it indispensable when employing techniques to manage uncertainty in model design. Using agent-based modelling, the analyst is largely free to explore whichever design features or de-idealisations of their model are of empirical interest, to test their results' robustness to those alternatives, and so on. This capacity is a formal property of computer simulation. Whichever features of a social system can be clearly articulated (in mathematical or natural language) can also be represented using computer code, known as Ostrom's third symbol system (Troitzsch, 1997). ABM's commitment to the explicit representation of low-level system components (what Squazzoni, 2012 refers to as 'ontological correspondence' pp.x) fully realises this capacity.

Decisions about what aspects of a system should be represented in a model, which aspects should be simplified, and so on, are therefore brought fully under the control of the analyst

using ABM. The analyst can employ the kinds of techniques that I've identified in this chapter, to try to manage this kind of uncertainty in a principled and systematic way.

This crucial capacity of ABM – to de-idealise the technical and conventional assumptions of analytical methods to investigate their robustness – is illustrated in the following examples.

Neo-classical microeconomic models of market exchange, based on analytical methodologies, assume that agent preferences remain constant over time, and generally predict that markets will arrive at price equilibrium. Epstein and Axtell (1996) used agent-based modelling to relax this assumption (whilst maintaining other neoclassical assumptions²²) and allowed agent preferences to vary over time according to plausible mechanisms of cultural transmission (Epstein & Axtell, 1996: 125). They found that classical models are *not* robust to the relaxation of the assumption of fixed preferences and that when the assumption is relaxed, the model does not attain price equilibrium.

Similarly, classical models of epidemic transmission, such as the Susceptibles, Infected, Removed (SIR) compartmental model, generally assume homogenous mixing – i.e. that all agents in the population who are infected (I) have some chance of infecting all agents who are susceptible (S). This allows the dynamics of the system to be described analytically through a device such as the Kermack-McKendrick equations (Epstein, 2006: 271). Chattoe-Brown (2020b) used ABM to relax the assumptions of homogenous mixing in the SIR model by explicitly representing individual agents and constraining their interaction through a network. He found that (whilst retaining other assumptions of the model) the dynamics of the model (especially the probability distribution of non-infected susceptibles remaining after the epidemic) showed meaningful non-robustness to the relaxation of this assumption.

Both of these examples (Epstein & Axtell, 1996, and Chattoe-Brown, 2020b) are illustrations of the technique of exploratory robustness analysis advocated in this chapter (and implemented in Chapter 5 for coordination models of FGM). Both required their respective analysts to utilise the extended representational capacity provided by the ABM framework.

Arguments for ABM from Complexity

So far, I have argued that ABM is a recommendable technical framework for this field because its representational capacity makes it sufficiently flexible to support a systematic

²² Specifically: 'agents [still] have Cobb-Douglas utility functions and engage only in Pareto-improving trades with neighbors' (Epstein, 2007: 20)

methodological treatment of uncertainty about the way a model is designed. This kind of argument is *design-content agnostic*. In advancing it, I don't claim that the kinds of distinctive model features supported by ABM are *necessarily* required in order to furnish an adequate model. I just assert that they *might be* and, therefore, that a prior exclusion of such features from considerations for technical reasons prevents model design uncertainty from being approached in a systematic or principled fashion.

The persuasive strength of this argument lies in its inclusivity. No pre-existing preference for the distinctive features of agent-based models is required. The argument only asserts that modellers should be free to represent whichever features of social systems *they* have reason to believe are most important. This perspective, I hope, will appeal to other modellers in this field, who may have only limited interest in the distinctive 'generative' perspective propounded by some ABM practitioners (Epstein, 2006; Squazzoni, 2012).

Nevertheless, the case for the use of ABM can only be strengthened if there are *also* reasons to believe that the *distinctive* model design features supported by ABM are *likely* to be needed to create an adequate model of the social dynamics of FGM. Promotion of the *general* use of ABM in social research has frequently presented arguments of this form, emphasising the virtues of ABM for studying complex systems driven by social interaction (Chattoe-Brown, 2017; Squazzoni, 2012; Wilensky and Rand, 2015).

A complex system is broadly defined as one whose aggregate dynamics are strongly driven by the interactions of its low-level constituent parts (Ball, 2012). In the context of a social system, these 'constituent parts' are social actors: individuals, households, firms, vehicles and so on. The properties of complex systems that are of particular interest to social scientists are *emergence*, *non-linearity* and lack of predictability from simple aggregate summaries.

Emergence is the phenomena that in certain systems, the *aggregate dynamics* of the system are difficult to describe (or are even unexpected – though see Epstein, 2006 for a critical discussion of this idea) in terms of the properties of individual system components, or mere summation of those individual components (Railsback and Grimm, 2012: 101). Emergent dynamics are particularly striking when they arise from relatively simply low-level interactions that bear limited resemblance to their aggregate effects and are not governed by any centralised controller (Wilensky and Rand, 2015). Immediately accessible illustrations of emergence in 'real world' complex systems include traffic flows (Ball, 2012; Wilensky and Rand, 2015) and the flocking behaviour of birds (Wilensky and Rand, 2015). Often, traffic jams, for example, don't arise from macroscopic 'inputs' to the system, such as a blockage

created by an accident, but from the low-level interactions of traffic entering and exiting highways (see also Wilensky, 1997). Likewise, the distinctive ‘V-shaped’ formations of geese flocks do not arise from a centralised set of ‘hierarchy’ rules known to the birds involved. Rather, they emerge from the low-level interactions of birds when they attempt to fly in shared direction, whilst avoiding others and staying near to neighbours (Wilensky, 1998; Wilensky and Rand, 2015). A more ‘sociological’ example of emergence is the (apparently) spontaneous emergence of grassroots social movements or protests in the absence of ‘centralized control or public communication’ (Watts, 2002: 5766).

Non-linear systems are broadly those whose outputs (e.g. their aggregate dynamics) change in ways that are not proportional to changes in their inputs (Hardesty, 2010). Stated more intuitively, the magnitude of some ‘effect’ in a system may be disproportionate to the apparent size of its ‘cause’ (Ball, 2012). A classic example of this kind of effect is a bifurcation, or tipping-point, whereby a small change in some feature of a system leads to large scale changes in its aggregate dynamics. An accessible (and sadly, topical²³) example of non-linearity in a ‘real world’ complex system is that of bifurcation in epidemic outbreaks. Epidemics can display a bifurcation pattern, whereby the introduction of a disease into a population can either lead to minor outbreaks that fizzle-out (formally ‘stochastic extinction’ or ‘epidemic quenching’) or devastatingly widespread outbreaks (Epstein, 2006: 290) – but rarely intermediate outcomes. This difference may potentially depend on low-level stochastic effects or the particular details of vaccination or immunity patterns in the population. The faddish popularity of certain media content (books, music, videos, memes etc.), might also be considered examples of non-linearities, since the difference in the prior quality of ‘hits’ and ‘flops’ may appear trivial compared to the magnitude of difference in their eventual success (Salganik et al., 2006; Watts, 2002).

Finally, the dynamics of complex systems are often thought to be difficult to explain or predict using simple aggregate summaries of their properties. Stock markets are generally thought to display this feature, whereby short-term fluctuations in stock prices are considered virtually unpredictable using available aggregate data on prior stock performance. This is related to the phenomena of emergence – short-term fluctuations in stock prices *emerge* in hard to anticipate ways from the low-level interactions of individual traders. This may be true for real-world economic systems in general – it has been claimed that there are no clear

²³ At the time of writing the world was in the grip of the deadly COVID-19 pandemic.

examples of statistical models successfully predicting turning-points in macro-economic cycles or financial markets (Moss and Edmonds, 2005: 1098).

A very wide range of social phenomena have been ascribed the label ‘complex system’, including traffic, crowd movements, social norms, the spread of crime, epidemics, financial systems and urban development (Ball, 2012). Indeed, it can be argued that social interaction itself frequently implies some level of complexity (Squazzoni, 2012). This perspective naturally translates into the recommendation that ABMs be used to model social phenomena where meaningful social interaction is present.

As described in the previous section, ABMs are well suited to explicitly representing the *interactions between social actors* that are thought to drive the dynamics of complex social systems, as well as the complicated social structures thought to mediate these (Chattoe-Brown, 2017). Moreover, because they employ a simulation approach, ABMs facilitate the discovery and study of complex aggregate dynamics (like emergence, non-linearity, etc) even if the analyst isn’t aware of (or doesn’t anticipate) these before the fact (Wilensky and Rand, 2015: 33). More specifically, ABM facilitates the representation of a range of features that have been linked to (and are considered important for modelling) complex social systems, especially: heterogeneity of social actors, local interaction, explicit spatial environments, bounded rationality, and sophisticated internal cognitive processing (Epstein, 2006; Gilbert and Troitzsch, 2005; Railsback and Grimm, 2012; Squazzoni, 2012).

An initial argument for the use of ABM to study the social dynamics of FGM, then, depends on a consideration of whether it too, is a complex system. There are good reasons to believe that the dynamics of FGM meet the definition of a complex system. First, as discussed in Chapters 1 and 2, there is widespread evidence that FGM is a ‘relational’ (UNICEF, 2013: 35) phenomena – that is to say that social interactions and social influence play an important role in decision-making. Moreover, the established ‘Social Convention’ model of FGM (see Chapter 2) predicts that tipping-points (i.e. non-linearities) can occur in the dynamics of the system, as a result of social interactions driven by coordination incentives. This, at the very least, suggests that ABM may be a suitable and useful tool in the study of the dynamics of FGM.

Nonetheless, this argument doesn’t yet establish that ABM is to be *preferred* over existing approaches (e.g. n -person game theory) for studying the social dynamics of FGM. Here we run into a fundamental problem of disciplinary conventions in the social sciences (Chattoe-Brown, 2020b). ABM does not have a monopoly on representing social interaction, nor on

modelling non-linear dynamics. Granovetter’s (1978) famous demonstration of threshold models did not use simulation or ABM, yet managed to model non-linearities arising from social interaction within a heterogeneous population (albeit only one dimension of heterogeneity – decision thresholds). Previous models of FGM (see Chapter 2) have grappled with the implications of social interaction and identified potential tipping points through the use of n -person game theory (and similar techniques) alone (Chesnokova and Vaithianathan, 2010; Mackie, 1996; Platteau et al., 2017; UNICEF, 2007).

The problem is that different kinds of complex system will require different degrees of low-level detail to model adequately, and different disciplines (ABM, game-theory) may make different assumptions about what these are. In order to make a compelling argument²⁴ (from complexity) that ABM is *preferred* for studying the dynamics of FGM, we need evidence that the distinctive forms of social complexity that only ABM is well suited to addressing are likely to ‘matter’ in the study of the phenomenon (Chattoe-Brown, 2020b).

The fundamental problem of model design uncertainty (see Chapter 1) means that we cannot *know* in advance what model features we will need in order to model FGM dynamics (complex or otherwise). However, there are at least two inductive arguments which make the importance of ABM in this field appear probable, these are detailed below. First, the trend in the field so far has been toward the identification of increasingly low-level details of social interaction (what Squazzoni et al., 2014, refer to as the ‘microscopicness’ of social reality) as potentially important determinants of the dynamics of the practice. This trend has probably not been exhausted, and further extending the representation of this low-level detail will likely require an ABM approach. Second, by examining the design of ‘parallel’ models (those with an analogy in design or topic to current models of FGM) we can potentially identify model features that require an ABM framework, and so are *also* likely to ‘matter’ in modelling the dynamics of FGM. These two arguments proceed as follows.

First, thus far in the field, as modellers have added more detail to their representation of low-level interaction in FGM-practicing communities, they have discovered that the social dynamics of these systems are sensitive to these additional details. Two crucial examples of this are the introduction of heterogeneity of preferences into models of FGM as a social norm of coordination (Efferson et al., 2015; Novak, 2016; Platteau et al., 2017), and

²⁴ I.e. compelling to modellers whose conventional assumptions about what ‘matters’ in models of social interaction don’t automatically align with those of the ABM community.

explorations of networked interaction (Efferson et al., 2019, and early analysis done as part of this dissertation, see Appendix G). The role of heterogeneity of preferences has been shown to be crucially important for coordination-based models in the field, as discussed in detail in Chapter 2. Early analysis that was done as part of this dissertation (see also, Chapter 5) showed that relaxing the assumption of global interaction in coordination models of FGM, by introducing social networks as a structure to mediate social interaction, can significantly affect the policy-relevant dynamics of coordination models (see Appendix G). Moreover, analysis by Efferson et al. (2019) of a similar kind of model also found important sensitivities to the introduction of networked social interaction.

These introductions of greater levels of heterogeneity, and more sophisticated low-level social structure (networks), have revealed important (potential) sensitivities in the corresponding social dynamics. It is reasonable to suppose that further elaborations will reveal yet more sensitivities (and I show that this is indeed the case, see Chapter 5). Such elaborations require ABM, since the introduction of increasingly detailed low-level social structure and individual heterogeneity quickly exhausts the possibilities for closed-form analytical modelling (Epstein, 2006).

A second argument follows a similar pattern but refers to ‘parallel studies’ using ABM in other domains (i.e. not FGM) yet which share some analogy with existing models of FGM. A number of ABM-specific elaborations in such models have been shown to have important implications for social dynamics – again suggesting the *potential* need for the special capacities of ABM to adequately model the social dynamics of FGM.

Arguably the set of ABM studies with the closest analogies to existing models of FGM, are those that utilise the binary threshold decision mechanism (see Chapter 2 and Granovetter, 1978) that has been popular among models of FGM (Efferson et al., 2019; Novak, 2016; Platteau et al., 2017)²⁵, in order to study collective behaviours. This mechanism is part of a class of decision-rules known as ‘binary decisions with [positive] externalities’ (Watts, 2002: 5766). It is based on the basic premise that the likelihood of a social actor choosing either one of two options, increases with the proportion of others who make a similar choice (Watts and Dodds, 2007: 443). The threshold-based version of the rule has foundations in a rational choice model of decision-making (see also the final section of this chapter) that predicts that

²⁵ Note that Platteau et al. (2017) and Novak (2016) derive an equivalent of the threshold mechanism from a rational choice account of the utility functions involved.

actors will switch their choice when the social incentives to do so reach a critical threshold (see Chapter 2 for a detailed discussion of how threshold rules are derived from rational choice assumptions about FGM).

A number of existing models of FGM implement a version of this mechanism, due to a theoretical conceptualisation of FGM decision-making as a coordination game (see Chapter 2). As such, ABM studies which implement a similar mechanism, but which *also include* elaborations which are ‘ABM-specific’ (i.e. out of reach of traditional n-person game-theory), may indicate the usefulness of the ABM approach for the field of FGM. Below, I discuss four examples from the ABM literature which employ the threshold mechanism to study collective behaviour. As I discuss subsequently, these models do not target FGM specifically, and there are a number of reasons not to consider them to be candidate models of the practice. Nevertheless, they reveal important additional sensitivities of threshold dynamics to elaborations of social structure, individual heterogeneity and cognitive processing that would be difficult to explore without ABM.

Watts (2002) presented a binary threshold model and agent-based simulation of collective behaviour, implemented over a simple random network. They explored the conditions under which small changes in individual behaviour could ‘cascade’ over the network – a notion with an analogy to tipping-points in models of FGM (albeit an imperfect one, see below). They found important sensitivities to social network structure, including overall connectivity. They chose a simple random network generation process in order to make some analytical results tractable. Yet they still had to rely on simulation for certain results, and simulations revealed ‘surprising’ results not anticipated by their analytical work (pp. 5769). This suggests that the introduction of low-level social structure using a very simple algorithm (i.e. a random network) places a threshold-based model of collective behaviour right on the cusp of what can be investigated analytically.

Subsequent work, using a similar model to study the role of ‘influentials’ (defined as those with high network degree) in cascade formation, relied entirely upon an ABM simulation approach (Watts & Dodds, 2007).

Hu et al., (2015b) employed a similar ABM to Watts (2002). They investigated the sensitivity of model dynamics to further low-level details in the social structure of the agent-population. In particular, they investigated ‘local opinion heterogeneity’ (i.e. homophily) – especially the extent to which individuals are connected to others that differ in preference from them. They found further sensitivity of aggregate dynamics to these low-level features.

Another study using a similar ABM of collective behaviour (Hu et al. 2015a), found (among other things) sensitivity of model dynamics to an additional dimension of heterogeneity in agent decision-making – conformity. This was implemented as the sensitivity of different agents to social incentives. They found that this could interact with the targeting-bias of simulated ‘interventions’ to significantly affect the aggregate dynamics of behavioural change.

Finally, Córdoba and García-Díaz (2020) used an ABM to investigate the potential sensitivity of diffusion-of-innovation models (which use a similar decision-mechanism) to a more sophisticated kind of cognitive processing, whereby agents ‘reflexively’ observe both their local social environment (e.g. network contacts) as well as the ‘global’ trend of behaviour in the population. They found that this had a substantial influence on the spread of collective behaviour through the population.

These studies show that models employing a threshold decision mechanism (based on positive social influence) may generally be highly sensitive to elaborations of their social structure, individual heterogeneity and representation of cognitive processing. Given the popularity of this decision mechanism in the FGM literature, this points evidentially to the distinctive usefulness of ABM for the field.

However, it is important to recognise that these are not models *of* FGM and that the *particular details* of the elaborations employed may or may not be appropriate for modelling the dynamics of the practice. In general (despite the analogous decision-mechanism used), there is no reason to expect that ABMs from other domains can be taken ‘off-the-shelf’ and deployed as candidate models of FGM. The design and application of such models will not have prioritised the key features of FGM (and associated policy problems) (most likely it will also not have followed a systematic model development strategy, including model calibration and validation – as I do and recommend in this thesis).

For example, certain design assumptions in the ABMs noted above are a poor fit for FGM, including the use of empirically implausible random networks (Watts, 2002, Watts & Dodds, 2007), the assumption that decision-making is uni-directional²⁶ (Watts, 2002, Watts &

²⁶ I.e. that all agents initially choose option A, and can only switch to option B, never the other way around. By contrast, in the case of FGM, it is clear families with or without a cutting tradition can choose to either practice or abandon FGM, and could even reverse their decision given changing social pressures.

Dodds, 2007) and the restriction of analysis to the case where preferences are uniformly distributed²⁷ (Hu et al, 2015a).

Likewise, the *applications* of these models don't translate well into the study of FGM. The exogenous 'interventions' that they use to explore system dynamics involve only a tiny fraction of agents in the population (Watts, 2002; Watts & Dodds, 2007; Hu et al. 2015a). This is fundamentally at odds with the kinds of policy problems that models of FGM have been used to study, in which interventions are often assumed to involve a substantive proportion of actors in the local communities (UNICEF, 2007; Platteau et al., 2017) – mimicking the collective interventions favoured by development workers seeking to reach a 'tipping point' in collective behaviour (UNICEF, 2007).

As such, while existing 'parallel' ABM research suggests important *potential* sensitivities of models of FGM to low-level details like social structure and individual heterogeneity, establishing this more concretely, requires the use of models targeted at FGM specifically. This should involve elaborations based on empirical evidence *about* FGM which are explored in relation to aggregate dynamics that are *directly relevant* to the problems that models of FGM are used to study (e.g. dynamics related to collective interventions).

That problem is tackled directly in this thesis (see 'Modelling Strategy' below, and Chapter 5). Also, as I show in Chapter 5, a proper consideration of low-level empirical features of social interactions surrounding FGM suggests a range of elaborations that are not present (to my knowledge) in 'parallel' ABM studies of collective binary decision-making *or in* existing formal models of FGM, such as *coalition formation* processes, *multiple* social influence networks and *organised diffusion* (see Chapters 5 & 6).

ABM as a Complementary (not Competing) Modelling Technique

It is important to note that advocating for the use of ABM in this field does not necessarily entail recommending that existing models (or insights derived from them) be discarded. Models of social dynamics developed using other techniques can be *translated into* ABMs, with little distortion of their underlying assumptions. Stated in terms of my primary argument for the use of ABM, the superior representational capacity of agent-based modelling *subsumes*

²⁷ This could never be considered an acceptable assumption in the FGM modelling literature, given significant interest in the sensitivity of results to the particular distribution of preferences assumed (e.g. Efferson, et al., 2015; Novak, 2016).

the designs of existing models of FGM. Rather than erasing the work of past analysts, it expands the capacity available to extend that work (and explore the robustness of its assumptions!).

All of the designs of existing analytical models of FGM depend on a set of assumptions (which may or may not be idealised) about underlying micro-level social processes. The heterogeneous threshold model of FGM (e.g. Novak, 2016), for example, assumes a population of social actors who repeatedly and simultaneously decide whether or not to engage in a behaviour, based on the proportion of other actors who engaged in that behaviour in the previous time-step, and their own threshold for acting. These assumptions could be represented *directly* in an ABM. When stimulated, this model would produce the same dynamics that are predicted by the analytical method used in Novak (2016, see also Chapter 5).

In general, agent-based models of social dynamics can reproduce the behaviour of analytical models to an arbitrary degree of precision²⁸. This kind of replication is a recognised activity, called ‘docking’ a simulation (Epstein, 2006: 275). Building a docked model can be a useful precursor to expanding or refining an existing analytical model because the docked model represents a rigorous point of comparison for more elaborate model designs (see Chapter 5).

Summarising the Case for Agent-based Models of the Dynamics of FGM

As noted above, a range of arguments are available that point to the value of ABM as a future framework for modelling the social dynamics of FGM.

A general ‘design agnostic’ argument for the use of ABM in the field is the additional representational capacity provided by the method. This capacity removes arbitrary constraints on model design that are imposed by more traditional methods, and thereby facilitates the kind of principled approaches to model design uncertainty outlined earlier in this chapter.

A consideration of the insights gained in the past through adding additional complexity to the design of models of social dynamics generally, both in this field and in parallel ABM research, points to the likelihood that the kinds of low-level details supported by ABMs (e.g.

²⁸ Analytical methods often assume infinite populations. ABMs can only represent finite populations, but these populations can be arbitrarily large.

detailed social structure and agent heterogeneity) will be important for producing an adequate model of the dynamics of FGM.

Ultimately, proponents of existing models in the field should have little reason to object to a shift in the field toward the adoption of ABM as a technical framework, because the ABM framework is able to subsume the design of these models.

A recent publication (published shortly after this thesis was first submitted) related to the social dynamics of FGM (Efferson, 2019), includes at least one formal model which is recognisable as an agent-based simulation. In keeping with the arguments made above, this appears to have been utilised in order to include a social network in the model. To the best of my knowledge, except for the early outputs of research for this dissertation, this is the first analysis of FGM to use ABM. As networks (and other model features, such as those that I introduce in Chapter 5 - which are difficult to examine without the use of an ABM) come to be recognised as essential for understanding the dynamics of the practice, the field may move organically in the direction of the ABM framework.

The ‘Theoretical Core’: Separating Models of Social Dynamics from Theories of FGM Decision-Making

In Chapters 1 and 2, I talked about the importance of distinguishing between theoretical ideas about FGM decision-making, and models of social dynamics which extrapolate from these. I pointed out that the dominant ‘social norm of coordination theory’ involves an important set of ideas about how actors make decisions about FGM, and that these are clearly separable from the social convention model, which extrapolates *from them* to generate claims about social dynamics. In Chapter 2, I showed that this distinction has important implications for the current status of social norm of coordination theory, and for recognising the distinctive challenge that *model design uncertainty* poses for the field. Here I consider further implications of this distinction for model development practice.

A number of methodological discussions of modelling complex systems recognise the distinction between the central theoretical ideas that motivate a model design and the details of the design of a model that extrapolates from those ideas. Kuorikoski et al. (2010: 547), for example, distinguish between ‘substantial assumptions’ which encompass “the causal mechanism about which the modeller endeavours to make important claims” and other kinds of assumptions in the design of a model. Levins (1993: 554) refers to the ‘core’ of a

model and distinguishes this from the various other aspects of the model used to facilitate the analysis, including simplifications.

I adopt Levins' (1993) terminology and refer to the theoretical 'core' of models of the social dynamics of FGM. In models of FGM, this 'core' is the theory of decision-making employed in the model. The social convention model, for instance, started with a theory about the decision-making process of actors in FGM-practicing communities, and then built on this to create an *n-person* game-theoretic model of the social dynamics that might arise under coordinated decision-making.

This conceptual distinction between the theoretical core of a model, and other particulars in its design has important practical implications. To a significant extent, once model development has started, the theoretical core of a model is taken as given (Kuorikoski et al., 2010; Levins, 1993). There are many plausible ways to formally represent 'social norm of coordination' decision-making (see Chapter 5), for instance. But if one believes that FGM *is* 'social norm of coordination' then one would not represent decision-making in a way which grossly *contradicted* the main claims of the theory. Furthermore, when exploring the robustness of such a model to possible modifications of its design, one wouldn't consider modifications that substantially *undermined* the core theory of the model. To do so would be somewhat incoherent. This makes the selection of a 'core theory' a crucial preliminary step in model development that will affect which design possibilities are considered and the way these are interpreted (see also 'Theory Development' in Railsback and Grimm, 2012). This also suggests that different core theoretical ideas about decision-making may correspond to different *classes* of possible models of the social dynamics of FGM. I discuss this issue, as well as different possible 'theoretical cores' for models of FGM, further in the next chapter.

It is important to note that this framing of the modelling problem doesn't imply that we are merely building models 'of theory' as an abstract exercise divorced from empirical data (see 'Theoretical Exposition' in Edmonds et al., 2019). It's not enough that the core theory is of intellectual interest; we want to be confident that the core theory is actually correct (Kuorikoski et al., 2010). Selecting an appropriate core theory is an empirical question. Indeed, preliminary selection of a core theory is and should be the first place empirical evidence enters the modelling process (Levins, 1993). In a sense, it is a kind of empirical calibration undertaken before the model is built. Rather than changing the purpose of the

modelling²⁹, distinguishing between ‘core-theories’ and other aspects of model design is a way of structuring model development and identifying where different substantive ‘claims’ enter the model design.

Nevertheless, one might object to this idea of selecting a core theory which constrains subsequent model design choices. Why not make every aspect of the model design ‘fair game’ for activities like robustness analysis? There is nothing to stop this in principle, but it might be difficult to implement in practice, and it might have undesirable consequences for the field of modelling FGM as a whole. In practice, without the constraint of a core theory of decision-making, the range of possible model designs to be considered would become much larger, with associated practical challenges. Furthermore, a great deal of non-modelling research on FGM, including the discussion of anti-FGM interventions has been organised around ‘core’ theoretical ideas about decision-making, especially social norm theory (Mackie, 2000, 2009; UNICEF, 2010, 2013). Building a model without recognising and justifying a ‘core theory’ risks denuding the analysis of its ability to ‘speak to’ this literature – either by disputing existing theories or by improving the understanding of existing theories’ implications for social dynamics³⁰.

Conclusion: The Modelling Strategy

Here I outline a modelling strategy which is designed to combine the different methodological techniques described in this chapter. It also aims to capitalise on the potential for these techniques to complement one another. I have argued that the widespread application of any of the techniques described in this chapter could benefit the field. However, here, I suggest a procedure to *combine them* into a complete strategy for model development and application. The remainder of this thesis is guided by this strategy, with Chapters 4-8 implementing the steps of the strategy. I outline the strategy here, with the practical issues involved discussed in the relevant chapters. The strategy consists of seven steps

1. **Selecting a ‘theoretical core’:** Different core theories of FGM decision-making are identified and subjected to empirical testing. The best-supported theory is selected,

²⁹ Which in this field had been to develop policy-relevant insights into social dynamics in real FGM-practicing communities.

³⁰ In Appendix H1 I critically discuss an alternative to using domain theory to build agent-based models: generic behavioral frameworks

and used as the basis for developing a ‘standard’ agent-based model of the social dynamics of FGM (Chapters 4/5)

2. **Exploratory robustness analysis:** The important dynamics of the ‘standard’ model are measured as a baseline, and then the model is manipulated by introducing de-idealisations (i.e. relaxing simplifying assumptions) and plausible elaborations (i.e. introducing previously omitted features) to its design. These possible elaborations of the model design are chosen with reference to existing models, theory and empirical research on FGM. Where the standard model is ‘non-robust’ to these new design elements (its dynamics are disrupted by them), this constitutes a demonstration that these elements represent important sources of uncertainty about the model design (Chapter 5).
3. **Constructing a ‘general’ model:** A ‘general’ new agent-based model is built which incorporates the different design possibilities (i.e. potential de-idealisations and elaborations) shown to be important in Step 4. These different possible designs are mapped to the parameter space of the model, such that the different design modifications considered in Step 4 (and different possible combinations of them) are ‘points’ in the parameter space of the model (Chapter 6).
4. **Global Sensitivity Analysis of the ‘general’ model:** The general model is subjected to global sensitivity analysis of its parameter space, so that uncertainty about its dynamics can be *apportioned* to uncertainty about different elements of its design. The results of this analysis provide a set of priorities for independent empirical calibration.
5. **Independent Empirical Calibration:** Based on the priorities established through sensitivity analysis, key uncertain elements in the design of the general model are calibrated using empirical data (Chapter 7).
6. **Independent Empirical Validation:** The adequacy of the (calibrated) general model is tested against multiple macro-level patterns relating to FGM practice in real communities (Chapter 7).
7. **Possibilistic Failure Scenario Analysis:** Using the (calibrated) general model, a ‘possibilistic’ approach is used to apply the model to policy-problems. The model is used to identify ‘possible scenarios’ under which social dynamics might disrupt the goals of community-level interventions hoping to promote the widespread abandonment of FGM.

In designing this strategy, I hoped to build on the complementary relationship between different methodological techniques. The early use of exploratory robustness is designed to help the modeller pre-emptively identify important model elements and problematic simplifications. It also provides a clear demonstration to the field as a whole that particular elements need to be considered when modelling the dynamics of FGM (i.e. that a simpler model might be inappropriate). Combining different possible designs into a general model makes it possible to use global sensitivity analysis to identify the relative contribution of these different uncertain elements of the design to uncertainty about the key dynamics of interest. In turn, this establishes priorities for empirical calibration. Validation of the calibrated model helps to establish its credibility. The ‘possibilistic’ approach to model application benefits from this credibility, whilst recognising remaining uncertainty about the model’s design.

Chapter 4: The Theoretical Core – Hypotheses about FGM Decision-Making Processes in Senegal

Introduction

The aim of this chapter is to identify the best-supported ‘theoretical core’, for the (further) development of formal models of the social dynamics of FGM (see Chapter 3). The identification of a ‘theoretical core’ allows selection of an initial model to begin the analysis and provides a conceptual boundary for model development activities (see discussion of ‘Theoretical Core’ in Chapter 3). It guides, for example, the way that new model features are formulated. To identify the best supported ‘theoretical core’ for modelling, the chapter evaluates three empirical hypotheses. These are hypotheses about decision-making in FGM practicing communities. These hypotheses represent foundational assumptions in attempts to understand the social dynamics of FGM.

First, much of the existing work on the social dynamics of FGM has operated within the ‘social convention/social norm of coordination’ theory (Mackie, 1996, 2000, 2017; Mackie and LeJune, 2017; Mackie et al., 2015; Shell-Duncan et al., 2011; UNICEF, 2007; UNICEF, 2010; UNICEF, 2013). The core of this theory is that actors make decisions about FGM that are interdependent and subject to coordination incentives. Decisions are interdependent because the decision of one actor depends on the decision of others (in the same community). Actors face coordination incentives because they face pressure to make the same decision as others. Thus, coordination is a type of interdependence. This theoretical contention has motivated model-based analyses of social dynamics. Specifically, the dynamics that arise when social actors’ decisions about FGM are coordinated (e.g. the social convention model, Mackie, 1996; UNICEF, 2007).

Second, other modelling work, while recognizing the potential for actors’ decisions to be interdependent, has diverged from an emphasis on social coordination. Some other modelers have focused on incentives favouring social *competition*. They have focused on competition among women within a marriage market to get better husbands (Chesnokova and Vaithianathan, 2010; Ross et al., 2016). Emphasis on competition as the main motivation for FGM can lead to different kinds of model building. As I discuss further below, competitive motivations can mean that actors face strong ‘mis-coordination’ incentives. Actors may be

motivated to be one of the few who practice FGM because this conveys greater relative marriage advantage.

Third, some studies have emerged which may challenge the idea that actors' decisions are interdependent (Bellemare et al., 2015; Efferson et al., 2015). There is strong evidence, noted in Chapter 2, that actors in FGM practicing communities influence one another in some way. However, questions have been raised about whether actors' decisions are interdependent. Decisions are interdependent if the decision of one actor depends directly on the decisions of another. However, actors could influence one another in other ways. It could be that actors only influence one another's beliefs about the intrinsic value of FGM. If this were the case, then actors' decisions would depend only on their personal attitude to the practice. They would not be coordinating or competing. They would simply be exchanging views. We could call this alternative hypothesis 'informational influence'. A hypothesis of informational influence points to another distinctive set of possible models. Models of informational influence are sometimes called models of 'opinion dynamics'. Flache et al. (2017) provide a useful survey of this type of model.

Thus far, there have been few concerted attempts to discriminate between these three hypotheses. However, at least some evidence has accumulated for each perspective. I contend that these hypotheses are clearly distinguishable both theoretically and empirically. But, they are not necessarily distinguishable using the arguments employed in existing discussions. I argue that the most appropriate way to adjudicate these different ideas is to examine decision-making directly. This is because the hypotheses differ in their claims about actors' motivations for practicing FGM.

In the first section of this chapter, I provide a brief overview of each hypothesis. I also outline relevant evidence that lends that hypothesis some basic credibility. In the second section, I undertake empirical analysis. I use national survey data from Senegal to evaluate the three hypotheses. To do this, I quantify the relative contributions of different kinds of beliefs about FGM to mothers' decisions to have their daughters cut.

My findings in this chapter support the social coordination hypothesis. They also strongly cohere with existing qualitative research in Senegal. So, I ultimately select social coordination as my 'theoretical core' for later model development.

Hypothesis 1: Coordinated Decision-Making

At least three kinds of observation of FGM practicing communities lend some support to the social coordination hypothesis.

First, in some communities, uncut women may not be considered eligible for marriage. Uncut women may be seen as promiscuous, unclean or in some other sense unfit to be brides (Mackie, 1996; Mackie 2017). Thus, women (and their relations) may face the social cost of being unable to marry if others practice FGM while they do not.

Second, individuals may experience negative social sanctions for not ‘participating’ in FGM. Uncut women may be ostracized, subject to ridicule, excluded from community events, and so on (Shell-Duncan et al., 2011). Pressures may be placed on family members and others in the community to participate in, and support, the tradition (Coyne and Coyne, 2014). This may include arranging for their relations to be cut, joining ceremonies, and so on. There is a commonly held view within some communities that the practice is a valued and necessary tradition. The maintenance of this tradition may be seen as the proper choice of a ‘responsible’ parent and family (Yoder et al., 2004: 13). The abandonment of the tradition may be seen as an act of “openly flaunting shared norms” by those involved (Bicchieri and Mercier, 2014: 64).

Third, ‘clustering’ of the prevalence of FGM can be observed within FGM-practicing countries (UNICEF, 2013). FGM is concentrated in particular socially interconnected populations. These can be defined in terms of both ethnicity and geography (c.f. UNICEF, 2013). This clustering of the prevalence of the practice is compatible with interdependence of decision-making about the practice (Mackie, 2017; UNICEF, 2013).

The social convention model of FGM (see Chapter 2) was initially motivated by the observation that uncut women may struggle to marry. However, as recognition of the immediate social costs of not practicing FGM has increased, the theory has shifted to a more general emphasis on coordination, which can be driven by immediate normative social pressure, or marriage concerns, or both. This shift is associated with the transition from the term ‘social convention’ to the term ‘social norm of coordination’ (Mackie and LeJeune, 2009; Mackie et al., 2015, Mackie 2017, Mackie, 2019).

The distinction between marriageability concerns and more immediate social pressure is potentially important for model building. I explore this as part of model development in

Chapter 5. However, the archetypal social convention model of FGM is generic with respect to these issues. It just depends on the existence of strong social coordination incentives.

Hypothesis 2: Competitive Decision Making

Distinctions between coordination and competition in FGM decision-making have received limited attention (though see Ross et al., 2016). Yet such distinctions are critical. Emphasis on competition incentives can motivate radically different models of social dynamics (see below). Chesnokova and Vaithianathan's (2010) model assumes that decisions about FGM are motivated by a wish to gain a relative *advantage* in the marriage market. This is distinct from a wish to avoid social exclusion. Ross et al. (2016) also consider that decision-making may be contingent on women gaining a relative advantage in the marriage market. In both models, these 'relative advantages' depend on there being variation in spousal value. If some spouses are more desirable than others and FGM improves women's 'mating value', then it can be an avenue to an advantageous marriage.

The assumption that FGM is about getting better husbands rather than avoiding exclusion can imply a decision-making process at odds with social coordination. First, competitive motivations can disappear if too many women practice FGM. If everyone cuts, then there is little competitive advantage to be gained from the practice. The opposite is true under social coordination incentives. Coordination incentives to practice are greatest when the practice is most prevalent. Second, competitive motivations can be strongest when no others practice FGM. If FGM is valued by potential spouses, then the relative marriageability advantage conveyed by the practice will be greatest when practicing alone. Conversely, coordination incentives to practice FGM are absent when no-one else practices.

The marriage-competition hypothesis appears to come from accounts of the origins of the practice. Mackie and LeJeune (2009) provide an example of just such an account. Their account associates the practice with extreme resource inequality in ancient African empires. They suggest that under extreme resource inequality, wealthy men were able to attract many wives. These men also exerted tremendous power over the marriage market. However, because of their many wives, these men were said to have faced paternal uncertainty (not being sure if their wives' children were their own). This incentivized them to monitor and control the fidelity of their wives. FGM appears as a method of fidelity control to help wealthy men address this 'problem'. These wealthy men then exerted pressure on families to cut their daughters. Families were willing to accept this to ensure that their daughters could

marry these 'elite' husbands. From there, the practice of cutting daughters to help them 'marry-up' is said to have diffused down the social hierarchy. Over time, FGM came to be viewed as a desirable assurance of fidelity even in monogamous unions.

Some support for this account is found in historical records indicating that FGM was seen as a way of ensuring fidelity (Mackie, 1996). It still has this status in some populations. Also, the theory may explain why “the practice [of FGM] is most prevalent and practiced in its most severe form around the former centres of the ancient Nubian and Malian empires” (Mackie and LeJeune, 2009: 4).

This historical account certainly implies the importance of marriage competition as part of the origins of FGM. The key question is whether (and where) this continues to be of central relevance in decisions about the practice today. Over time, the tradition may have come to be valued in other ways and for other reasons, including religious ones (Mackie and LeJeune, 2009).

There are at least two kinds of contemporary evidence in favour of the marriage competition hypothesis. First, evidence that men in some FGM practicing populations prefer women who have been mutilated as marriage partners. Second, evidence that women who are cut have 'better' marriage outcomes.

Chesnokova and Vaithianathan (2010) cite a survey of young Somali men in London (UK) conducted by Morison et al. (2004). This survey found that a majority of young men who arrived in England aged 11 and older said they would prefer to marry a woman who has been cut. Chesnokova and Vaithianathan's (2010) own analysis found that women in Burkina Faso who were cut tended to marry earlier and into more wealthy households. They interpret this as consistent with the idea of FGM conveying a relative marriage advantage. Similarly, recent cross-national analyses of survey data have suggested that cut women have more children (Howard and Gibson, 2017). Both of these potentially indicate a relative marriage advantage from FGM practice.

Hypothesis 3: Informational Influence

There can be little question that social influence of some kind operates in FGM practicing communities. Analyses of national surveys show that FGM practice is strongly correlated within communities (see Chapter 2). Also, this correlation is robust after taking into account wealth, education, and other shared features (see Hayford, 2005, for example). However, this

does not necessarily establish interdependence of decisions. The 'informational influence' hypothesis is that actors influence one another's non-social beliefs about the practice and then take decisions based on these non-social beliefs.

This perspective appears to have become much more prominent since Efferson et al.'s (2015) study was published in *Science*. Their study raised doubts about the original social convention model. They claimed to have found stable interior rates of cutting within practicing communities in Sudan. A popular reaction to this study was to interpret it as revealing that decisions about FGM are based on 'private values', rather than social coordination pressures (e.g. 'Female Genital Cutting is Based on Private Values Rather than Social Norms' *Medical News Today*, 25th September 2015). The authors of the study also indulge this view (without necessarily endorsing it). They point to ways in which social actors could plausibly value FGM for its own sake:

“[our findings] points toward other potent forces [than coordination] sustaining cutting. Families may value cutting because they see it as a religious obligation (13), or they see cutting as the only way to produce feminine women in a society where gender must be clearly marked (13).” (Efferson et al., 2015: 1447)

Whether or not these values are truly 'intrinsic' might be open to interpretation. However, cultural preferences for femininity, or religious motivation, are likely not to depend on the immediate decisions of one's neighbours. As Dellenborg's (2004) ethnography of the Jola in Senegal highlights, women may view FGM as crucial for their spiritual status:

“The prayers of a woman who is not circumcised will not 'take' as well, they will not give her as many 'points' as had she been excised.” (Dellenborg, 2004: 82)

The plausibility of the hypothesis of informational influence is supported by two observations. First, some actors in FGM practicing communities view it as having intrinsic advantages. These advantages include: promoting women's health and hygiene, fulfilling the edicts of religion, and having aesthetic appeal (Mackie, 1996; UNICEF, 2013). Second, practicing and non-practicing families can, on occasion, be found living 'door-to-door' (Powell, 2017: 2). This is compatible with the idea that such households are making decisions based on their own (diverging) private values.

These observations suggest that the hypothesis of informational influence should be taken seriously. The hypothesis would also be in line with theories that have proved useful in public health, such as the 'stages of change' model. That model also relies on individualistic notions

of decision-making. Shell-Duncan and Hernlund (2006) offer a critical discussion of the stages-of-change model in the context of FGM.

Avenues for Discriminating Between the Three Hypotheses

Above, I aimed to show that each hypothesis has some basic plausibility, making it worth considering. Now, I turn to the question of how one might effectively adjudicate the three hypotheses. I start with some of the ways this has been attempted in the past.

In the existing literature, there appear to be two broad approaches to this problem. The first approach involves using macro regularities to evaluate the different hypotheses. Macro regularities are patterns in the prevalence of the practice across multiple communities. Different aggregate regularities have been said to indirectly support the different hypotheses about decision-making. The second approach involves attempts to examine decision-making processes directly.

Supporters of the coordination hypothesis have pointed to the concentration of FGM within ethnic groups, geographies and communities, within which it is ubiquitous (Mackie, 2017; UNICEF, 2013). Critics of the coordination perspective have emphasized variation within these groupings. They point to communities where not everyone practices FGM and neighbours make different decisions. Supporters of the marriage competition hypothesis point to a different set of regularities. They provide evidence that cut women are married more quickly, are wealthier and have more children.

I am sceptical of the idea that core theoretical assumptions about FGM decision-making can be litigated in this way. The different decision-making hypotheses are claims about micro-social processes. That is, they are claims about individual social actors. Yet the patterns noted above are at the macro-level. As I have emphasized in previous chapters, extrapolating from micro-processes to macro-outcomes is fraught with uncertainty. Fundamentally, it is a model-based activity. We need a model to extrapolate from micro to macro. Therefore, our confidence about the connection between a micro-hypothesis and a macro-regularity depends on our confidence in the particular formal model used to connect the two. We currently have little reason to imbue any particular model of FGM with great confidence (see Chapter 2).

As discussed in Chapter 2, the prediction of the social convention model: that FGM should be locally universal, is fragile. It can be disrupted, for example, by the introduction of

heterogeneity of preferences. Thus, the coordination hypothesis is not necessarily condemned-by or confirmed-by such findings. Flache et al., (2018) show that opinion-dynamic models, which embody the hypothesis of informational influence, can predict both convergences of opinion and divergence of opinion within communities. As such, evidence of variation in decision-making within communities has no clear-cut bearing on the informational influence hypothesis. The observation that cut women marry better is compatible with the competition hypothesis. However, models of coordination and informational influence do not necessarily contradict this prediction. As such, findings of marriage advantage for cut women do not really adjudicate the hypotheses either.

The macro-level observation that the prevalence of FGM often exceeds its popularity (UNICEF, 2013) arguably *could* falsify the informational influence hypothesis. If this hypothesis were true, we would expect only those who approve of FGM to practice it. However, we must allow for the possibility that such discrepancies in empirical data reflect temporal differences in the available measures. The prevalence of FGM among adults reflects actions that may have been taken a long time ago (Yoder et al., 2004). The popularity of FGM, by contrast, represents beliefs at the time of the survey. There is also the possibility that individuals under-report their approval of FGM (Efferson et al., 2015; Powell, 2017). Furthermore, while the regularity that the prevalence of FGM exceeds its popularity *might* falsify the informational influence hypothesis, it's not clear that it can discriminate between hypotheses of coordination and competition. Both imply social incentives that might lead actors to make choices that don't match their personal attitudes.

Instead of looking at macro-patterns, when adjudicating 'core' theoretical ideas about FGM, one should avoid conflating hypotheses with models. One should investigate hypotheses directly. This means directly investigating decision-making.

A case in point is Shell-Duncan and colleagues' in-depth mixed methods research on FGM decision-making in Senegal and the Gambia (Hernlund and Shell-Duncan, 2007; Shell-Duncan et al., 2011; Shell-Duncan and Hernlund, 2006). These studies investigated how decisions about FGM are reached. They primarily used interviews and focus groups in different local sites in Senegal and the Gambia. They investigated participants' concerns about marriageability. They also asked about the role of marriage in FGM decision-making. The findings of these studies supported a hypothesis of coordination, rather than marriage competition or informational influence.

First, study participants rarely cited FGM as related to marriage, and rejected the idea that it was an avenue to a better marriage:

"All interviewees and [focus group] participants were asked if FG[M] was necessary for a girl to find a "good" husband; most, whether from a circumcising tradition or not, agreed that circumcision was not necessary for marriage, and none asserted that it was an avenue to get a richer or better husband." (Shell-Duncan et al., 2011: 6)

Second, participants emphasized the *social costs* faced by uncut women. Practicing FGM meant social support and inclusion in important aspects of community life. Non-practice meant social criticism and ostracization. Shell-Duncan et al. (2011) concluded that FGM is a 'peer convention'. They suggested that interdependence of decision-making operates strongly among female peers within the local community. As such, in line with the original social convention model, they recommended that "change must be coordinated among members of [the social network of the local community]"(p. 12).

Shell-Duncan and colleagues' research provides relevant evidence for evaluating the hypotheses considered here. However, their findings depended on limited samples from only one region of Senegal and two regions of the Gambia. Moreover, their method of investigating decision-making relied primarily on qualitative interpretation of interviews. Their research did include a survey component. However, the use of this survey was limited primarily to two tasks. First, they surveyed marriage practices and calculated rates of FGM-discordant marriage. Second, they polled respondents about what they perceived to be the advantages of FGM. Neither task is a direct investigation of decision-making (I discuss the relevance of polling belief about FGM, below). None of this invalidates their research of course. However, it leaves open the possibility that their findings are subject to the idiosyncrasies of their sample and qualitative methodology.

In my analysis (below), I am able to assess whether Shell-Duncan et al.'s (2011) findings can be replicated using a very different methodology and sample. I discuss this further below. The methodology I use quantifies the contributions of different beliefs to decisions about FGM. I also use a nationally representative survey sample from Senegal. This is a classical use of a triangulation approach to social research. It assesses whether "a hypothesis can survive the confrontation of a series of complementary methods of testing" (Webb et al. 1996: 174, as cited in Korgen, 2017). The results of my analysis cohere with Shell-Duncan et al.'s (2011) findings. This successful triangulation lends considerable support to the coordination hypothesis.

Beliefs about FGM in National Surveys

One alternative ‘quantitative’ approach to studying FGM decision-making has been to poll actors’ beliefs about the practice.

National Demographic and Health Surveys (DHS) have been conducted in many countries in which FGM is concentrated. Many of these surveys have questioned women about the perceived advantages of the practice (UNICEF, 2013).

One strategy for interpreting these surveys has been to look at which ‘advantages’ of FGM are stated most frequently. Coyne and Coyne (2014: 140), for instance, point out that marriage is *not* the most commonly cited advantage of FGM in many national surveys. They treat this as evidence against Mackie’s (1996) original account of the practice. Mackie (2017:15), arguing in favour of a broader ‘social norm of coordination’ hypothesis, follows similar reasoning. He examines the frequency with which ‘social’ advantages of FGM are cited in national surveys. Such beliefs are cited frequently, and Mackie treats this as evidence that FGM is a social norm.

However, I would argue that this strategy is not an investigation of decision-making. The beliefs which are cited most often are not necessarily the beliefs which most strongly motivate decisions. Bicchieri et al., (2014: 6), make a similar point in discussing possible measures of social norm decision-making. They point out that even if we know the range of beliefs people hold about a practice “we still need to find out what *causal role* these potential motives play”.

Results presented below show, for instance, that ‘suffering’ is among the most commonly cited disadvantages of FGM in Senegal. Yet this belief explains relatively little of the variation in decisions about the practice (other examples can be found below).

What one really needs to know is whether a given belief motivates actors to participate in the practice. In other words, one needs to quantify the relative contributions of different beliefs to decision-making. This is a task that lends itself to statistical modelling.

There are many past statistical analyses which have modelled decisions about FGM (see Chapter 2). However, these have almost always modelled decisions as a function of demographic characteristics, social and environmental factors (Hayford, 2005; Kandala et al., 2009; Kandala and Shell-Duncan, 2019; Modrek and Liu, 2013). I show below that this

does not help us to adjudicate the hypotheses of interest. Instead, we need to model decisions about FGM as a consequence of beliefs.

This is a relatively novel approach in the field. One exception is a study conducted by Pashaei et al. (2016). The researchers conducted a small survey of 300 mothers in a single country of Iran. They used structural equation modelling to estimate the impact of personal attitudes, ‘subjective norms’ (i.e. social pressures) and ‘perceived behavioural control’ on mothers’ intentions to cut their daughters. They found that personal attitudes and subjective norms played an important role in decision-making. Their findings are largely consistent with my own (see below). The other is Hayford and Trinitapoli (2011), but they only considered a single belief (in FGM as a religious requirement), and didn’t compare the impact of different beliefs.

Testing Decision-Hypotheses Against the Relative Contribution of Beliefs to Decision-Making

The three hypotheses *do* make different claims about which beliefs are important in social actors’ decisions about FGM. However, there is some overlap between them. All three hypotheses allow for an important role of non-social beliefs. These are beliefs about the intrinsic value of FGM. However, the coordination hypothesis and the competition hypothesis *also* emphasize the importance of social beliefs. These are beliefs about the value of FGM that are contingent on the decisions of others in the community. By contrast, the informational influence hypothesis allows no important role for social beliefs.

We can further distinguish the competition and coordination hypotheses by which social beliefs they treat as most important. Under the social norm of coordination hypothesis, the major contribution comes from beliefs about social acceptance. The belief that FGM conveys social acceptance has been treated as an important indicator that it is a social norm in past analyses (UNICEF, 2013; Mackie, 2017). The difference in my analysis is that, consistent with the arguments of Bicchieri et al. (2014), I evaluate the hypothesis against the *causal role* played by that belief. Under the competition hypothesis, the major contribution to decision-making shouldn’t come from belief that FGM conveys social acceptance, but from belief that it conveys *marriage advantage*. Both of these can involve marriage. However, under the coordination hypothesis, the primary concern is *social exclusion* from the marriage market *and or* other positive aspects of community life. Under the competition hypothesis, the

primary concern is gaining relative *advantage within* the marriage market, in order to attract the most desirable spouse.

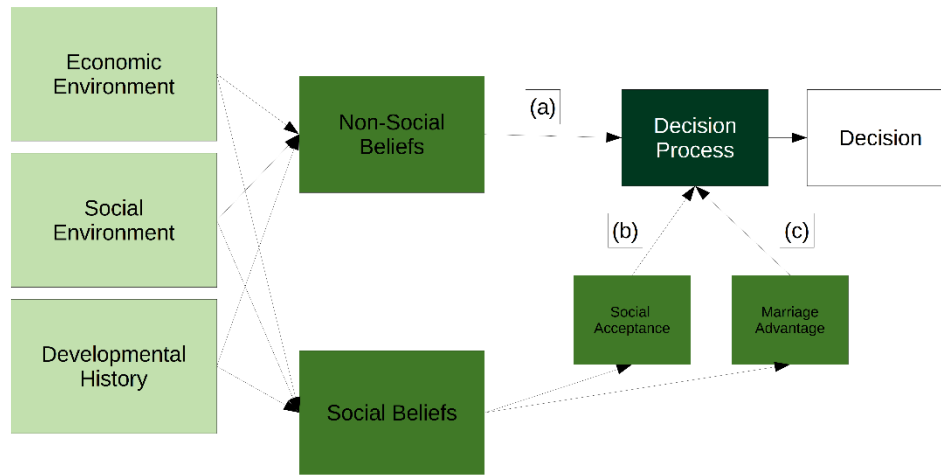


Figure 7: Causal Structure of FGM-related Decision-Making (generic to the three hypotheses).

Figure 7 illustrates this argument through a causal diagram which is generic with respect to the three hypotheses. This diagram begins with the environment (both social and economic) of the actor, as well as their developmental history. These are treated as 'antecedent factors'. They are assumed to influence actors' beliefs about FGM. In turn, actors' social and non-social beliefs about FGM influence their decisions about whether to engage in the practice.

It is important to note the mediating role of beliefs in this diagram. Decisions themselves are assumed to depend on beliefs only. Environmental and development factors operate on decisions through the formation of beliefs. This is a fundamental theoretical assumption in any sociological analysis of decision-making and is explicit in formulations of social norm theory (see 'expectations' in Mackie et al., 2015).

The technical importance of this assumption is that it implies that actors' decisions about FGM are independent 'conditional on' their beliefs about the practice. This assumption provides a technical justification for the regression-based analysis employed below.

The three hypotheses make different claims about the three inputs to the decision-process in Figure 7. The informational influence hypothesis asserts that the most important input is 'Non-Social Beliefs'. It asserts that 'Social Acceptance' and 'Marriage Advantage' contribute little. The competition hypothesis asserts that the most important contributions are from 'Non-Social Beliefs' and 'Marriage Advantage'. The coordination hypothesis asserts that the most important contributions are from 'Non-Social Beliefs' and 'Social Acceptance'.

Some national DHS surveys (see below) in FGM-practicing countries contain measures for each of the components of Figure 7 (except the decision-process). As such, it is surprising that statistical analyses of FGM have usually focused solely on antecedent factors (environment and development), with few exceptions. This has been the case even when analysts cite the social convention account as motivating their analysis (e.g. Hayford, 2005). Such analyses use actors' social environment as a proxy for coordination pressures.

Figure 7 illustrates why this strategy will not distinguish between the three hypotheses. The problem is that the antecedent factors could plausibly influence any of the proximate beliefs. An association, for instance, between levels of FGM practice in actors' social environment, and their own decision to practice FGM, does not imply the correctness of the social coordination account. The social environment may influence non-social beliefs. Maybe interacting with others who practice FGM leads one to believe that the practice has intrinsic benefits, and so on.

Quantifying Variable (Belief) Importance

The target of the empirical analysis is a decision-process. This decision has some inputs, which are beliefs about FGM. It also has an output, which is a decision to practice FGM (or not). As described below, empirical measures of the inputs and outputs are available. Our hypotheses make different claims about the decision-process. Specifically, they make different claims about the *importance of different inputs*. However, the decision-process is not observed directly.

The decision-process is a 'black-box'. To adjudicate the hypotheses, one needs to develop insight into the relations between its inputs and outputs. The natural approach to doing this is to use statistical analysis. One can observe variation in the inputs (beliefs about FGM) and corresponding outputs (decision to practice FGM, or not). Therefore, one is in a position to analyze the relationship between the two.

One's immediate instinct in measuring the importance of different beliefs for decisions about FGM might be to use main-effects regression. A simple example of this would be to fit an ordinary least-squares (OLS) regression of the form:

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_p x_{pi} + \epsilon$$

Where β_p are estimated coefficients, and x_{pi} are the beliefs held by actor i and y_i represents actor i 's decision to practice FGM (or not). The importance of each belief would then be captured by the β coefficient associated with it. The problem with this approach is that it imposes a very simple functional form on the relationship between beliefs and decisions. It imposes that each belief contributes independently and additively to the decision. It doesn't allow for interactions between beliefs (i.e. the presence of one belief might alter the effect of another). One can add interaction terms to the model. However, it would then become hard to delineate the contribution of individual beliefs through inspection of the model alone. What is required is a metric to estimate the importance of individual variables in the context of a more complex functional form that allows for interactions.

Murray and Conner (2009) provide an overview of variable importance metrics that are compatible with complex regression models. Two, in particular, stand out for their ease of interpretation. These are the squared zero-order correlation (SZOC), and the squared semi-partial correlation (SSPC, sometimes also called the coefficient of semi-partial determination). Both of these work on the principle of partitioning the variance explained by the regression model, into contributions of different variables³¹.

The SZOC is the proportion of variance in an output explained by a single input alone. This includes the unique contribution of the variable, as well as 'redundant' variation in the output that could be explained by other correlated inputs. The SSPC represents the unique additional contribution of the input to explaining variation in the output. It is expressed as the absolute contribution to overall variance explained. The SSPC takes into account interactions with other inputs. SZOC and SSPC were applied to measure the variable importance in this study.

Methods, Context, and Data

Context: Senegal

Senegal is the westernmost country of mainland Africa. It receives frequent mention in this thesis. The country has a national prevalence of FGM practice (% of women aged 15-49 who have been mutilated) of around 25%. It has also been a geographic reference-point for a number of important ideas in discussions of FGM. It was the original site of the much-

³¹ Quantifying the contribution of different inputs to variance in an output is also a classic technique in sensitivity analysis of complex simulation models (Saltelli et al., 2008; ten Broeke et al., 2016)

celebrated Tostan program (see Chapter 2). Tostan's work has been associated with social convention/social norm of coordination theory (Mackie, 2000). Close attention has been paid to the achievements of Tostan, and to levels of FGM in Senegal. Successful abandonment has been heralded as vindicating the theory (Mackie, 2000). Continuing FGM has been treated as undermining it (Platteau et al., 2017).

Perhaps due to its prominent status, Senegal has been the site of high-quality empirical research on FGM. This includes Shell-Duncan and colleagues' studies of decision-making, discussed above (Hernlund and Shell-Duncan, 2007; Shell-Duncan et al., 2011; Shell-Duncan and Hernlund, 2006). It also includes Diop et al.'s (2008) detailed evaluations of the Tostan program.

Senegal has theoretical importance and has been the site of high-quality research. Also, in 2005, it hosted a national survey with detailed questions about participants' beliefs about FGM. Theoretical relevance, past-research and available-data make Senegal ideal for the analysis of this chapter. I also continue to focus on the national context of Senegal in later chapters.

Data: Senegalese Demographic and Health Survey (2005)

The Demographic and Health Survey (DHS) program conducts nationally representative household surveys of populations in the developing world. These surveys provide wide-ranging information. Measures include sociodemographic, health and nutritional characteristics. The 2005 Senegalese survey also included questions about FGM.

The central part of the survey is the women's questionnaire. This questionnaire was given to all women between the ages of 15 and 49 residing within households targeted by the survey. The structure of the survey included 22 households which were surveyed at random from within pre-selected geographic clusters in Senegal. These clusters were existing census enumeration areas; typically a village or city-block. 376 clusters were selected from the Senegal census, spanning rural and urban parts of the country, and including all major regions.

The questionnaire was administered verbally, in-person, by trained fieldworkers who spoke the local language. Fieldworkers obtained informed consent prior to data collection and participation was voluntary. Participants were assured of the confidentiality of their answers. All interviews were conducted in privacy (including the separation of husbands and wives). Discussion of sensitive subjects was terminated if privacy was breached.

The empirical analysis in this chapter, and in Chapter 7, relies primarily on the 2005 Senegalese DHS survey. These data is publicly available and was collected from 7412 households in Senegal between February 2005 and May 2005. A total of 14,602 women between the ages of 15 and 49 from these households were surveyed.

Surveyed women were asked (twice, in different ways) whether they had ever heard of FGM. If they answered in the affirmative, then they were asked a series of questions about FGM. These questions covered their own experience of FGM, their daughter's FGM status and their views about FGM. If none of their daughters had been cut, women were also asked about their intentions to cut daughters in the future. Questions about women's views on FGM covered their beliefs about the advantages of the practice.

The 'decision-inputs' for quantitative analysis in this chapter were participants' stated beliefs about the advantages and disadvantages of FGM. The 'decision-output' was an indicator of whether the participant practiced FGM. Further details are as follows.

Independent Variables: Beliefs about FGM

Women were asked what the advantages were for girls to be cut³². They were also asked what the advantages were for girls *not* to be cut. In both cases, they were prompted for further responses until they stated that there were no other advantages.

Advantages stated in *favour* of cutting were (pre) classified as:

- FGM Improves Women's Hygiene
- FGM Improves Social Acceptance
- FGM is a Religious Necessity
- FGM Improves Marriage Prospects
- FGM Preserves Women's Virginity/ Prevents Premarital Sex
- FGM Improves Sexual Pleasure for Men

Advantages stated in favour of *not cutting* were (pre) classified as:

³² The survey did include three other questions in a different format, about pre-marital sex, religion, and men's attitudes to FGM. These partially duplicate the measures noted in this section. Moreover, they were asked directly, rather than stated spontaneously after prompting. Due to redundancy and non-comparability of the measurements I did not include them in the analysis.

- Avoids health problems
- Avoids suffering
- Improved Sexual Pleasure for Women
- Improved Sexual Pleasure for Men
- (Non-Practice) Accords with Religion

Note that each of these classifications is a (potentially imperfect) translation of the original French questionnaire.

In the analysis, each of these beliefs was represented as a dummy independent variable.

Some of the beliefs are straight forward to interpret in terms of the three hypotheses in question. Beliefs in improved social acceptance clearly accord with the coordination hypothesis (see above, also Mackie et al., 2015; UNICEF, 2013). Beliefs in improved marriage prospects accord with the marriage competition hypothesis. Beliefs about health, hygiene, women's sexual pleasure and religious-necessity imply non-social motivations.

Other beliefs are more ambiguous. Consider, for instance, beliefs about men's sexual pleasure or the preservation of virginity. Men's perceived sexual 'pleasure' from cut-women might motivate marriage competition. Or, it might be relatively incidental to marriage decisions themselves. It might be more relevant for an expectation of 'happy' married life itself. In that case, it would be more like an intrinsic quality of the practice. Or, perceived connections between FGM and virginity/pre-marital sex might imply a coordination incentive because pregnancy outside of wedlock can be a family dishonor in Senegal (Diop et al. 2008: 24). Or, it might imply an intrinsic benefit of the practice. Parents may view pregnancy outside of wedlock as inherently reprehensible (and expensive). In that case, FGM would have 'value' irrespective of other's choices.

Fortunately, these areas of ambiguity did not impinge significantly on the analysis. The beliefs shown to play the most important role in decision-making were also the most straightforward to interpret.

Dependent Variable: Decision to Practice FGM

Unfortunately, the DHS survey does not ask women about their participation in FGM related activities. This is perhaps due to the illegality of the practice. However, it does ask questions that can be used as a proxy of women's decisions. Women are often the final decision-makers about the FGM status of their own daughters or have a central role in such

decisions (Yoder et al., 2004: 13, 21; Eldin et al., 2018: 24). As such, if one of a woman's daughters has been cut, this can *usually* be taken as an indication of their participation in the practice. The DHS survey asks all women with at least one currently living daughter how many of their daughters have been subjected to FGM. If none of a woman's daughters has been cut, the participant is also asked if they intend to cut their daughters in the future.

The dependent variable used in the analysis was a dummy indicator of women's participation in the practice. This indicator was coded as '1' (FGM participation) if at least one of the woman's daughters had been cut or if they intended to cut their daughter(s) in the future. Otherwise, it was coded as '0' (Non-participation). Only women who had at least one living daughter were included in the analysis.

Sample

Out of the 14,602 women in Senegal who were surveyed, the analysis targeted women who had heard of FGM ($n = 13,747$), who had at least one living daughter and who responded to questions: about their daughters' FGM status ($n = 7280$), and their own beliefs about FGM ($n = 7245$).

Analytical Strategy

The technical aspects of calculating squared zero-order correlation (SZOC) and squared semi-partial correlation (SSPC, a.k.a. coefficient of semi-partial determination) in this case, are described in Appendix B1.

Results

If belief in improved social acceptance is the most important input to decision-making, this strongly supports the social norm of coordination account of decision-making, relative to hypotheses of competition or information influence. My findings suggest this is the case in Senegal.

Key results are reported in Table 2. Table 2 presents the SSPC and SZOC scores for each belief variable. It also shows the frequency of that belief (in the available sample). Figures 8 and 9 plot the position of each belief on each dimension of importance (SZOC and SSPC). The absolute values, and rank ordering of the results, under both frameworks, were very similar. The complete linear probability model (all main effects and second-order interactions) explained 47% of the variation in decision-making. The complete logistic regression model explained 48% of the variation in decision-making (see Limitations). To avoid redundancy, I just discuss the logistic regression results.

Table 2: Variable Frequency and Importance (Beliefs about FGM, Senegal DHS 2005)

Belief	Sq. Zero- Order Cor. (LMP)	Sq. Semi-Partial Cor. (Coef. Semi-Partial Determination) (LMP)	Sq. Zero- Order Cor. (Logit)	Sq. Semi-Partial Cor. (Coef. Semi-Partial Determination) (Logit)	Frequency of Belief (%)
FGM Improves					
Hygiene	0.129949	0.039918	0.129168	0.040334	8.81
FGM Improves					
Social Acceptance	0.213896	0.104285	0.213037	0.107401	26.07
FGM Improves					
Marriage Prospects	0.023258	0.002613	0.023046	0.004569	3.49
FGM Preserves					
Virginity (no p.m. sex)	0.022283	0.020843	0.022211	0.021448	14
FGM Improves					
Men's Sexual					
Pleasure	0.009813	0.002036	0.009736	0.003244	0.8
FGM Is a Religious					
Necessity	0.202256	0.09844	0.201291	0.103696	11.7
no-FGM Means					
Fewer Medical					
Problems	0.029625	0.022699	0.029915	0.026978	25.36
no-FGM Avoids					
Suffering	0.0148	0.006108	0.014945	0.009206	20.2
no-FGM Improves					
Female Pleasure	0.01475	0.007931	0.014936	0.008949	9.22
no-FGM Improves					
Male Pleasure	0.005751	0.001859	0.005824	0.003123	4.58
no-FGM Accords					
with Religion	0.007576	0.00238	0.007678	0.002164	3.31

Two beliefs stand out as having substantial importance for decisions about FGM. The first is the belief that FGM provides social acceptance. The second is the belief that it is a religious necessity. The belief that FGM promotes social acceptance was the most important input to decision-making. This is shown by both metrics and across both the LPM and Logistic-Regression analysis frameworks. Under the logistic framework, the belief that FGM promotes social acceptance was able to account (alone) for more than 21% of the variation in mothers' decisions to practice FGM. Social acceptance also contributed a unique additional 10.7% to the variation explained by a model with all other variables. The belief that FGM is a religious necessity explained just-over 20% of the variation in practice decisions on its own. It also added 10.3% unique additional variation-explained to a model with all other variables. Results for the LPM framework were almost identical.

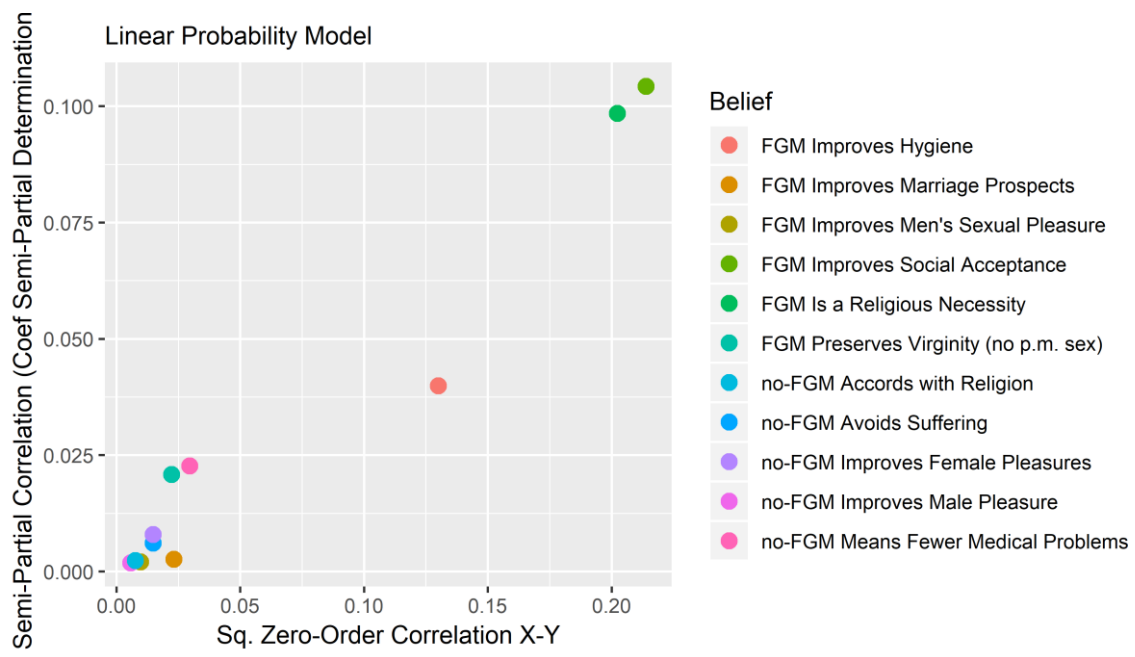


Figure 8: Influence of Beliefs on Decisions about FGM (LPM)

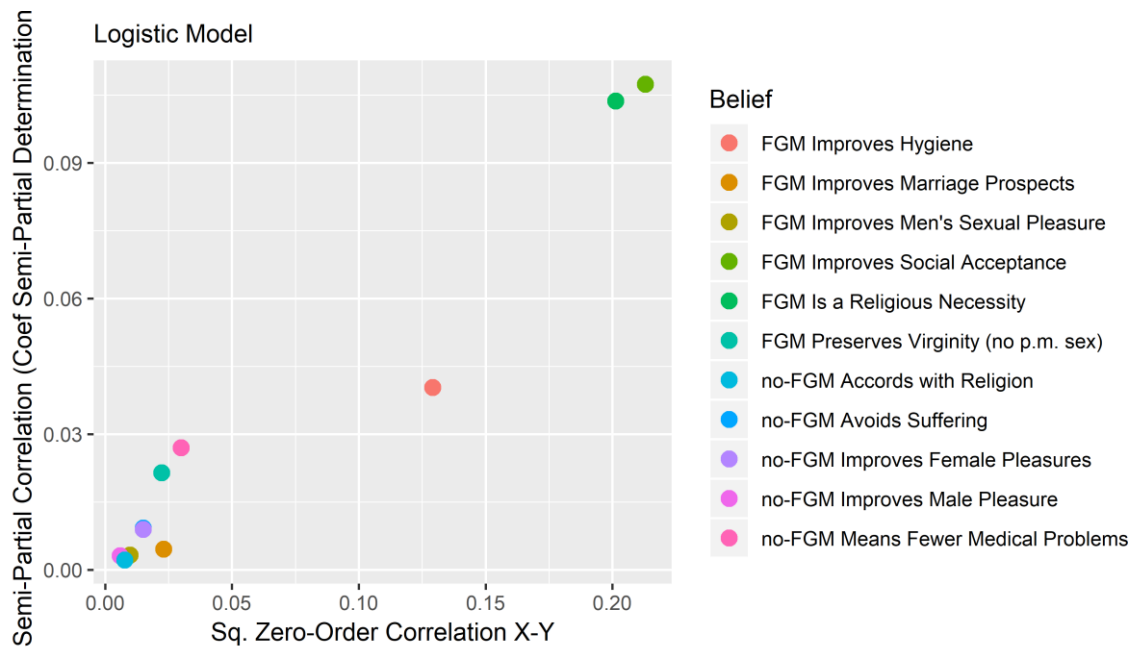


Figure 9: Influence of Beliefs on Decisions about FGM (Logistic Model)

As evident in Figures 8 and 9, only one other variable stands out as being a relatively important input to decisions. This was the belief that FGM improves hygiene. However, even this variable provided only a relatively weak contribution to decision-making. Apart from this, other beliefs appeared to make only a minor contribution to decision-making. Few other variables contributed more than 1% or 2% unique explained variance.

The remaining belief of key theoretical interest was that FGM improves marriage prospects. Yet, this explained just over 2% of the variation in decisions alone. Its inclusion in a larger model contributed less than 1% unique additional variation-explained. Other variables which might be connected to marriageability, including male sexual pleasure (for or against) and preservation of virginity, had similarly low importance scores.

As I argued above, belief frequency and belief importance evidently need not be the same. The belief that FGM preserves virginity was cited by 14% of respondents. Yet, this belief explained only 2% of the variation in decisions about the practice. By comparison, the belief that FGM is required by religion was cited by only 11.7% of respondents. Yet, this belief could account for almost ten times as much of the variation in decision-making.

Robustness of Methodological Assumptions

One assumption of the analysis is that the causal process linking beliefs to decisions is similar across the population. Although this is not to say that actors are identical. Actors may have different beliefs.

But, what if decision-processes were fundamentally variable at the individual level? In this case, a regression-based analysis couldn't be used. In such a scenario, the three hypotheses would also be meaningless. Each assumes a significant degree of continuity of decision-making across actors. Any assumption of 'perfect' homogeneity of decision-making is likely to be false. However, the analysis depends on there being meaningful regularities in decision-making.

Variation in decision-making processes at the individual level cannot be estimated directly. A separate model cannot be fit for each individual. However, some test of heterogeneity can be attempted. Specifically, the assumption that there are regularities in decision-making supported if results are relatively consistent across population strata (Kohler et al., 2019). In Appendix B2, I report the results of the analysis undertaken separately for each of 11 regions of Senegal covered by the 2005 DHS. This analysis shows that the conclusions of the analysis are relatively robust to geographic variation:

1. Out of the 11 regions, social acceptance is shown to make the largest unique contribution to the explanatory power of the model (SSPC) in 7 regions.
2. Social acceptance is among the two most important variables (by SZOC or SSPC) in 9 of the 11 regions.
3. Social acceptance has a higher importance score than belief in improved marriage-prospects on SZOC and SSPS in 10 out of 11 regions.
4. Belief in improved marriage prospects is not the most important variable in any region, on either metric.

Since the aim of the analysis is to explore a single causal process, the national representativeness of the data is of limited concern (c.f. Kohler et al., 2019). Due to the oversampling of rural communities, descriptive measures from the Senegalese DHS are nationally representative only when sampling weights are employed. Nevertheless, to reassure the reader that the results are not sensitive to the omission of sampling weights, Appendix B3 presents the results of the analysis using sample-weighted logistic regression. The results remain entirely consistent with the conclusions of this chapter.

Discussion

Empirical Support for the Three Hypotheses

The findings of this chapter unambiguously favoured the hypothesis of coordination over the other two. The belief that FGM improves social acceptance was found to be the most important input in decision-making. It was able to explain (alone) more than a fifth of the total variation in decisions about the practice. It also contributed the most unique additional variation-explained to a larger model.

One of the contentions of the informational influence hypothesis was supported by the analysis. This is the idea that beliefs about the intrinsic value of FGM play an important role in decisions. The belief that FGM is a religious necessity was shown to be especially important. However, the informational influence hypothesis also asserts no major role played by social beliefs. This assertion is contradicted by the analysis.

Finally, little support was found for the competitive decision-making hypothesis. This hypothesis asserts that decisions to practice FGM are motivated by a wish to gain a relative advantage in a competitive marriage market. However, beliefs that FGM improves marriage prospects, or that it improves men's pleasure, played little role in decisions.

Triangulation with Qualitative Research

Like any, this analysis is subject to a number of limitations (see below). These limitations caution against treating it as a final adjudication of the three hypotheses. Even so, extra weight is added to conclusions of the study because it triangulates with existing research.

The substantive conclusions of the analysis have a strong coherence with Shell-Duncan et al. (2011)'s qualitative findings. The study used a national sample collected in a different way from Shell-Duncan et al. (2011)'s localized qualitative samples. It also used a very different methodology to elicit insights about decision-making. Yet, both studies arrived at a similar conclusion. Both concluded that actors' decisions about FGM are not motivated by marriage competition. Both suggest that FGM practice is not merely a personal preference. Both suggest that actors face strong social pressures to practice FGM.

Limitations of the Empirical Analysis

There are limitations to both the measurement of beliefs about FGM and the measurement of FGM-related decisions.

Measures of belief lacked detail. Participants were only recorded as stating, for example, that FGM promotes social acceptance. Yet, no information was recorded about the form or extent of the relation between FGM practice and social acceptance. A simple list of possible advantages may omit considerable information about actors' perceptions of the costs and benefits of FGM. This may account for why the complete regression models only explained about half of the variation in decisions.

Decisions about FGM were measured in terms of the FGM status of mothers' children (or their intentions if no children had been cut). This assumes that mothers are the final decision-makers about their children. This will often, but not universally, be the case.

Equally importantly, mothers' decisions about whether to cut a daughter are only one part of participation in the practice of FGM. FGM participation may include ceremonial activities, payment arrangements, educational activities, norm enforcement, and so on (Eldin et al., 2018; Shell-Duncan et al., 2011). These wider activities will also involve other adults in the community. They may include husbands, grandparents, neighbours and so on (Eldin et al., 2018). I discuss this further in Chapter 5. As such, the outcome measure captures only a part of actors' decision-making about FGM participation.

Finally, the strength of the analysis is further limited by uncertainty about the timing of decisions. Mothers' decisions about FGM are inferred from the FGM status of their children. However, their children's mutilation will have occurred in the past. We have to assume (but cannot establish empirically) that mothers' beliefs at the time of the survey reflect their beliefs at the time of deciding the FGM status of their daughter.

A final limitation that I want to re-emphasize is that decision-processes have been tested only in the Senegalese context. It is perfectly possible that the other two hypotheses would be better supported in a different national context.

Conclusions

The aim of this chapter was to identify the best-supported theory about decision-making in FGM practicing communities. I used quantitative analysis, triangulated with qualitative research. I considered three possible hypotheses, each of which has, or could, act as a 'theoretical core' for a model of the dynamics of FGM.

Results supported the coordination hypothesis. This points to existing coordination models as the best-supported starting point for new modelling efforts. Here, I assert that the 'core'

of the theory is well supported. Yet, in the next chapter, I show that the ‘standard’ coordination model remains potentially flawed in a number of ways. Specifically, it is non-robust to a range of de-idealisations and elaborations. As I have argued throughout, the challenge of understanding the social dynamics of FGM has at least two components. First, adjudicating theory. Second, tackling the thorny problems of model design uncertainty.

Chapter 5: Exploring the Robustness of Coordination Models of FGM to De-Idealisations and Elaborations

Introduction

This chapter reports an exploratory robustness analysis of a ‘standard’ model of FGM as a social norm of coordination, which I define below. This activity is the next step in the overall modelling strategy of the thesis. The immediate aim is to identify credible and important design possibilities for coordination models of FGM. These ‘design possibilities’ are referred to as de-idealisations and elaborations because they involve making simple aspects of the model more realistic (de-idealisations) and adding elements to the model that have previously been omitted or abstracted-away (elaborations). Design possibilities are considered important if they can be shown to disrupt the key dynamics of the ‘standard’ coordination model (see below). This is a crucial step in managing model design uncertainty, since it involves seeking out and testing credible modifications of the model design (see Chapter 3). I refer to modified versions of a model as ‘model variants’.

The design possibilities considered were drawn from existing literature on the social dynamics of FGM, including qualitative research in Senegal (Dellenborg, 2004; Hernlund and Shell-Duncan, 2007; Shell-Duncan et al., 2011; Shell-Duncan and Hernlund, 2006), theoretical discussions (Bicchieri, 2005; Cloward, 2015; Mackie, 2000; Mackie and LeJeune, 2009; Mackie et al., 2015), evaluations of the Senegalese anti-FGM Tostan program (Diop et al., 2008; UNICEF, 2008) and existing coordination models of FGM (Efferson et al., 2019; Novak, 2016; Platteau et al., 2017; UNICEF, 2007). Although I focused on the national context of Senegal, I also considered research from other contexts where it was especially relevant. I focused on features of social dynamics suggested by these literatures that were (a) absent (or abstracted away) in the ‘standard’ coordination model (see below), (b) compatible with the ‘theoretical core’ of the model (see below) and (c) suggested to be important by the given accounts. See Appendix C1 for further discussion of the source material and my approach to it.

Where some credible modification to the design of the standard model is shown to disrupt its key dynamics, this is treated as evidence that the modification in question represents an important area of ‘uncertainty’ in the design of the model. Such a demonstration has important implications for further model development, and for the field as a whole.

Important ‘uncertainties’ identified in this chapter set the agenda for building a ‘general’ model in the next chapter. The general model in the next chapter aimed to encompass the different areas of uncertainty identified in this chapter. This facilitated, among other things, sensitivity analysis of the global ‘space’ of design possibilities, and targeted empirical calibration of the *most important* uncertainties in the design of the model.

Also, important ‘uncertainties’ identified in this chapter help to advance the agenda of the field as a whole. They show that when building coordination models of FGM in future, modellers should pay attention to *at least* all the elements of model design showed to be important in this chapter. To do otherwise would be to deliberately ignore aspects of social dynamics that one has reason to believe are important. In this sense, the findings of this chapter are cumulative – they add to the range of model design elements known to be important in the field. They, therefore, build on insights from past analyses. In turn, future analyses may build on these further by identifying other important design elements not considered here.

The Theoretical Core: Social Coordination

In the previous chapter, I conducted an empirical analysis of three ‘core’ theoretical hypotheses about FGM decision-making. I concluded that a hypothesis of social coordination is best supported for the Senegalese context – which the national context I focus on in this thesis.

Heuristically, the basic premise of this theory is simple. Actors’ decisions about FGM are interdependent and subject to coordination incentives (Mackie et al., 2015). If others practice FGM, actors are incentivised to do so as-well. If others abandon FGM, actors are incentivised to do so as-well (Bicchieri, 2005: 52)³³.

Here I define this in a more rigorous way for the purposes of formal modelling and establish some key terminology that I rely on throughout the remainder of the thesis.

³³ Some philosophical formulations of social norm theory treat decision-making as mediated by a combination of *empirical expectations* (whether you expect others to follow the norm) and *normative expectations* (whether you think others expect you to follow the norm or might sanction you for not doing so). Although the latter may depend on the former. From a modelling perspective, the standard practice has been to model the contingencies of decision-making as depending on the behavior of others (Bicchieri, 2005: 52) and therefore having the incentive structure of a coordination game (Bicchieri, 2005; Bicchieri et al., 2018; Mackie, 2018). See also the treatment of explicit norm enforcement in this chapter.

Each actor in the population has one or more *reference groups*. Reference groups are groups of other actors whose FGM activity (broadly defined) the actor responds to (Mackie et al., 2015). The standard modelling assumption has been that the *reference group* of an actor is all others in the local community.

All or *most* actors also have *conditional preferences* for practicing FGM (Bicchieri, 2005). In a modelling context, each actor prefers to *practice* FGM conditional on certain kinds of activities of their reference groups. Each actor also prefers to *abandon* FGM conditional on certain *other* kinds of activities in their reference group. The activities in question can be generically divided into *pro-FGM activity* (key-term) and *anti-FGM activity* (key-term). I refer to these two generically as *social pressures* (key-term). The ‘standard’ assumption in coordination models of FGM has been that the level of *pro-FGM activity* is equal to the proportion of actors in the reference group(s) who practice FGM (Novak, 2016; Platteau et al., 2017; UNICEF, 2007). Whereas *anti-FGM activity* is equal to the proportion of actors in the reference group(s) who do *not* practice FGM. Thus, *social pressures* depend on the proportion of others practicing FGM.

However, under social norm theory, perceived *social pressures* might also depend on things like explicit norm enforcement. This is *deliberate* social sanctioning of those not following the norm (i.e. FGM practice, Bicchieri, 2005; Mackie, 2018; Mackie et al., 2015; c.f. Hernlund and Shell-Duncan, 2007 and Shell-Duncan et al., 2011 for empirical examples).

In general, an actor following a social norm will be willing to practice FGM if pro-FGM activity in their reference group is high enough (and anti-FGM activity low enough), and willing to abandon it if anti-FGM activity in their reference group is high enough (and pro-FGM activity low enough)³⁴.

The Standard Coordination Model

The ‘standard’ model based on social coordination is a heterogeneous coordination model of the kind discussed in Chapter 2. Although the *game-theoretic n-person* social convention model is the archetypal model in the field (UNICEF, 2007), recent analyses have already convincingly demonstrated its non-robustness to one of its key simplifying assumptions: that all actors are equally willing to abandon FGM (Novak, 2016; Platteau et al., 2017). In light

³⁴ The standard assumption has been that pro-FGM activity is the complement of anti-FGM activity (if one is high, the other is low), because it is modelled as the proportion of actors practicing FGM. This need not necessarily be the case when norm-enforcement is involved. I discuss this further in Chapter 6.

of this, ongoing analysis in the field which is compatible with a social coordination theory of decision-making has assumed that preferences are *heterogeneous*, such that actors vary in their willingness to abandon FGM (e.g. Efferson et al., 2019). The coordination model with heterogeneous preferences is, therefore, the current ‘starting-point’ for coordination models in the field.

In some recent formal analyses (Efferson et al., 2019; Novak, 2016), explicit representation of the utility functions which drive actors’ decision-making has been abstracted away when analysing model dynamics, in favour of the generic representation of ‘thresholds’ which emerge from these functions. A threshold is typically a fixed proportion of others practicing FGM, above which the actor will do so as well. In my own modelling, I treat ‘threshold analysis’ as a useful heuristic tool, but I don’t abstract away the components of individual utilities that *create* an actors’ threshold.

Thresholds are a highly abstract concept that is difficult to measure (see discussion of Novak, 2016 in Chapter 7). And at the same time, recent analysts have highlighted the potential value of actually measuring the characteristics of communities which make coordinated abandonment of FGM viable (see Efferson et al., 2019, although they did not actually use empirical measures of key characteristics like the distribution of thresholds). Since thresholds are hard to measure, if we hope to understand communities’ readiness to change (under coordination dynamics), we will need to actually represent the components of decision-making that create different distributions of thresholds (see Chapter 7 for an example of the value of this for measurement). Thresholds are often treated as indicated by individual actors’ perception of the intrinsic value or cost of FGM (e.g. Efferson, 2015 – Supplement; Platteau et al., 2017 use the term ‘aversion’). However, this is only strictly true under certain kinds of decision-function. As I show in this chapter, issues such as variation in individual autonomy can create complex and non-intuitive relationships between the distributions of individuals’ perceptions of the value of FGM, and the distribution of ‘thresholds’ in the population.

The other significant problem with a reliance on the notion of ‘thresholds’ is that there are some ‘valid’ models of FGM as a social norm of coordination in which the decision-function of actors does not ‘reduce’ to a single threshold. Broadly speaking, this can occur when the level of pro-FGM activity is not the relative-complement of the level of anti-FGM activity. In Chapter 6, for example, I show a modelling scenario (involving norm enforcement) in which individual actors have multiple ‘thresholds’. As such, limiting the representation of decision-making to a single ‘threshold’ can, without justification, limit the range of model

designs that can be considered. Nevertheless, the concept of a threshold remains very useful for analysing the implications of *particular* decision-functions if and when they can be reduced to a simple threshold.

I implemented the ‘standard coordination model’ of FGM using a simple agent-based social simulation model (ABM) built in the NetLogo development environment. This and other simulation models used in this chapter are included in the DVD-ROM attached to this thesis.

The key terminology and formal notation developed here are crucial for the remaining chapters of the thesis.

This model simulates a local population of social actors, called *agents* who repeatedly decide to practice FGM, or not, in random sequential order. Their decision is based on the following utility-functions, based on UNICEF’s (2007) specification of the social convention model of FGM. Each agent i has two utility functions³⁵:

$$U(\text{abandon})_i = p_i \cdot -\hat{M}$$

$$U(\text{practice})_i = -H_i$$

Here, p_i is the proportion of other actors in the reference-group of actor i who practice FGM. So, p_i represents the level of *pro-FGM* activity in the reference-group of actor i . I use set-notation to represent the *group of* actors who make-up the reference-group of actor i : $\{\text{social-ref}\}_i$. In the standard model, $\{\text{social-ref}\}_i$ is the set of all other actors in the community. As such, p_i is defined as:

$$p_i = \sum_{j \in \{\text{social-ref}\}_i} \frac{d_j}{n(\{\text{social-ref}\}_i)}$$

Here one sums over the set of actors in the reference group $\{\text{social-ref}\}_i$. d_j is a binary variable (0 or 1) representing actor j ’s most recent decision to practice FGM ($d_j = 1$), or not ($d_j = 0$). Also, $n(\{\text{social-ref}\}_i)$ represents the number of actors in the reference group (a.k.a. the cardinality of the set).

\hat{M} is an arbitrary constant representing the maximum social cost for practicing FGM. Therefore $p_i \cdot -\hat{M}$ represents the expected *social cost* (key term) for abandoning FGM. A

³⁵ Note that here, like Novak (2016) I have dropped reference to benefits for unilateral practice, since the social competition hypothesis was rejected in Chapter 4. Also, the corresponding term \mathfrak{S} in the original social convention model actually had no impact on the dynamics of the model (see footnote 11).

crucial concept here is the *relationship between* social pressure (p_i) and social costs \widehat{M} . The standard assumption (also made here) has been that this relationship is *linear* (e.g. Platteau et al., 2017: 8). However, other assumptions about the *relationship between social pressures and expected social costs* are possible (see *Heterogeneity of Autonomy* below).

H_i is an individual variable in the interval $[0, \widehat{M}]$, representing the intrinsic cost that actor i attributes to FGM. To maintain the conformist properties of the model, H_i is defined as:

$$H_i = q \cdot \widehat{M}$$

$$q \in [0,1]$$

Where q is a beta-distributed random variable.

In each time step of the simulation, agents decide to practice FGM or not. Actors' decisions between practicing or abandoning FGM depend on the expected utility for each choice at the point at which the decision is made:

$$\begin{cases} U(\text{practice})_i \geq U(\text{abandon})_i \Rightarrow d_i = 1 \\ U(\text{practice})_i < U(\text{abandon})_i \Rightarrow d_i = 0 \end{cases}$$

The dynamics of the simulation occur through the agents repeatedly applying this decision-rule, which takes into account the most recent choices of the other agents.

Just as in analytical heterogeneous threshold models, all agents in the standard model have a threshold. This threshold 'emerges' from their decision-process. Formally the 'threshold' of agent i , which is denoted p_i^* , will be the value of p_i for which:

$$U(\text{practice})_i = U(\text{abandon})_i$$

Which is when:

$$p_i \cdot -\widehat{M} = -H_i$$

Substituting in the definition of H_i , we find that p_i^* is equal to q , the random component of H_i :

$$p_i^* = \frac{H_i}{\widehat{M}} \Rightarrow p_i^* = \frac{(q \cdot \widehat{M})}{\widehat{M}} \Rightarrow p_i^* = q$$

As such, this ‘standard’ model will have nearly identical³⁶ dynamics to an analytical heterogeneous threshold model with q distributed thresholds. Based on UNICEF’s (2007: 49) informal speculation about the shape of the distribution of actors’ willingness of abandon FGM, I assume that in the ‘standard’ coordination model, q is beta-distributed with a mean of between 0.4 and 0.6 and a variance of between 0.001 and 0.031.³⁷

In my analysis, the agents in the model are construed as individual people in FGM practicing communities. Some formal analyses of FGM have treated agents as households, families, or parental couples (Novak, 2016; Platteau et al., 2017; UNICEF, 2007) others treat the players as individuals (Coyne and Coyne, 2014). I prefer the latter. Treating agents as households prevents the analysis from considering a range of important issues. These include social pressures exerted between household members (Hernlund and Shell-Duncan, 2007), the role of authoritative individuals within communities (Cloward, 2015) and the reality that community-level interventions target individuals, not households (Diop et al., 2008; UNICEF, 2008). The ‘community’ itself is construed as a local population of actors who coordinate on FGM decisions. Depending on the assumed importance of marriage (vs other social pressures) in coordination incentives, this might consist of a group (or network) of intermarrying families (Mackie, 2017), or a group of actors socially connected actors in a territorial community (Shell-Duncan et al., 2011). Given studies showing a lack of strong connection between FGM and marriage in Senegal (Hernlund and Shell-Duncan, 2007; Shell-Duncan et al., 2011), the latter may be more appropriate. Once again, neither construal is integral to social norm theory, and few parts of the analysis depend on the distinction (although see discussion of ‘*Decision Maker*’ *Reference Groups* below).

The decisions of agents to practice FGM are broadly construed in terms of participation (versus refusal to participate) in FGM related activities (Coyne and Coyne, 2014). This includes contributions to decisions about whether to cut a girl in one’s family, as well as the many other actions that may be involved in the tradition, including arranging the event, paying the circumciser, participating in celebrations, providing the ‘training’ of girls that may proceed cutting, and sanctioning uncut women (Shell-Duncan et al. 2011, Eldin et al., 2018).

³⁶ Some minor stochastic effects can occur because a finite population of agents is used, and decision-making is random-sequential.

³⁷ This produces a roughly symmetric unimodal distribution with values bounded at 0 and 1. See Chapter 6 for further discussion of the beta-distribution as a general way of representing bounded parameter distributions.

These activities can involve a wide range of members of the community, including immediate family (parents and siblings) as well as grandparents, uncles and aunts, friends and neighbours (Eldin et al., 2018). Limiting the construal of decision-making (and social coordination pressures) to narrowly reflect the FGM outcome for specific girls in a particular household would neglect the fact that individuals in the community face a range of situations in which they must to decide whether to participate in the practice, or risk the social consequences of unilaterally deviating from the shared tradition (Bicchieri and Mercier, 2014; Cloward, 2015; Shell-Duncan et al., 2011).

Key Dynamics of the Standard Coordination Model

Any changes to the design of a formal model of social dynamics will be likely to affect its dynamics in *some* way. However, the question for robustness analysis is whether the *important* dynamics are *substantially* altered by plausible modifications of its design.

The important dynamics of the standard coordination model can be divided into dynamics of *persistence* and dynamics of *interventions*.

Dynamics of persistence in the standard coordination model concern the *rates* of FGM practice that can persist in the model (defined as the proportion of individuals in the local community practicing FGM). This, in turn, provides an indication of the levels of FGM that are expected to be observed in real communities if the model is correct (Efferson et al., 2015; Mackie, 2017). The original social convention model, for instance, predicted that only rates of FGM near to 0 or 1 would be stable, a prediction which Efferson et al. (2015) translated into an empirical test of the model (they claimed it failed, though see Mackie, 2017).

Important dynamics of intervention in the standard coordination model encompass a range of issues all related to the expected success of community-level interventions aiming to end FGM. In discussions of social dynamics, interventions are characterised as having at least two aims, (a) to create a stable reduction in the rate of FGM (Mackie, 2017) and (b) to generate positive spillover effects (Efferson et al., 2019, 2015; UNICEF, 2007). Positive spillover effects occur when the effect of the intervention spreads from participants to non-participants as a result of coordination dynamics. In theory, positive spillovers can allow a ‘limited’ intervention to create widespread change (Efferson et al., 2015, 2019).

The standard coordination model makes a number of important predictions about the circumstances under which these aims will succeed or fail. These ‘circumstances’ include the size of the intervention, the extent of preference change created by the intervention, and the

targeting bias of the intervention (Efferson et al., 2015, 2019; Platteau et al., 2017). ‘Targeting bias’ in this context means the tendency of the intervention to preferentially select those ‘reluctant’ to abandon FGM (low H_i values, hereafter ‘reluctant actors’), or those ‘willing’ to abandon FGM (high H_i values, hereafter ‘willing actors’)³⁸.

The specific dynamics of the standard coordination model with respect to these issues are described in this chapter. However, the key point of the chapter is to identify credible modifications to the design of the model which *disrupt* these predictions. To do this, simulated experiments were used to measure the persistence and intervention dynamics of the standard coordination model, as well as variants of the model (i.e. with de-idealisations and elaborations). The design of these experiments was as follows.

Dynamics of Persistence Experiment (330 simulations):

1. For conditions $x \in \{0, 0.1, 0.2, \dots, 0.8, 0.9, 1\}$, each repeated 30 times:
 - a. Simulate a population of 800 agents³⁹
 - b. A randomly selected proportion x of agents begin the simulation practicing FGM ($d_i = 1$)
 - c. The simulation is run for 30 time-steps⁴⁰ (each agent decides to practice FGM or not in a random sequential order, 30 times).
 - d. The proportion of agents in the simulation practicing FGM is measured

Here, I present an example of applying the “Dynamics of Persistence” experiment to the standard coordination model (with ‘normal’ distribution of H_i). The results of this experiment are shown in Figure 10:

³⁸ Efferson et al. (2019) use the terms ‘resistant’ and ‘amenable’ actors. I prefer ‘reluctant’ and ‘willing’, which are the terms development by Shell-Duncan and Hernlund (2006) in their empirical work in Senegal and The Gambia.

³⁹ This is largely arbitrary (and doesn’t affect the results), although Mackie and LeJeune, (2009: 13) suggest the average rural village size in Senegal is around 800

⁴⁰ This is sufficient for the models to reach a ‘steady state’ in almost all cases, except in the deliberate case where explicit ‘temporal’ features are added to the model which can delay change (see discussion of ‘decision-maker’ reference groups in this chapter).

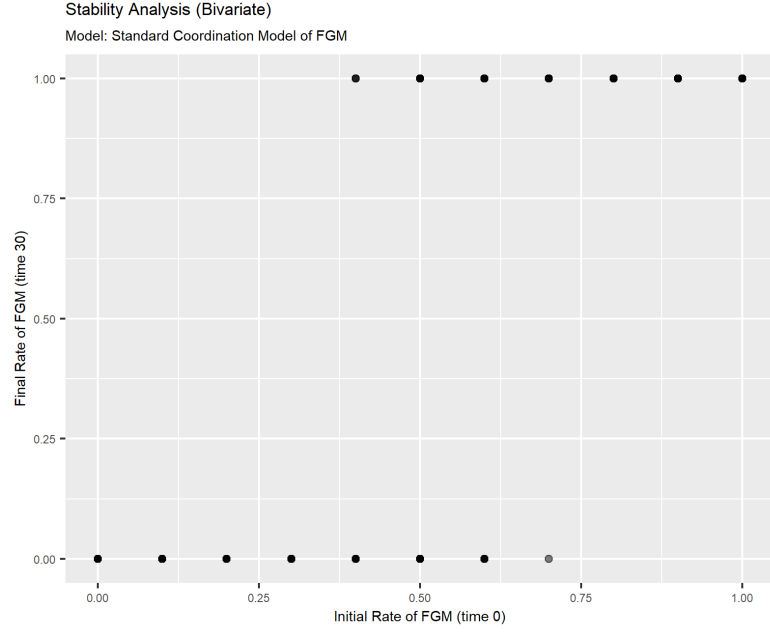


Figure 10: Persistence Dynamics of the Standard Coordination Model

Figure 10 shows the initial rate of FGM in the model (x-axis), as well as the final rate of FGM after 30 time-steps (y-axis). We can see that although the standard coordination model was initialised with rates of FGM *between* 0 and 1, all final rates of FGM were *either* 0 or 1. The average final rate of practice across the whole experiment was 0.5. Thus, in the standard coordination model, there were no interior rates of FGM, and there was also no bias towards practice or abandonment, given a uniform distribution of initial starting rates (the x values). As I discuss in Chapter 2, these dynamics can be understood in terms of the distribution of thresholds in the model (which depend on the distribution of the H_i attribute of agents, see above). Since this distribution of thresholds is assumed to be roughly ‘normal’ (symmetrical and unimodal) in the standard model, interior rates of FGM are driven to 0 or 1.

Dynamics of Intervention Experiment (3300 simulations):

1. For conditions $z_1 \in \{0.1, 0.2, \dots, 0.8, 0.9, 1\} \times z_2 \in \{0, 0.1, \dots, 0.9, 1\} \times z_6 \in \{\text{reluctant}, \text{random}, \text{willing}\}$, repeated 10 times, where z_1 is the proportion of the population targeted in the intervention, z_2 is the degree of preference change created by the intervention (‘educational strength’, see below) and z_6 is the ‘targeting bias of the intervention’:
 - a. Simulate a population of 800 agents, all of whom practice FGM ($d_i = 1$)
 - b. Target z_1 of actors according to targeting rule z_6 . If $z_6 = \text{reluctant}$, target actors in ascending order of H_i . If $z_6 = \text{willing}$ target actors in descending order of H_i .

- c. Apply the educational effect of the intervention z_2 , defined as an increase in the H_i values of targeted agents ($H_i \rightarrow \hat{M} - [(\hat{M} - H_i) \cdot (1 - z_2)]$, see section below)
- d. Simulate the intervention process, including ‘coalition formation’ (see section below)
- e. The simulation is run for 30 time-steps (each agent decides to practice FGM or not in a random sequential order, 30 times).
- f. Measure the rate of FGM in the population.

The way the intervention process is simulated is discussed further in a later section. It requires further discussion because the representation of the intervention process is *itself* is an important area of non-robustness in the design of the standard coordination model. The key point to communicate at this stage is that one can define a range of types of intervention, in terms of size, preference change and targeting bias. In turn, the outcomes of these interventions predicted by a given model design can be measured using the simulated experiment above.

Recognised Areas of Non-Robustness

Recent analyses of coordination models of FGM have been framed as making ‘new’ positive predictions based on improved model design, rather than as robustness analyses of model designs. However, because they have involved model designs which deviate from the ‘standard’ coordination model, they still provide important insights into areas of non-robustness. These insights, which are built upon in this chapter, are outlined here.

The most important set of insights concern the sensitivity of the important dynamics of the standard coordination model to assumptions about the distribution of thresholds in the population. At-least five distinctive ‘types’ of threshold distribution can be identified, each which has a substantial effect on the important dynamics of the model. These effects are summarised in Table 3. In the table, I focus on three issues:

- a. Rates of FGM that are predicted to be stable or unstable
- b. Predictions about whether interventions will be able to create ‘stable change’ (i.e. stable reductions the rate of FGM)
- c. Predictions about whether interventions will be able to create positive spillovers

In the table, the dynamics related to interventions assume that interventions only affect behaviour, not preferences (i.e. the distribution of thresholds doesn’t change). As I discuss below, if the intervention also changes preferences, the situation is complicated further.

Table 3: Types of Threshold Distribution and their Effects on Dynamics

Distribution Type	Dynamics of Persistence	Dynamics of Intervention (No preference change): Stable change	Dynamics of Intervention (No preference change): Positive spillovers
‘Normal’ (central unimodal)	Only exterior ($\{0,1\}$) rates of FGM will be stable, other rates of FGM will be driven down or up to these rates, depending on whether they are above or below a central ‘tipping-point’.	Stable change will occur if the intervention is large enough to cross the tipping point.	Large positive spillovers leading to complete abandonment will occur if the intervention crosses the tipping point.
‘Uniform’ (high variance, non-modal)	All rates of FGM will be stable.	Stable change will occur for interventions of any size.	Positive spillovers will never occur (without preference change).
‘U-modal’ (central bimodal)	Only interior rates of FGM will be stable.	Stable change will occur for small interventions, but under larger interventions, some actors will return to practicing FGM.	Small interventions will achieve small positive spillovers, larger interventions will not.
‘J-modal’ (left-skewed, unimodal)	Only rates of FGM near to universal practice will be stable.	No stable change can occur unless virtually the whole population is involved.	No positive spillovers will occur.
‘L-modal’ (right-skewed, unimodal)	Only rates of FGM near to universal <i>abandonment</i> will be stable.	Stable change will occur for any intervention size.	Any intervention size will result in large positive spillovers leading to complete abandonment.

Graphical illustrations of each of the distributions described in Table 3 are given in Chapter 2. Readers can apply the graphical technique for heterogeneous threshold models described in Chapter 2, to see why the corresponding dynamics occur.

The above examples assumed that the intervention had no effect on preferences. This is meant to be equivalent to the intervention merely acting as a ‘coordination device.’ The classic example would be if the intervention *only* involved organising a public declaration by a group in the community that they were going to abandon FGM (Platteau et al., 2017).

However, further subtleties are added when the intervention is also assumed to change the *preferences* of actors (i.e. the thresholds of individuals targeted by the intervention). Broadly speaking, stable change is always possible with sufficiently large interventions that sufficiently change actors’ thresholds. Intuitively, the less conducive a threshold distribution is to reductions in the rate of FGM *without* preference change, the greater the preference change needed for the intervention to be effective.

A further important result, in the context of interventions which produce preference change, is due to Efferson et al. (2015, 2019). This result is that if the intervention increases the threshold of all targeted actors to 1 (so that they will only practice FGM if *all* others do), targeting bias can play a significant role in whether positive spillovers occur. Specifically, given the above assumptions, and *if* the threshold distribution is roughly symmetric, simulated interventions which target reluctant actors (those reluctant to abandon FGM) will produce much greater positive spillovers than simulated interventions which target willing actors (those willing to abandon FGM). Also, regardless of the distribution of thresholds, simulated interventions (given the above assumptions) are guaranteed to do no *worse* by targeting supporters of FGM than they would by targeting opposers of FGM, or targeting actors at random.

However, this result *has itself* been shown to be non-robust to analyses conducted by Efferson et al. (2019). Their analyses show that advantages (in terms of positive spillovers) from targeting those reluctant to abandon can be heavily attenuated if (a) those who are reluctant are also less likely to respond to the intervention or (b) the population of actors is ‘fragmented’ by a low-density social network. Furthermore, the advantages (in terms of spillovers) can actually be eliminated if the interaction occurs across a social network with strong homophily of preferences (this means actors with similar views on FGM are more likely to be connected to one another).

These results show that the persistence dynamics of the standard coordination model are highly sensitive to the way that threshold distributions are represented, as are dynamics related to achieving stable reductions in FGM through intervention. Dynamics related to achieving positive spillovers are also sensitive to assumptions about the structure of social interactions (i.e. social network structure), especially homophily of thresholds (where actors with similar thresholds are more likely to be connected).

These represent important areas of uncertainty in the design of coordination models of FGM because the choices made about designing these elements (including the choice to omit them) can have substantive effects on the important dynamics of the model.

The analyses undertaken in this chapter build on, and expand, these results in a number of respects.

First, I consider a wider range of possible modifications to the standard coordination model, drawn from the theoretical and empirical literature. These include:

- Incorporating additional kinds of heterogeneity into decision-making processes
- Changing the definitions of social pressure (e.g. to include explicit norm-enforcement)
- Incorporating additional kinds of social structure (e.g. by constraining the composition of actors' reference-groups)
- Incorporating additional elements into the representation of the intervention process

Second, for each of the modifications considered, I assess the effects on a range of dynamics, including persistence, stable reductions in the rate of FGM *and* positive spillovers (whereas Efferson et al., 2019, only addressed positive spillovers⁴¹). Among other things, my analysis shows that levels of FGM *persistence* can also be strongly affected by social structures (i.e. social networks). This, in turn, has important implications for the evaluation of the standard coordination model.

In the following sections of this chapter, I outline the model design modifications that I considered. I also summarise the results of the robustness analyses.

⁴¹ In fact, since they assume that interventions transform actors into unconditional abandoners of FGM (see below) changes of behavior are always stable in their analysis.

Reports showing a full set of results and graphs for each model design considered are included in the DVD-ROM attached to the thesis. Nevertheless, I begin with a full ‘worked example’, looking at current formal representations of the intervention process itself. Based on these considerations, I define a new ‘standard’ intervention scenario which better captures real anti-FGM intervention processes at the community-level. I also show that this elaboration is important, in that it undermines some of the established results noted above.

‘Worked Example’: Representing the Intervention Process

In past formal analyses of FGM as a social norm of coordination, at-least two distinctive representations of community-level interventions can be identified. Under some representations, the intervention has no effect on preferences (e.g. ‘expressive effect’ in Platteau et al., 2017, although they also consider alternatives involves preference change) under others, it completely alters preferences, such that participants become maximally opposed to FGM (Efferson et al., 2019). The first case is meant to correspond to a situation in which the intervention organises a public declaration (or other ‘coordination device’) to allow actors to publicly commit to abandoning FGM. In a simulation context, this effect can be achieved by changing the behaviour of agents ($d_i \rightarrow 0$) for a single time-step. The second case is meant to correspond to a situation in which the intervention has a strong educational component, and then public declarations occur after an education process. In a simulation context, this effect can be achieved by setting targeted agents’ perceived cost of FGM (H_i) to the maximum value (\hat{M}) and *also* changing their decision to practice FGM ($d_i \rightarrow 0$). In this case, all *targeted* actors are guaranteed to stop practicing FGM permanently, because their threshold is set to 1 and they abandon together.

In some ways, the latter is probably a more realistic representation of the intervention process, since community-level interventions to end FGM have always included an educational component (Diop et al., 2008; Mackie, 2000; UNICEF, 2007; UNICEF, 2008). However, the assumption that the educational effect is always strong enough to set all participant’s thresholds to 1, is optimistic, to say the least. However, we can generalise over the two ways of representing the intervention process by parameterising the educational effect of the intervention. To do so, we can define the educational effect of the intervention to be:

$$H_i \rightarrow_{intervention} \hat{M} - [(\hat{M} - H_i) \cdot (1 - z_2)]$$

Here, the educational effect of the intervention is controlled by a parameter z_2 which is between 0 and 1. When z_2 is 1, all participants become maximally opposed to FGM ($H_i \rightarrow \hat{M}$). When z_2 is 0, all participants retain their original perception of FGM ($H_i \rightarrow H_i$). When z_2 is 0.5, the distance between \hat{M} and H_i is halved for each participant, and so on⁴². I refer to z_2 as the educational strength of the intervention.

Using the ‘Dynamics of Intervention Experiment’ (see above), we can explore the full persistence dynamics of the standard coordination model across different intervention sizes, and strengths, and with different targeting biases. The results are shown in Figures 11-13. There is one figure for each kind of targeting bias (reluctant – low H_i , willing – high H_i , and random). The x-axis of each figure shows the educational strength of the intervention; the y-axis shows the size of the intervention (proportion targeted). The size of the points shows the average rate of FGM after the intervention (over ten simulations) and after 30 time-steps of coordination. The colour of the points shows the size of the spillover. This was calculated as the average absolute difference between the size of the intervention, and the proportion of the population who had abandoned FGM at the end of the simulation. As such, spillovers could be positive or negative (i.e. if change was not stable).

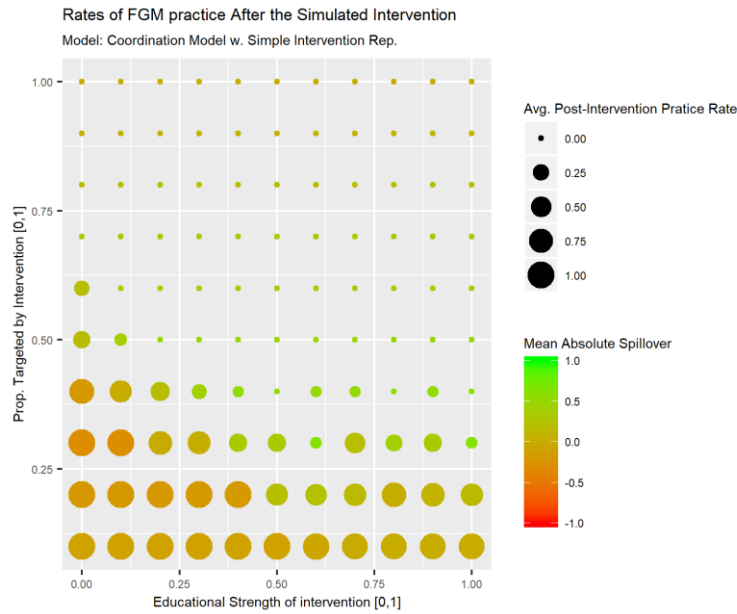


Figure 11: Simulated Intervention Effects Under Simple Intervention Representation (Random Targeting)

⁴² Note that when z_2 is in the interval $(0,1)$ the rank ordering of participant's preferences is maintained.

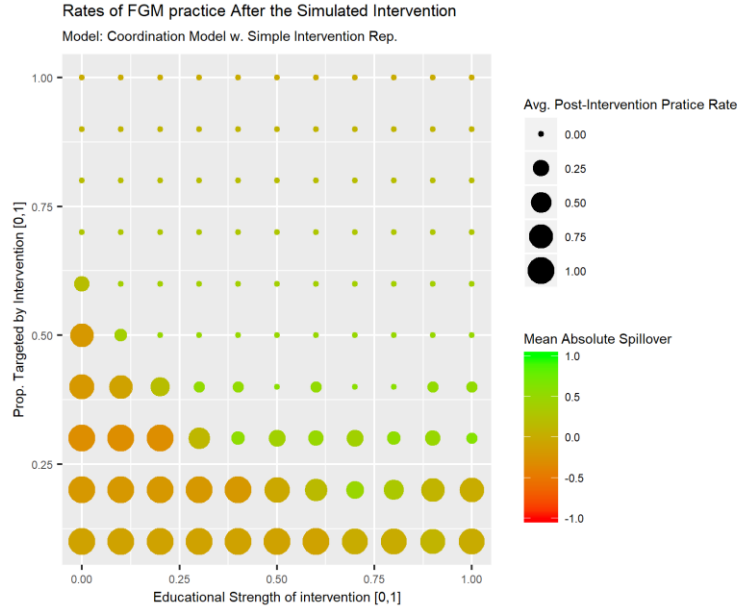


Figure 12: Simulated Intervention Effects Under Simple Intervention Representation (Targeting Those Most Reluctant to Abandon)

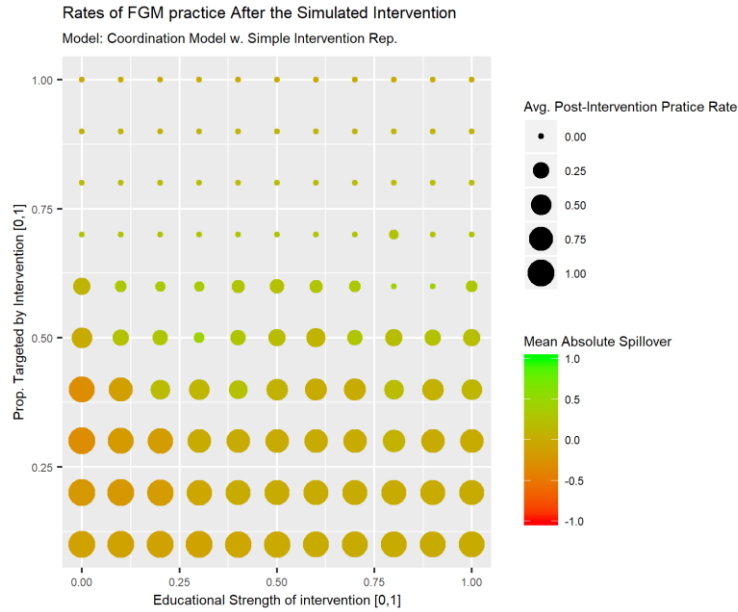


Figure 13: Simulated Intervention Effects Under Simple Intervention Representation (Targeting Those Most Willing to Abandon)

One can see that the expected results for a heterogeneous threshold model with ‘normal’ distribution of preferences were observed. Moderately large interventions were able to achieve substantive reductions in FGM and large positive spillovers. When the intervention included large preference changes, these changes were achieved with smaller interventions. The claimed advantages of targeting those reluctant to abandon FGM are clearly visible and appear even at low levels of preference change (i.e. Efferson et al.’s, 2019, assumption that preference change is extreme is not actually required). Intuitively, when targeting those

willing to abandon FGM, the intervention does not really benefit from preference change, since these actors are already very willing to abandon.

All of this suggests that the simulations ‘line up’ with past analytical results. However, there is a significant problem with this kind of simulation of the intervention process. It assumes that actors who participate in the intervention automatically transition to publicly abandoning FGM (temporarily or otherwise)⁴³. Yet in real community-level interventions, this transition is not automatic. Instead, as Mackie and LeJeune (2009), UNICEF (2010) and Mackie (2017) note, the initial stage of the community intervention involves building a group who are “*conditionally* committed to abandon if enough others do” [emphasis added] Mackie (2017: 39). In other words, individuals’ decisions to publicly abandon FGM are conditional on there being enough others willing to abandon at the same time (Mackie and LeJeune, 2009: 11).

There is a process of *coalition formation* which moderates the transition from the educational part of the intervention to the decision to abandon FGM. I incorporated this into the formal representation of the intervention process as follows. As before, the intervention targets some proportion of the local population (z_1) and educates them about the harmful effects of FGM (z_2). These participants represent an *initial coalition* (key term) of actors ready to abandon FGM. However, then, rather than ‘automatically’ abandoning FGM, these initial coalition members make an assessment of whether they would prefer to abandon FGM, conditional on all other participants doing the same. If some actors would *still* prefer to practice FGM even *if* all other coalition members abandon it (e.g. because they attribute a low intrinsic cost to FGM), then they leave the coalition. Those remaining in the coalition then *reassess* whether they would *still* prefer to abandon, given the smaller coalition, and so on. This process continues until the coalition stabilises into a *final coalition* (key term), or collapses entirely (I refer to this outcome as ‘coalition collapse’⁴⁴).

It turns out that introducing this feature into the simulated intervention process can have a substantial impact on the dynamics of the standard coordination model. Figures 14-19 show the results of repeating the ‘dynamics of intervention’ experiments with the standard coordination model, including ‘coalition formation’ as an interim step. Figures 14, 16 and 16

⁴³ And this is guaranteed to be permanent if the educational effect is extreme enough.

⁴⁴ There is evidence that dynamics of this kind can occur, UNICEF (2010: 15) notes, for example, that some interventions participants were unwilling to make public declarations of abandonment because of “fear of public criticism or social exclusion”.

have the same format as Figures 11-13 shown above. In Figures 15, 17 and 19, the meaning of the size of points is changed, so that the size reflects the size of the ‘final coalition’ that formed as a result of the intervention (prior to coordination dynamics).

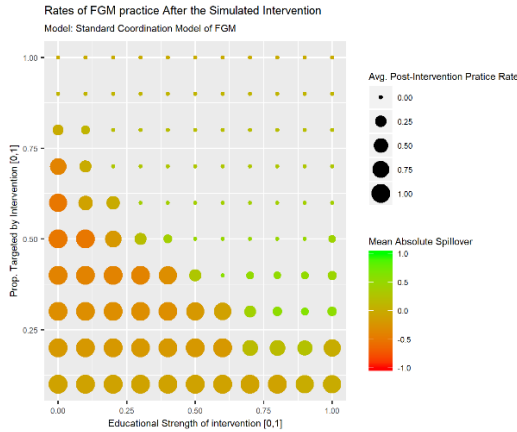


Figure 14: Simulated Intervention Effects Under Coalition-based Intervention (Random Targeting)

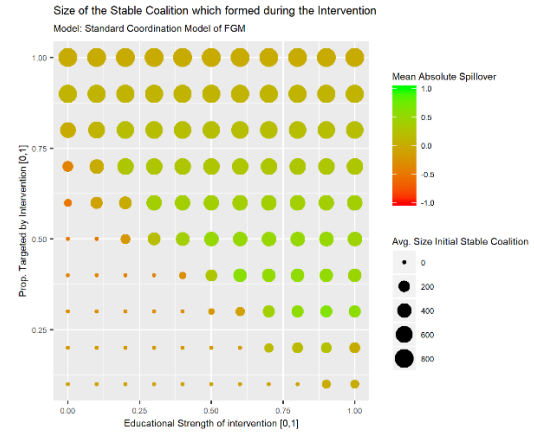


Figure 15: Coalition Formation Under Coalition-based Intervention (Random Targeting)

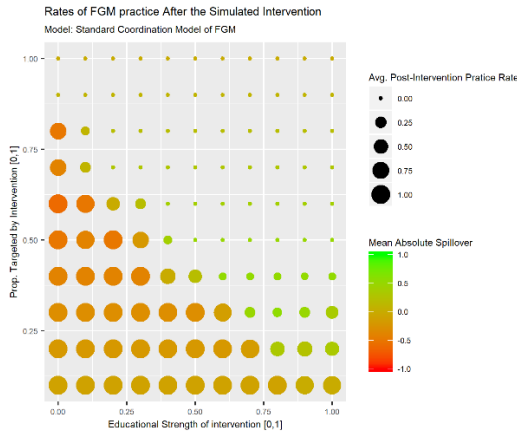


Figure 16: Simulated Intervention Effects Under Coalition-based Intervention (Targeting Reluctant Actors)

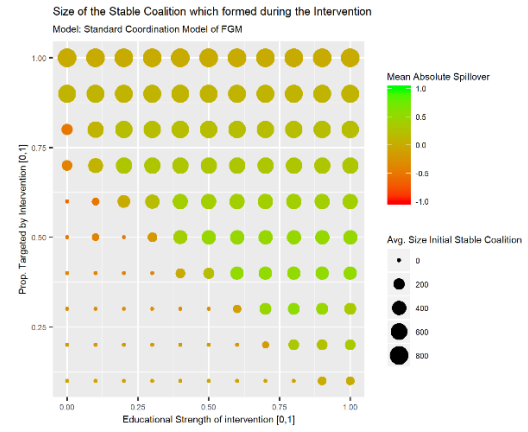


Figure 17: Coalition Formation Under Coalition-based Intervention (Targeting Reluctant Actors)

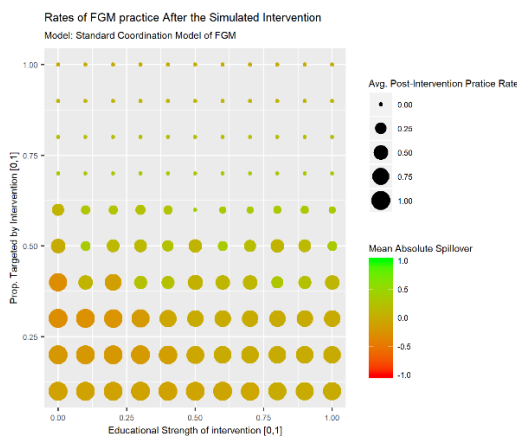


Figure 18: Simulated Intervention Effects Under Coalition-based Intervention (Targeting Willing Actors)

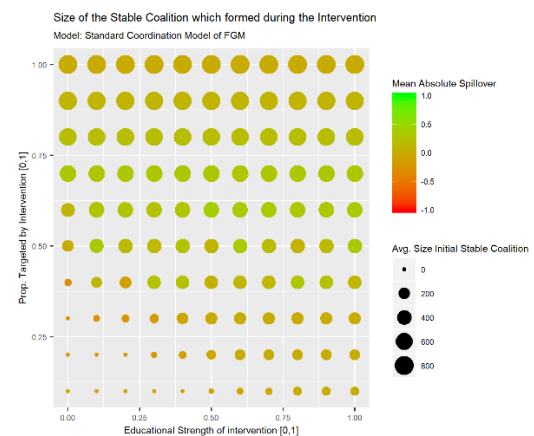


Figure 19: Coalition Formation Under Coalition-based Intervention (Targeting Willing Actors)

Incorporating a coalition formation process created an important change in intervention dynamics. The advantages of targeting supporters of FGM were significantly undermined.

When targeting reluctant actors, a wide range of interventions of small size or weak educational strength failed to produce any reduction in the rate of FGM. This occurred because the initial coalition of participants collapsed. By contrast, when targeting willing actors, coalition collapse was less likely to occur. The intuition underlying this is that willing actors are more likely to find smaller coalitions ‘sufficient’ for them to prefer to abandon FGM.

This finding shows a new way in which the established result: that targeting reluctant actors leads to a greater reduction in the rate of FGM, can be non-robust to de-idealisation of existing model designs. In fact, reflection on the implications of assuming that actors ‘automatically’ abandon FGM suggests that the established result may not just be fragile and simplified, but starkly wrong. Without a coalition formation process in which a local subset actor makes a *choice* to abandon FGM or not, we are forced to make one of two unacceptable assumptions. Either interventions are always so effective that participants will always prefer to unilaterally abandon FGM. Or, some interventions ‘automatically’ lead actors to abandon FGM despite being too small to actually persuade them that abandonment is in their interests. Both assumptions are implausible as general principles

Given the significant logical problems with assuming that interventions ‘automatically’ lead participants to abandon FGM, I treat coalition formation as a standard part of the intervention process. All subsequent simulations of the intervention process in the thesis include coalition formation. Since the dynamics of the standard coordination model are sensitive to the inclusion of the coalition formation process, I also repeat analyses of each of the four other threshold distributions noted in this chapter (but this time, with coalition formation included in the intervention process). Intervention dynamics remain highly sensitive to these different distributions, but not always in the way expected under more simple representations of the intervention process.

Modifying the Decision-Making Process

Past formal analyses of FGM which have included an explicit decision process (i.e. including explicit utility functions) have tended to make simplifying assumptions about the costs and benefits that actors face, or have obfuscated more complex possibilities by relying on the notion of thresholds. Here I consider a range of credible elaborations of the explicit decision-process used in the standard coordination model, based on the theoretical and empirical literature, as well as alternative assumptions found in existing model designs.

Model Variant 1: FGM as a Subjective Benefit

Formal analyses of FGM have typically assumed that FGM is an *intrinsic* ‘cost’. In other words, all other things being equal, actors would prefer to abandon it. Thus, although actors may vary in the *magnitude* of the cost that they associate with FGM, they all view it as intrinsically costly. Following the assumptions of the social convention model (UNICEF, 2007), the standard coordination model also assumes that actors continue to practice FGM because they face *social* costs for unilaterally abandoning it. Yet, empirical research (including my own in Chapter 4) clearly indicates that despite the *objective* medical harm of FGM some actors in Senegal may *subjectively* view FGM as intrinsically beneficial, including because they view it as a religious requirement or spiritual advantage (Dellenborg, 2004; UNICEF, 2013).

This suggests that heterogeneity of decision-making may extend not only to variation in the perceived cost of FGM but to some actors viewing FGM as intrinsically *beneficial*. We can represent this possibility by modifying the distribution of H_i represented in the standard coordination model:

$$H_i = (-\hat{M} \cdot s_2) + q \cdot (1 + s_2) \cdot \hat{M}$$

Here, the distribution of intrinsic ‘cost’ of FGM (H_i) is extended by the s_2 parameter to include potential intrinsic benefits of FGM, which scales the maximum perceived intrinsic *benefit* of FGM in the population. The original definition of H_i is a special case of this equation, where s_2 is equal to 0. If s_2 is set to 1, the possible range of H_i values is extended from $-\hat{M}$ to \hat{M} , such that the perceived *benefit* of FGM can be just as large as the perceived *cost*. Depending on the value of s_2 , this can have substantial effects on the decision-making process of actors. First, it can change the location of their ‘thresholds’:

$$p_i^* = \frac{H_i}{\hat{M}} \Rightarrow p_i^* = \frac{(-\hat{M} \cdot s_2) + q \cdot (1 + s_2) \cdot \hat{M}}{\hat{M}} \Rightarrow p_i^* = q(s_2 + 1) - s_2$$

In general, if s_2 is greater than 0, actors’ thresholds for practicing FGM will go down (they will be more willing to practice FGM).

Second, if other assumptions about social influence are unchanged, it can create a subset of actors who unconditionally practice FGM⁴⁵ (they have a threshold less than 0). This will happen whenever q is less than $\frac{s_2}{s_2+1}$:

$$q(s_2 + 1) - s_2 < 0 \Rightarrow q < \frac{s_2}{s_2 + 1}, s_2 \in [0,1]$$

Table 4 illustrates some of the possible effects of allowing subjective benefits of FGM. It shows three thresholds (based on q) in the standard coordination model, followed by the effect on these thresholds with increasing values of s_2 . To explore the potential impact of representing FGM as having ‘intrinsic value’ to some actors in the standard coordination model, simulated experiments were undertaken using an s_2 value of 0.4

Table 4: Effects on Thresholds of Representing FGM as an Intrinsic Benefit

q / Original Threshold	Threshold if $s_2 = 0$	Threshold if $s_2 = 0.5$	Threshold if $s_2 = 1$
0	0	-0.5 (Unconditionally Practice)	-1 (Unconditionally Practice)
0.5	0.5	0.25	0 (Unconditionally Practice)
1	1	1	1

Model Variant 2: Symmetric Social Pressures

In the standard coordination, based on assumptions expressed in UNICEF (2007), coordination incentives only apply to the decision to *abandon* FGM. Actors deciding to *practice* FGM don’t experience social pressure. They only experience the intrinsic cost of FGM practice. However, some other models have assumed that social pressures are symmetrical, such that actors also face social costs for practicing FGM while others abandon it (e.g. Novak, 2016). These costs to FGM practitioners could arise from strain to social relationships with those who don’t participate (Hernlund and Shell-Duncan, 2007), especially if abandoners exert social pressure on others to abandon as well (Mackie, 2017), including following intervention activity (Diop et al., 2008).

One can represent this scenario by modifying the definition of actors’ utility functions in the coordination model as follows:

⁴⁵ As noted in Bicchieri’s (2005) philosophical treatment of social norm theory (which is widely referred to in social norm theory treatments of FGM, e.g. Mackie et al., 2015), social norm theory allows that *some* actors in the coordinating population may unconditionally obey (or disregard) the norm. Introducing subjective ‘intrinsic benefits’ of FGM into decision-making in the standard model allows for this possibility

$$U(\text{practice}) = -H_i + (1 - p_i) \cdot -\hat{M}$$

$$U(\text{abandon}) = -\hat{M} \cdot p_i$$

Here, actors *also* face social costs for unilaterally *practicing* FGM. Social pressure to abandon is the relative-complement of social pressure to practice $(1 - p_i)$. This leads to a new definition of actors' thresholds in terms of \hat{M} and H_i :

$$-H_i + (1 - p) \cdot -\hat{M} = -\hat{M} \cdot p_i \Rightarrow p_i = \frac{\hat{M} + H_i}{2\hat{M}}$$

Substituting in the standard definition of H_i (i.e. in terms of q) we arrive at a new definition of actors' thresholds (in terms of q):

$$H_i = q \cdot \hat{M} \Rightarrow p_i^* = \frac{\hat{M} + q \cdot \hat{M}}{2\hat{M}} = \frac{q + 1}{2}$$

In the standard coordination model, actor's thresholds are equal to q . Once symmetric social pressures are introduced, they are equal to $\frac{q+1}{2}$. The effect is that now, rather than being between 0 and 1, all actors' thresholds are shifted upward and compressed to between 0.5 and 1. As summarised in Table C6.1, the effects on the dynamics of the model are dramatic.

Model Variant 3: Heterogeneous Autonomy (Relations Between Social Pressure and Social Costs)

In the standard coordination model (see also Platteau et al., 2017), I assumed that the relation between *social pressure* to abandon FGM (p_i) and the costs of abandoning FGM (\hat{M}) is linear ($p_i \cdot \hat{M}$). Analyses which obfuscate social costs through threshold analysis, do not make particular claims about the relation between FGM activity and social costs, except to say that as more people do FGM, the cost of abandoning it goes up (e.g. Novak, 2016). What's missing from both kinds of analysis is an explicit examination of what might happen if the relationship between social pressure and social costs of abandonment *varied between actors*.

Theoretical discussions of the social dynamics of FGM clearly suggest that this might be the case. Certain actors, because of their status in the community, their personal qualities, or their ability to avoid the social costs of unilaterally abandoning FGM, may be less sensitive to the social costs of abandoning FGM (Cloward, 2015: 387). These actors may be able to 'withstand' a certain amount of social pressure. Shell-Duncan et al. (2011: 10) make a similar observation, noting that female elders in the community are 'less likely' to 'rely on their connections for social or material support'. Conversely, others, such as young women (Shell-Duncan et al., 2011), may rely on others more heavily, and so be more vulnerable to social pressure.

These ideas can be subsumed under the concept of *heterogeneity of autonomy* within the community. They can be modelled by introducing a potential *non-linearity* in the relation between social pressure and social costs that *varies between social actors*. To introduce this modification into the standard coordination model one can replace p_i with a function $f(p_i)$ defined as:

$$f(p_i) = p_i^{2^{\alpha_i}}$$

$$\alpha_i \in [-1, 1]$$

Where α_i is the *autonomy* of agent i , and varies between actors. High values of α_i imply that an actor is highly autonomous. Low values imply that an actor has low autonomy and is vulnerable to social influence⁴⁶. The representation of p_i in the standard coordination model can be seen as a special case of $f(p_i)$ in which α_i is equal to 0 for all actors (i.e. autonomy is homogenous).

As the following equations, and Figure 20, illustrate, the ‘limits’ of social pressure remain the same ($f(p_i) \in [0, 1]$). However, when autonomy is high, the social costs of abandoning FGM ‘grow’ less quickly as social pressure increases. When autonomy is low, they grow more quickly. When autonomy is at its lowest value (-1) the social costs that actors experience for abandoning FGM at 25% social pressure, are equivalent to the social costs experienced by actors in the standard coordination model at 50% social pressure.

$$\alpha_i = 1 \Rightarrow p_i^{2^{\alpha_i}} = p_i^2$$

$$\alpha_i = 0 \Rightarrow p_i^{2^{\alpha_i}} = p_i$$

$$\alpha_i = -1 \Rightarrow p_i^{2^{\alpha_i}} = p_i^{\frac{1}{2}}$$

⁴⁶ Note that the assumption that non-linearity is second-order (χ^2) is arbitrary, and relaxed in Chapter 6.

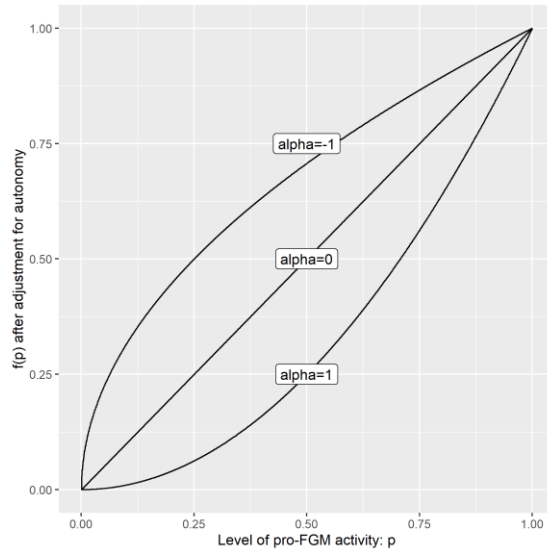


Figure 20: Effects of Autonomy (α) on the Relation between Social Pressures and Social costs

When exploring the robustness of coordination dynamics with heterogeneous autonomy I represented autonomy as uniformly distributed, and having a perfect rank-order correlation with support for FGM (i.e. lower H_i , higher autonomy). This is arbitrary (and relaxed in Chapter 6), but consistent with the idea that supporters of FGM within communities are often also those with the most power, and the least reliance on the support of others (Shell-Duncan et al., 2011).

Incorporating Heterogeneity of Social Influence

The assumption of existing coordination models of FGM (including heterogeneous threshold models) is that *social pressure* depends on the *proportion* of others in the reference-group practicing FGM. This implicitly assumes that social influence is *homogeneous*, all actors (through their choices) exert *equal* social pressure on others. Yet other literature on FGM strongly suggests that this may be an idealisation. In this section I address three credible ways of representing *heterogeneity of social influence* within the standard coordination model.

Model Variant 4: Heterogeneous Authority

Theoretical discussions of the dynamics of FGM have emphasised the idea that there may be ‘elites’ (notables, leaders, elders, etc.) in local communities who have a disproportionate influence on the decisions of others (Cloward, 2015: 394; Mackie and LeJeune, 2009: 13, Mackie et al., 2015: 39). Empirical studies reach a similar conclusion. Shell-Duncan et al. (2011: 10) emphasised the influence of female elders on the decisions of others in the community to practice FGM (likewise Shell-Duncan and Hernlund, 2006 and Dellenborg, 2004). Hernlund and Shell-Duncan (2007: 54) note other figures, including village leaders and religious authorities as highly influential. Evaluation of the Tostan program in Senegal, found that some intervention participants felt it essential that important figures, such as the village leader, were included in the abandonment of FGM (UNICEF, 2008: 29). These accounts suggest that actors may vary in their *authority* with the community, with some authoritative actors exerting more social influence over FGM decision-making than others.

I incorporated *heterogeneity of authority* into the standard coordination model of FGM by assigning each actor an authority ‘weight’ ($w_i \in [0,1]$) and redefining the level of social pressure (p_i) in terms of the *weighted proportion* of actors in the reference-group practicing FGM:

$$p_i = \frac{\sum_{j \in \{\text{social-ref}\}} w_j \cdot d_j}{\sum_{j \in \{\text{social-ref}\}} w_j}$$

When exploring the robustness of coordination dynamics with heterogeneous *authority* I represented authority as uniformly distributed, and having a perfect rank-order correlation with support for FGM (i.e. lower H_i means higher authority). Again, this assumption is arbitrary (and relaxed in Chapter 6) but compatible with empirical research suggesting that powerful actors in FGM practicing communities are often supporters of FGM (Shell-Duncan et al., 2011; Dellenborg, 2004).

Model Variant 5: Norm Enforcement (Explicit Versus Implicit Social Pressure)

The standard assumption in coordination models of FGM has been to represent social pressure as *implicit*. Actors perceive there to be social pressure to practice FGM *if* others practice it. The practice of FGM by others may create a marriage market that non-practitioners are excluded from (Mackie, 1996), or it may involve important community activities from which non-practitioners are excluded (Shell-Duncan et al., 2011; Dellenborg, 2004). Also, in a cultural context where FGM is recognised as a rule of behaviour, practice by others may imply that they recognise this ‘rule’ as legitimate and would disapprove of others not followed the shared custom⁴⁷ (Bicchieri, 2005; Bicchieri and Mercier, 2014).

Yet, this representation abstracts-away the possible impact of *explicit* norm enforcement on social dynamics. Empirical research, including in Senegal, emphasises the potential *explicit* negative social sanctioning of those who don’t practice FGM, including name-calling and criticism (Hernlund and Shell-Duncan, 2007; Shell-Duncan et al., 2011; Shell-Duncan and Hernlund, 2006). The important implication of this is that actors may have different *roles* in the social dynamics of FGM. In the standard model, *all* actors implicitly influence others through their choices. Yet, it is reasonable to suppose that only some actors *explicitly* influence others by *enforcing* the norm. This suggests an alternative possible representation of the social pressure that actors face to practice FGM, involving *explicit* influence by a subset of *norm enforcers*.

I designed a version of the standard coordination model with *explicit norm enforcement* by selecting a subset of agents (10%) who attribute the highest value to FGM (lowest H_i values) and assigning them the role of ‘norm-enforcers’. These agents will ‘enforce’ FGM practice as long as they practice it themselves ($d_i = 1$). Other agents never try to enforce FGM practice. As such I redefined social pressure (p_i) with a different reference group, representing all norm-enforcers in the community ($\{\text{norm-enforcers}\}_i$):

$$p_i = \sum_{j \in \{\text{norm-enforcers}\}_i} \frac{d_j}{n(\{\text{norm-enforcers}\}_i)}$$

⁴⁷ Bicchieri (2005: 52) refers to this as a mapping between what you expect others to do, what you think you ‘ought’ to do (i.e. your ‘normative expectations’).

Here the level of pro-FGM social pressure depends on the level of norm enforcement in the norm-enforcement reference-group. The effects of this change on the dynamics of the standard social coordination model are summarised in Table C6.1.

Model Variant 6: ‘Decision-Maker’ Reference Groups

Representations of actors’ reference-groups as involving all others in the community could be modified in a different way. The relevance of marriage in the Senegalese context is disputed (see Hernlund and Shell-Duncan, 2007 and Shell-Duncan et al., 2011 *against*; Mackie, 2000 *for*). Yet it is certainly possible that in some contexts social actors are primarily concerned with the future state of the marriage market (from which uncut women will be excluded). In this case, actors may not be concerned with participation in FGM in the community in general, but only with the final decisions being made about the FGM status of girls at a given moment in time.

In this case, we would expect actors’ reference-groups to be the *final decision-makers* about the cutting status of girls in the community. As such, we would define pro-FGM social pressure (p_i) in terms of the decisions of this subset of the population:

$$p_i = \sum_{j \in \{\text{decision-makers}\}} \frac{\theta_j}{n(\{\text{decision-makers}\}_i)}$$

Where θ_j represents the latest final decision about the FGM status of a girl, made by actor j (who is a final decision-maker).

Distinguishing participation in FGM from final decisions about particular girls would also raise important questions about *temporal dynamics* in coordination models of FGM that have previously not been considered. Opportunities to participate in FGM practice occur frequently (many girls are cut every year). But, a decision-maker will only have to make a ‘final decision’ about one of the girls in their household occasionally. One possible implication is that it will take longer for changes in decisions about FGM to ‘diffuse’ through the community, since decision-makers will not have the opportunity to ‘update’ their choice to stop cutting until such a decision actually arises.

In the context of a social simulation, one can conceptualise this as follows. There are a set of ‘time-steps’, and in each step, actors decide to practice FGM or not. In each time-step, ‘final decision-makers’ in each household may or may not have to decide whether to cut one of the girls in their household. This will depend, among other things, on the age of the girl.

One can model this, in turn, by assigning a subset of agents in the community the role of ‘decision-maker’ and assigning them a probability that, in a given time-step, they will have to decide on the FGM status of one of the girls in their household (which is assumed to match their decision to practice FGM in general). To assign a meaningful value to this probability one has to consider a range of quantities. First, one has to assume that there are a certain number of time-steps per-year. In other words, an explicit temporal scale has to be added to the simulation. Then, one can estimate the probability that, in a given time-step, a decision-maker will need to decide on the status of a girl in their household. This depends on the number of households in the community, the typical number of girls per-household, and the age by which girls are expected to be cut (if they are to be cut at all).

I designed a version of the standard coordination model which incorporated coordination on the specific cutting decisions of a subset of ‘decision-makers’ and calculated the above probability using figures for Senegal from an international survey of household characteristics (United Nations Department of Economics and Social Affairs, 2017, see calculations in Appendix C2). I assumed arbitrarily (though see Chapter 6) that there were 12 time-steps per-year. Results of simulations using this ‘variant’ model are summarised in Table C6.1.

Introducing Social Structure

Model Variant 7: Social Reference Network

With the exception of Efferson et al. (2019) and early outputs of this project (see Appendix G), coordination models of FGM have always assumed global interaction. This means that actors’ social reference groups consist of all others in the community.

Yet, empirical accounts emphasise that social interactions related to FGM are structured around important social relationships within communities, including between intermarrying families (Mackie, 2017:16), within households (Hernlund and Shell-Duncan, 2007; Shell-Duncan et al., 2011; Shell-Duncan and Hernlund, 2006), within extended families (Camilotti, 2016; Hernlund and Shell-Duncan, 2007; Shell-Duncan and Hernlund, 2006) and across social networks (Shell-Duncan et al., 2011). Put simply; actors may be focused on coordinating with important social relations in the community, rather than the community as a whole. As such their *social reference group* may consist of this more limited set of relations.

A natural implementation of this feature in a simulation context is to include a *social network* which connects actors in the community. In so doing, we construe the connections between

actors in the network as representing important social relationships. An actors' 'reference-group' with respect to decisions about FGM is then the set of actors to whom they are connected in the network. In this way, we *constrain* social interaction to interactions between connected actors.

I incorporated a social network into the standard coordination model using a version of the spatial network algorithm developed by Hamill and Gilbert (2009). This algorithm randomly scatters agents across a finite 2D space (I refer to this as 'network space') and then connects agents who are within a given distance of one another (see Appendix C3 for further details). This tends to produce a more realistic network structure (e.g. terms of network clustering) than more traditional algorithms such as random networks, or preferential attachment networks (Hamill and Gilbert, 2009).

To explore the impact of social network structure on the dynamics of the standard coordination model, I set the *average* number of social connections per actor to 20 to part of simulated experiments. The effect on the standard coordination model is summarised in Table C6.1.

Model Variant 8: Households as a reference group

Above, I address the idea that actors have some important and some *unimportant* social relationships in the local community. This is captured by the idea of a social network spanning the local community. Generally, this network is a single 'clique' (a 'path' exists between all actors in the network).

However, one aspect of 'localised' social interaction noted above is that actors also interact *within households*. Shell-Duncan and Hernlund (2006), for example, found that interactions *within households* between co-wives could be an important source of social pressure on individuals to practice FGM. This, again, suggests a degree of 'localisation' of social interaction. However, it suggests a rather different kind of network structure. Specifically, households are non-overlapping *network cliques*, with all members of the household connected to one another, and individuals typically living in only one household.

I incorporated a social structure of this kind into the social convention model by dividing the population of agents into a set of fully-connected network cliques (see Appendix C4 for details). Naturally, if agents then *only* coordinate within households, then there can be no community-level social dynamics. However, I was interested in how the dynamics of the social coordination model might be disrupted if the social pressures facing actors were a

combination of influence from a household reference group ($\{\text{household}\}_i$) and a social reference group ($\{\text{social-ref}\}_i$).

This was implemented formally by redefining the social pressure (p_i) component of agents' utility functions as a combination of social pressures from both reference groups.

To do this, agents' utility functions were defined as before:

$$U(\text{practice})_i = -H_i$$

$$U(\text{abandon})_i = -\hat{M} \cdot p_i$$

But with p_i defined as:

$$f(p_i) = (1 - s_3) \cdot p_{\text{social}_i} + s_3 \cdot p_{\text{household}_i}$$

Where $p_{\text{household}_i}$ represents the proportion of the social reference group ($\{\text{social-ref}\}_i$) practicing FGM, and p_{social_i} represents the proportion of the household reference group practicing FGM ($\{\text{household}\}_i$).

Here the balance of influence from these two sources is controlled by the $s_3 \in [0,1]$ parameter. When conducting a robustness analysis with this model feature (household influence), s_8 was arbitrarily set to 0.8 (but this is relaxed in Chapter 6), so that household reference groups had a strong effect on the social pressures felt by agents.

Incorporating Additional Intervention Features:

The final elaborations of the standard coordination model that I considered were further elaborations of the representation of the intervention process itself. Both empirical and theoretical accounts of anti-FGM community-level interventions highlight additional aspects of these which may have important effects on social dynamics.

Model Variant 9: Heterogeneity of Responsiveness

The first elaboration I considered is a simple one also employed in Efferson et al. (2019); this is the notion of *heterogeneity of responsiveness* to the intervention. Efferson et al. (2019) point to evidence from an intervention experiment conducted in Sudan (Vogt et al., 2016) to suggest that actors who are initially more reluctant to abandon FGM may respond differently to intervention activities. They incorporated this into (some of) their analyses by assigning each actor a probability that they would respond to the intervention. In their strongest implementation of this feature, this probability was 0 if the agent had a threshold of 0 for practicing FGM, and was 1 if the agent had a threshold of 1 for practicing FGM (and linearly increasing between the two).

I incorporated a similar assumption into my representation of the intervention process (i.e. including coalition formation), with the probability of a targeted actor joining the ‘initial coalition’ of abandoners set as:

$$Pr(join\ intervention)_i = \frac{H_i}{\bar{M}}$$

Under this version of the model, all targeted actors are exposed to the educational effect of the intervention⁴⁸, but they are not guaranteed to be willing to join the initial coalition of those willing to abandon the practice (they will do so with probability $\frac{H_i}{\bar{M}}$). The impact of incorporating this assumption into the model is summarised in Table C6.1.

Model Variant 10: Organised Diffusion

A further credible elaboration of the simulated intervention that can be combined with the notion of heterogeneity of responsiveness is the idea of ‘organised diffusion’ (Mackie and Lejune, 2009; UNICEF, 2007). Mackie and Lejune (2009) argued that the initial group of

⁴⁸ Note that to assume otherwise would be to assume that interventions can never directly affect actors who don’t initially view FGM as costly. Also, it would ignore the fact that anti-FGM activities often taken place as part of larger community-education projects which are not exclusively focused on FGM (Cheikh Seydil Mactar Mbacke, 2018)

participants in an intervention will be incentivised to *recruit others* to join their coalition. This is because the greater the level of abandonment, the greater the reduction in social costs for those abandoning. They suggested that community interventions could create a process of ‘organised diffusion’, where the initial participants recruit others to join the abandonment coalition and pass on their knowledge about the harms of FGM.

“a relatively small core group of first movers, called the critical mass, can conditionally resolve to abandon FGM/C, and then has an incentive to recruit remaining members of the community to conditionally join in the effort... After second movers conditionally commit to abandon, third movers are recruited, and so forth” (Mackie and LeJeune, 2009: 11)

Qualitative evaluation of the Tostan interventions in Senegal (Diop et al., 2008) indicated that a pattern of this kind could occur, with initial participants mobilised to persuade others in the community to join them in their commitment to abandoning FGM.

I incorporated ‘organised diffusion’ into the standard coordination model as a recursive process. An initial group of actors is targeted, some of whom form an initial coalition who are conditionally willing to abandon FGM. Then, the actors in the coalition spread the educational effect of the intervention to others and try to recruit them to join the coalition. New recruits also spread the intervention to others, and so on, until no-one else agrees to join the initial coalition. The key parameter in this process is the number of ‘others’ that each actor tries to recruit (z_4). Further details of the implementation of this process are given in Appendix C5.

The impact of incorporating organised diffusion into the intervention process is summarised in Table C6.1. For those simulations, I assumed that actors who join the initial coalition attempt to recruit 2 others ($z_4 = 2$).

Model Variant 11: Building Social Relationships Between Intervention Participants

The final elaboration of the simulated intervention that I considered was the formation of new social relationships between social actors. Intervention programs such as Tostan have been predicated on the idea of empowering communities through education, and through dialogue among participants. Also, evaluations of the program report positive changes in social relations within practicing villages (Diop et al., 2008). Yet, formal representations of the community intervention process have omitted any possible effect of the intervention on the social structure of the community itself.

Changes in social structure may be a moot point in the standard coordination model. That model assumes that social interaction is global. However, if interactions occur within a social network, then a plausible effect of the intervention on social relationships is that it *connects* previously unconnected coalition members⁴⁹. The effect of this, within the model, would be to create *new* social pressures (based on other coalition members' choices) but also to relieve existing social pressures, by changing the composition of actors' social reference groups in a way which favours abandonment.

To explore the robustness of the standard coordination model to this elaboration of the simulated intervention, I included a social network with average connectivity of 20 (using the spatial network algorithm described above) and added a component to the simulated intervention wherein network connections were created between all members of the initial coalition.

⁴⁹ This is, admittedly, only a possible impact of community-level interventions on social structures, but it is certainly at-least as credible as the idea that such interventions have no effect on social structure at all.

Results of the Robustness Analysis

Simulated experiments measuring dynamics of persistence and dynamics of intervention (see section: *key dynamics of the standard coordination model*, above) for:

- The standard coordination model (results given in the chapter)
- The eleven variant models outlined above
- The four non-‘normal’ thresholds distribution summarized in Table 3, but including coalition formation in the intervention process

Full reports, including graphics, for each model variant are included in the DVD-ROM attached to the thesis (each would be multiple A4 pages to print, making inclusion in an appendix impractical).

Detailed summaries of the results are given in Table C6.1 (Appendix C). For each model variant, Table C6.1 includes a summary of notable changes (if any) in social dynamics created by the modification of the model. I divide the discussion into dynamics of persistence, and dynamics of intervention. I also offer an interpretation of the underlying ‘processes’ which are generating these changes in dynamics. To help the reader to conceptualise the magnitude of changes, I include quantitative summaries of the simulations’ output, where appropriate. For the outputs of the ‘Dynamics of Persistence’ experiments, these quantitative summaries include the average rate of FGM in the population and the proportion of ‘interior’ rates of FGM (conservatively defined as rates of FGM between 10% and 90%). For outputs of the ‘Dynamics of Intervention’ experiments, quantitative summaries include the average reduction in the rate of FGM over all simulations, the average spillover from simulated interventions⁵⁰, and the proportion of simulated experiments resulting in coalition collapse (i.e. no stable change in FGM practice).

In assessing the dynamics of interventions in each model variant, I considered both overall changes in the predicted effectiveness of the intervention *and* changes (if any) in the relative advantages and disadvantages of different targeting biases. As previously, I refer to agents who perceive FGM to have high intrinsic cost (high H_i) and so are more willing to abandon

⁵⁰ Defined as the difference between the size of the intervention and the proportion of agents abandoning FGM, a positive value means that the intervention has a positive spillover.

FGM, as ‘willing agents’. I refer to agents who perceive FGM to have low intrinsic cost (low H_i) as ‘reluctant agents’.

The ‘headline’ result is that virtually all of the modifications to the standard coordination model considered in this chapter substantially altered the dynamics of intervention or persistence (or both) of the model. The only exception was heterogeneity of autonomy, which appeared not to have a meaningful effect (but see section below).

As expected, dynamics of persistence were strongly affected by the distribution of preferences, with the U-shaped distribution, in particular leading to interior rates of FGM persistence. However, interior rates of persistence were also created by change assumptions about social structure. Interior rates of FGM occurred after the introduction of a social reference network, or when social influence was shared between social reference groups and household reference groups. Apart from revealing the potential importance of these design elements, this finding underscores an argument made in Chapter 2. This argument is that *a heterogeneous distribution of preference favouring interior rates of FGM* in a coordination model is only one of a number of ways that interior rates of FGM might be predicted by such a model. This is a general methodological issue. Model design uncertainty presents a considerable challenge when trying to *diagnose* particular problems with a model’s design using macro-level data. Put simply; interior rates of FGM could be observed because there is strong variation in preferences, a particular kind of social structure, or for some other reason.

Other modifications to the standard coordination model strongly ‘biased’ the rate of FGM in the model towards universal practice or abandonment. As expected, the L-shaped distribution of preferences tended to drive the population to universal practice (and vice-versa for the J-shaped distribution). However, heterogeneity of authority, as-well-as explicit norm enforcement, also tended to drive the population to universal practice. This occurred because social influence was concentrated in actors who were the least willing to abandon FGM (lowest H_i value). Introducing FGM as an intrinsic value or allowing symmetry of social costs, had countervailing effects, with the former driving the population to universal practice, and the latter driving the population to universal abandonment.

Outcomes of simulated interventions were highly sensitive to a wide range of elaborations of de-idealisation in model design. Under random targeting, the success of the intervention was undermined by heterogeneity of authority and norm-enforcement. Once again, this was because of concentration of social influence among those least willing to abandon. The

localisation of social influence (e.g through a social reference network) had important *benefits* for intervention outcomes in that it prevented coalition collapse.

Different model elaborations interacted with the effects of targeting bias. A key (previously unrecognised) advantage of targeting agents willing to abandon FGM was that this prevented coalition collapse. However, the advantage could also be reversed under certain model variants. When, for example, social influence was concentrated among those reluctant to abandon FGM (norm enforcement / heterogeneity of authority), targeting those willing to abandon FGM was virtually useless, since these actors had so little influence. The key advantage of targeting willing agents was could also be attenuated by modifications which made coalition collapse unlikely across all targeting strategies. This was especially the case when the intervention involved organised diffusion or created new relationships between coalition members. In these cases, targeting those reluctant to abandon became superior again – because the advantages of targeting these actors for positive spillovers *after* coalition formation process were retained.

Finally, some contradictions were found with Efferson et al.'s (2019) results. These authors concluded that targeting bias would only affect intervention outcomes if the distribution of thresholds were roughly symmetric and unimodal (e.g, normal). However, their analysis focused exclusively on positive spillovers and did not include a coalition formation process. Incorporating a coalition formation process into the simulated intervention created strong sensitivities of intervention outcomes to targeting bias under non-symmetric (i.e. L-shaped and J-shaped) distributions, particularly through the possibility of coalition collapse and reduced coalition size.

Model Complexity and Interactions Between Features

The above results show that the ‘standard’ social coordination model is non-robust to a range of credible modifications of its design. There were 14 model variants considered, including analyses with the four additional distributions of preferences (see Tables 3 and 5). These were considered ‘one-at-a-time’, i.e. taking a baseline ‘standard’ model and adding a modification. This is *sufficient* to show that the modification in question ‘matters’ since it can *potentially* disrupt the key dynamics of the model. However, it’s possible that a feature is important despite not changing dynamics when considered in isolation. This could occur because one modification *interacts with* others. My choice to consider features one-at-a-time was a pragmatic one. If I were to consider all *pairs of* modifications, for instance, I would need 91 variant models. Recognition that model modifications can interact together in their effect on simulated dynamics, potentially in unintuitive ways, underscore the formidable challenge created by model design uncertainty. Here I present two examples of such interactions. One example shows that although introducing the *heterogeneity of autonomy* assumption had no substantial effect on the standard coordination model, it can have a substantial effect when *combined with* a uniform distribution of preferences.

Norm Enforcement Within Networks

In the above analysis, norm enforcement was considered in the context of global interaction (all actors interact with all others). The effects on dynamics of persistence are shown in Figure 21. This shows that the persistence of FGM was often driven to universal practice, and no stable interior rates of FGM occurred.

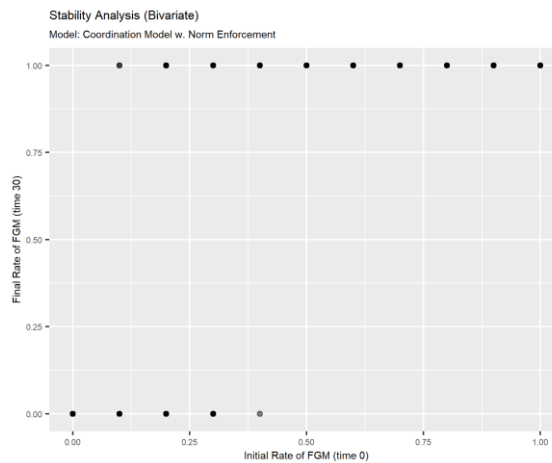


Figure 21: Dynamics of Persistence with Explicit Norm Enforcement by a Subset of Actors (values overlap, and darker dots indicate more data-points)

Next, Figure 22 shows the dynamics of persistence of the standard coordination model without norm enforcement, but when a social network is used (with an average connectivity of 20).

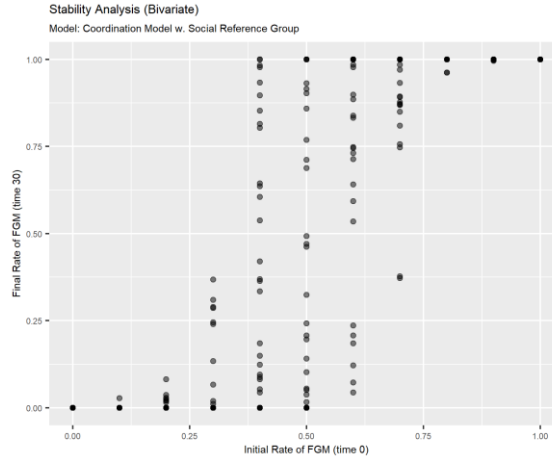


Figure 22: Dynamics of Persistence with a Social Reference Network (average connectivity of 20, values overlap, and darker dots indicate more data-points)

Here we see that a range of interior rates of FGM occur, although they tend to be limited to the case in which the initial rate in the simulation is close to 0.5. Interior rates of FGM occur because of clustering in the network, with some clusters practicing FGM and some abandoning.

Now consider the dynamics of stability that occur when norm enforcement and a social network are combined together (using the same parameters for both features, Figure 23).

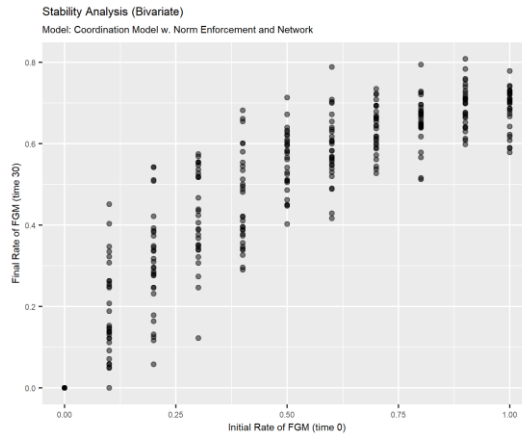


Figure 23: Dynamics of Persistence with Explicit Norm Enforcement by a Subset of Actors and a Social Reference network (average connectivity of 20, values overlap, and darker dots indicate more data-points)

Here we see that interior rates of FGM in the simulation are strongly *enhanced*. This effect isn't attributable to either element of the model in isolation but occurs because they interact together. The reason for this enhanced stability of interior rates is as follows. When a social

network is present, and norm enforcement is concentrated in a small number of actors (here 10%), the network can actually *shelter* some actors from norm enforcement. For some actors; none of their important social contacts are norm enforcers. This enhances the capacity of subsets of the network to resist pressures to abandon or practice FGM. Figure 24 (below) illustrates this effect, showing a random output of the simulation at initialisation. Here, actors are coloured blue if none of their immediate social relations are norm enforcers.

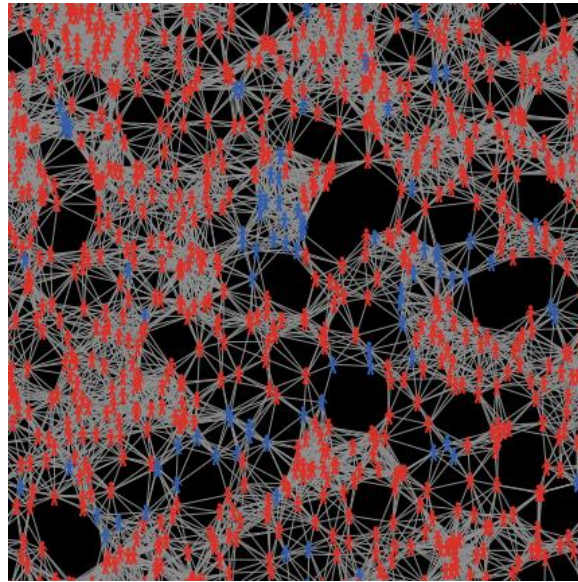


Figure 24: Simulation Output: Social Reference Network with Norm-Enforcement (actors coloured blue if none of their network contacts are norm enforcers)

Heterogeneous Autonomy with Uniform Approval of FGM

My second example involves the combined effects of heterogeneity of autonomy (which is individual variation in the relationship between social pressures and social costs) and a *strict* uniform distribution of preferences in the population.

Figure 25 shows the dynamics of stability of the standard coordination model with a strict uniform distribution of preferences. As expected, one can see that a wide range of rates of FGM can occur, most them interior.

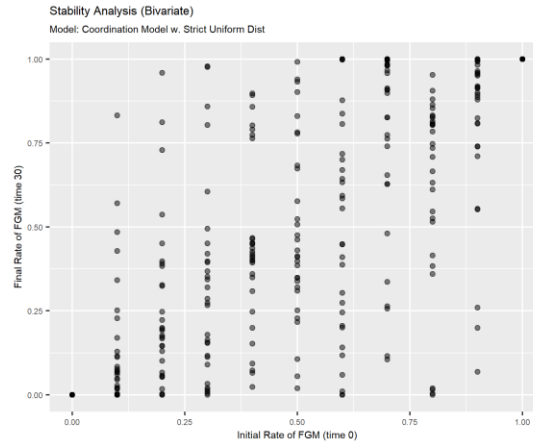


Figure 25: Dynamics of Stability with a Uniform Distribution of Preferences (values overlap, and darker dots indicate more data-points)

The results of the robustness analyses showed that introducing heterogeneity of autonomy had no meaningful effect on persistence dynamics in the standard coordination model (with ‘normal’ distribution of preferences).

However, when *combined with* a uniform distribution of preferences, the effects are stark, as shown in Figure 26. Introduction of heterogeneity of autonomy effectively *restores* the dynamics of the standard coordination model with a norm preference distribution.

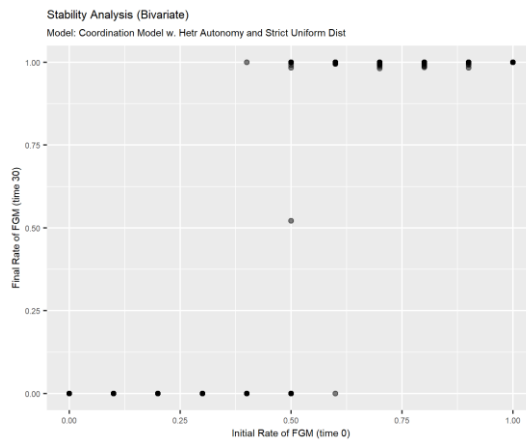


Figure 26: Dynamics of Stability with a Uniform Distribution of Preferences and Heterogeneity of Autonomy

To understand why this occurs, we need to look at the potential effect of heterogeneity of autonomy of the distribution of thresholds. When heterogeneity of autonomy is introduced, the threshold of individual agents becomes:

$$p_i^* = \left(\frac{H_i}{\widehat{M}} \right)^{\frac{1}{2\alpha_i}}$$

Which is a function of *both* the actor's approval of FGM (H_i) and their autonomy (α_i). In the analysis, I assume that α_i is uniformly distributed with a perfect (negative) rank correlation with (H_i), such that those reluctant to abandon FGM are the most autonomous. The effects of this on the distribution of thresholds are illustrated in Figures 27 and 28. We can see that heterogeneity of autonomy can *compress* the distribution of thresholds, making it unimodal. The effect (Figure 28) is to create a cumulative distribution of thresholds whose expected dynamics match those of the standard coordination model with 'normal' distribution of preferences. The cumulative distribution crosses the diagonal from below, meaning no stable interior rates of FGM.

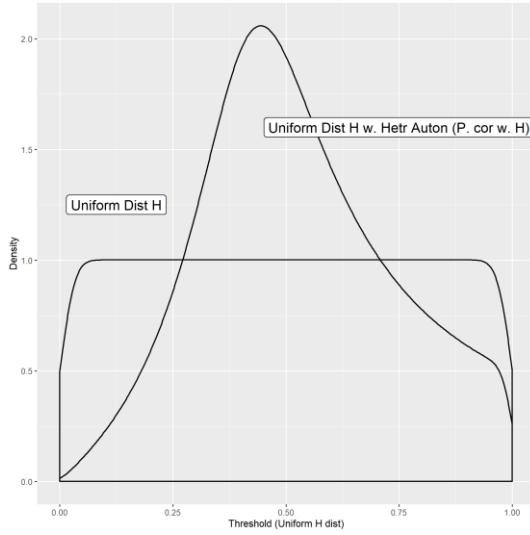


Figure 27: Effect of Heterogeneous Autonomy on Threshold Distributions (PDF)

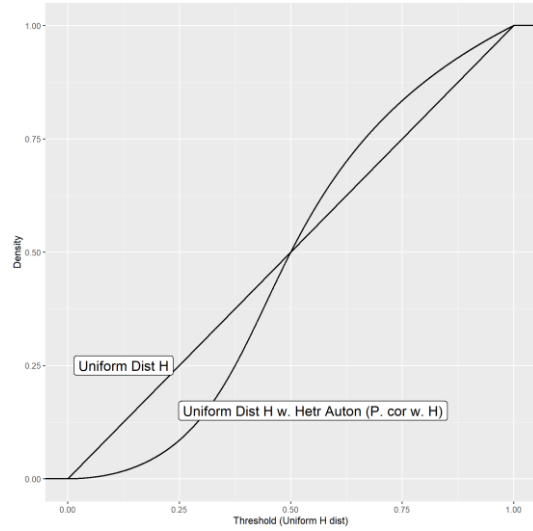


Figure 28: Effect of Heterogeneous Autonomy on Threshold Distributions (CDF)

Conclusions

The key dynamics of the standard coordination model are highly non-robust to a wide range of credible modifications of its design, drawn from empirical research and theoretical literature and existing models. Demonstration of the potential impact of these design features reveals them to be important areas of *uncertainty* in the design of coordination models of FGM. This sets the agenda for subsequent model development in the thesis. It also provides important information to the field as a whole. Insights drawn from existing coordination models may be fragile. They may depend on the choice to omit or simplify the features considered here. In order to manage uncertainty when building coordination models of FGM, researchers should *at least* consider the set of design ‘possibilities’ outlined in this chapter. Future research may also identify *other* areas of design uncertainty not considered here. This is meant to be a cumulative exercise.

In the next chapter, I implement a strategy for managing these areas of uncertainty in the design of coordination models of FGM. This strategy involves building a ‘general’ model which incorporates and *parameterises* all of the elements considered here. The variant designs considered here are all locations in the parameter space of that model. The model is also able to, to some extent, generalise over the different possible versions of the design features considered here. In the general model, for example, the connectivity of the social network is not arbitrarily set at 20 but can vary according to a flexible parameter controlling connectivity. Among other things, this activity allows me to conduct global sensitivity analysis to establish the *relative importance* of these different areas of uncertainty, taking into account *interactions between* different possible elements of the model design.

Many of the modifications employed in this chapter, including social networks, coalition formation and organised diffusion, would be difficult or impossible to explore using closed-form analytical methods. Moreover, combining them *together*, as I do in the next chapter, would be extremely difficult⁵¹. By contrast, they are straightforward to implement using agent-based simulation. This underscores the necessary connection between the methodological strategy employed, and the use of simulation as a technical framework.

⁵¹ Closed form analytic solutions might exist, but would almost certainly be uninterpretable.

Chapter 6: A General Model of FGM As a Social Norm of Coordination

Introduction

In the previous chapter, I analysed a range of plausible de-idealisations and elaboration of a ‘standard’ coordination model of FGM using a social simulation (i.e. agent-based modelling) framework. These were drawn from relevant theoretical and empirical literature on FGM, as well as existing model designs. The explorations spanned aspects of agent’s decision-functions, different forms of social influence, different kinds of social structure (e.g. social networks) and different representations of community-level interventions designed to end FGM.

The different model design elements considered, represented areas of *uncertainty* in the design of coordination models of FGM. It may or may not be the case, for example, that there are typically social costs for unilaterally *practicing* (as opposed to abandoning) FGM. If there are, one doesn’t know that they will be of the same *magnitude* as costs for abandoning the practice. It seems reasonable to assume, for example, that interactions with respect to FGM in large communities occur within the constraints of a social network. Yet, the structure of this network is uncertain, it may be very highly interconnected (and therefore approximate global interaction) or more sparsely connected (and therefore impose strong localisation of interaction). The analysis presented in Chapter 5 demonstrates only that potentially important areas of uncertainty exist, not what their correct formalization is. In the process of building a single model of the social dynamics of FGM, a strategy for managing these different possibilities is required.

There are broadly two possible strategies for incorporating a range of important yet uncertain design features into a single model of social dynamics. The first is to address this uncertainty prior to model construction by attempting to independently calibrate each of the design features before adding it to the model. One might, for instance, decide to try to assess the network structure of the real population before adding a network structure into the model, and so on. The second is to attempt to build a model which *generalises over* the uncertainty associated with each of the design issues, with the calibration of this general model a separate activity. As the reader will know from foreshadowing in previous chapters, I prefer the latter strategy. In this chapter, I develop and analyse such a general model, using this analysis to inform the calibration of this model in the subsequent chapter (which also includes testing,

i.e. empirical validation). I outline my reasons for following this strategy (general model first, calibration second) below. However, before doing so, it is useful to provide a more detailed discussion of what I mean by ‘building a *general* model’, in this context.

‘Generality’ in Modelling

The concept of the generality of models is often misunderstood, or at least, it has a number of distinct possible interpretations (Levins, 1993). A common, and often criticized (Edmonds and Moss, 2005), notion is that simple models are general models. Here generality seems to refer to *generality with respect to the predictions of the model* or *generality with respect to the construal of the model*. The former interpretation seems to depend on the idea that if a simple model relies only on a minimal set of features (e.g. neighbourhood similarity preference in the Schelling Segregation model - Chattoe-Brown, 2013) to produce some dynamic, then this can act as a general model of (the large set of) real-world targets which share this minimal feature set. Yet clearly such a perspective depends on at least two suppressed (and highly suspect) premises: that the dynamics of the simple model depend *only* on features genuinely shared with the real world targets (rather than idealisations not found in those targets) and that all systems with the shared features in question will produce the dynamics in question, irrespective of other real-world features.

It is difficult to see why either suppressed premise should hold universally (or even frequently). As such, widespread reference to simple models as *general* likely refers implicitly to generality of *construal* (Weisberg, 2012). Since simple models rely on a fairly minimal set of features (which may be shared with a wide range of targets), it may be possible to *construe* them as models of a wide range of phenomena (Levins, 1993). Yet, while this kind of generality may be conceptually or pedagogically valuable, it is not clear that it conveys any strong epistemological justification of the model itself (the relevant shared features may or may not furnish an adequate model in each case).

The concept of a general model which I use here doesn’t refer to creating a set of predictions which can generalise over a range of targets or producing a model which can be construed as a model of a range of targets (except perhaps incidentally). Instead, it refers to a model which generalises (in a technical sense) over a space of possible but uncertain model design choices – with a consequent *expected* loss of precision in its predictions (Levins, 1993). In this case, it also refers to a coordination model for which the *particular* coordination models considered in the thesis so far (i.e. the model variants developed in Chapter 5), can be considered *special cases* (i.e. particular parameterisations). A model of this kind does not (and

is not intended to) eliminate this uncertainty, instead, it provides a vehicle to systematically explore the *relative contributions* of different kinds of uncertainty to the dynamics of the model. An immediate illustration of this form of generalisation is as follows. Consider that in the previous chapter we explored the dynamics of a number of particular distributions of the intrinsic value that actors attach to FGM, including ‘normal, U-shaped and uniform. One way to further generalise over different possible threshold distributions, prior to attempting calibration, is to parameterise a continuous space of plausible distributions and to explore the full range of dynamics predicted across this space (see relevant section below).

This example (generalising over preference distributions) will make the predictions of the model less, rather than more, precise. However, this just means that we are incorporating into the model (rather than ignoring) the extent of our (current) uncertainty about dynamics, given our uncertainty about the distribution of actor’s characteristics. The general model developed in this chapter simultaneously incorporates (and parameterises) uncertainty about the full set of design options considered in the previous chapter. Given the potential for complex interactions between different design features, this general model *represents* a global space of uncertainty about model design and permits an *assessment of* the full range of dynamics which are predicted within this design space.

Advantages of Building a General Model Prior to Calibration

A core practical reason that producing a general model of this kind prior to calibration is useful is that independent empirical calibration is a heavily resource-intensive activity. Whether it involves: distilling empirical literature, analysing qualitative (Edmonds, 2015; Ghorbani et al., 2015) or quantitative data (Hedström, 2005), or consulting experts, empirical calibration takes time, and often financial resources. The more extensive the model and the weaker the availability of existing data, the more remote the possibility of adequately calibrating *all* model design features prior to model construction. In this context, efficiency is paramount. In a given research project, it is desirable to prioritise resources to focus on calibrating those areas of uncertainty in a model which most strongly drive the dynamics of interest (Saltelli et al., 2008). At the extremes, if some area of uncertainty is irrelevant to the important dynamics of the model, it would be a waste of resources to try to reduce this uncertainty (some particular specification can be chosen arbitrarily, Saltelli et al. 2008). Conversely, if the dynamics of a model depend entirely on a single uncertain feature, all of the available resources of the project should be dedicated to calibrating this feature.

Global sensitivity analysis exists precisely to perform the function of decomposing the uncertainty about the dynamics of a model into uncertainty about its constituent features (as Saltelli, 2008, describes it, the model's 'inputs' and 'outputs'). A general model which parameterises the different areas of uncertainty in model design facilitates global sensitivity analysis of this kind so that uncertainty can be apportioned to the different parts of model design, and calibration prioritised accordingly. This is exactly the strategy undertaken in this chapter - with the results of the global sensitivity analysis used to guide empirical calibration in the next chapter. The contrast with a 'calibrate-first' strategy is clear. Since the importance of features may depend on interactions with other features (e.g. one feature may make another irrelevant) there is no way to estimate the importance of individual features in the final model *without* a general model, and thus no rigorous way to prioritise different features for calibration.

Goals of the Analysis

The conceptual goals of this chapter are to generalise over the areas of model-design uncertainty (i.e. fragility) shown to be *potentially* important in Chapter 5 and to quantify the relative importance of each of these areas of uncertainty within a general model that combines them. The associated technical goals are to build a model which achieves this generality and to perform a global sensitivity analysis of this model. The first part of chapter focuses on summarizing the strategies used to generalize over the design features explored in Chapter 5.

The remainder of the chapter deals with the global sensitivity analysis of this general model, which reveals a number of parameters which disproportionately affect dynamics of intervention in the model— suggesting that these should be prioritized in calibration efforts.

I try to balance the necessary discussion of technical aspects of the model, with the broader conceptual issues of 'generalising' over design possibilities. Some of the denser technical material is confined to appendices. In keeping with best-practice in social simulation, a full ODD specification (Railsback and Grimm, 2012) for the general model is provided in the DVD-ROM attached to the thesis, along with the source-code itself (this ODD document is far too long to include in a printed appendix).

Generalising Over Model Design Possibilities

Generalising over the Decision-Functions of Agents

In the previous chapter, I began with a simple definition of the utility functions of agents:

$$U(abandon)_i = p_i \cdot -\hat{M}$$

$$U(practice)_i = -H_i$$

Where $p_i \in [0,1]$ is the level of pro-FGM activity (defined as the proportion of others practicing FGM, in the standard model). $H_i \in [0, \hat{M}]$ is the perceived intrinsic cost of FGM. \hat{M} , which is an arbitrary constant, is the maximum social cost for abandoning FGM. H_i was then defined as:

$$H_i = q \cdot \hat{M}$$

Where q is a beta-distributed random variable in the interval $[0,1]$.

I then considered three modifications to the structure of these functions. The first was that some actors may intrinsically value FGM (so $-H_i$ becomes positive). The second was that actors may also experience social costs for unilaterally practicing FGM. The third was that actors may vary in their autonomy, in a way which affects the relationship between social pressure (p_i) and $-\hat{M}$.

For the general model, I used a decision-function that generalises over these features and adds even further flexibility:

$$U(abandon)_i = p_{pro_i}^{V(-\delta_1 + \alpha_i)} \cdot -\hat{M}$$

$$U(practice)_i = H_i - (s_1 \cdot p_{anti_i}^{V(-\delta_2 + \alpha_i)} \cdot \hat{M})$$

Where H_i is defined as:

$$H_i = -\hat{M} + (q \cdot [\hat{M} + (\hat{M} \cdot s_2)])$$

Note here that H_i is redefined as the intrinsic **value** of FGM to the agent (rather than the intrinsic cost), which can be positive (FGM is viewed as intrinsically beneficial) or negative (FGM is viewed as intrinsically costly). Since H_i can be positive or negative in the general

model, I adopt the term ‘FGM supporters’ to refer to actors who value FGM ($H_i \geq 0$) and the term ‘FGM opposers’ to refer to actors who view FGM as costly ($H_i < 0$)⁵².

These functions generalise over all the possibilities discussed in the previous chapter, as follows:

- H_i can vary between $-\hat{M}$ and $\hat{M} \cdot s_2$. As such, the $s_2 \in [0,1]$ parameter controls the maximum perceived positive value of FGM in the population.
- The maximum social cost actors pay to unilaterally practice FGM varies from 0 to \hat{M} and is controlled by the s_1 parameter.
- The relationship between social pressure (p_{anti_i}/p_{pro_i}) and social costs (\hat{M}) is controlled by parameters $V (\geq 1)$, δ_1 (or δ_2 , both in the interval $[-1,1]$) and $\alpha_i (\in [-1,1])$, representing the autonomy of individual actor i . V controls the overall non-linearity of the relation between social pressure and social costs. δ_1 and δ_2 allow for global control of the relation between social pressures and social costs for abandoning or practicing FGM (respectively). As δ_1 increases, for example, the costs to abandon FGM ‘scale’ faster with social pressure, making it harder for actors to abandon the practice (and vice-versa for δ_2). δ_1 and δ_2 allow the model to accommodate the possibility that the relationship between social pressure and social costs varies *overall* for those practicing versus abandoning FGM. The effect of these parameters is eliminated if they are set to 0. The α_i component allows variation in the social-pressure social-cost relationship at the individual level, with costs scaling faster for less autonomous agents (lower α_i).

The utility-functions of agents in the standard model are a special case of this more general formulation, in which $V = 1$, $s_1 = 0$, and $s_2 = 0$. Other previously seen variations on the decision process can be achieved by appropriate manipulations of these parameters.

In the above formulation, I don’t define p_{anti_i} and p_{pro_i} , beyond that they represent pro-FGM and anti-FGM activity by others. However, I assume that they are both bounded between 0 and 1. We can turn now to their (general) definition in the model.

⁵² This replaces the use of the term ‘willing agents’ and ‘reluctant agents’ in the previous chapter.

Generalising over Sources of Social Influence

In the standard coordination model pro-FGM activity (p_i) is simply defined as the proportion of the social reference group ($\{\text{social-ref}\}_i$) who practice FGM. Also, the social reference group is defined as all other actors in the population.

However, I considered a range of modifications of these assumptions in Chapter 5. I considered that actors may vary in their influence (i.e. heterogeneity of authority) and I implemented this through weighted influence. I also considered alternative kinds of reference-group, including defining the social reference network ($\{\text{social-ref}\}_i$) in terms of network connections, an alternative reference group based on households ($\{\text{household}\}_i$) and an alternative reference group based on decision-makers in the community ($\{\text{decision-makers}\}_i$). I also considered different kinds of social pressure. I distinguished between *explicit* social pressure, and *implicit* social pressure, with the former supplied by agents with an assigned ‘role’ as a norm-enforcer.

The definitions of pro-FGM activity (p_{anti_i}) and anti-FGM activity (p_{pro_i}) in the general model *generalise over* these possibilities. They allow social pressure to be a flexible *combination* of influence from the three reference group types (social, household and decision-maker), and from the two kinds of social pressure (explicit and implicit). Social pressure can also be limited to only one of these sources. The technical details are rather involved, so further details are given in Appendix D1. However, the key points, and parameters, are as follows.

Social pressure is a weighted sum of implicit and explicit influence from three reference groups, consisting of decision-makers, the household, and a social reference group. p_{pro_i} and p_{anti_i} bound between 0 and 1. The parameters readers need to be aware of are:

- s_4 which controls the relative influence of the decision-maker versus the household/social reference group
- s_3 which controls the relative influence of the household versus social reference group
- s_5 which controls the relative influence of explicit versus implicit normative pressure.

The different assumptions about the sources of social influence considered in Chapter 5 can be easily recovered using these parameters. For the standard coordination model, s_3 , s_4 and s_5 should be set at 0. For the model with 80% household influence s_3 should be set at 0.8,

while s_4 and s_5 are set at zero, and so on. Beyond merely parameterising these different possibilities, the formulation developed here allows for a wide range of combinations of these different sources of social influence.

In the case of social pressure from norm-enforcement, the pro-FGM social pressure that actors experience is *not* necessarily the relative-complement of the anti-FGM social pressure that they experience. As such, the notion that actors have a single threshold, expressible as a level of pro-FGM activity (i.e. p_{pro_i}), can break down (see Appendix D2 for further details).

The definition of social pressure from household and social reference groups has *weighted* social influence (i.e. heterogeneity of authority) ‘built-in’. But one can recover the *homogeneity of authority* assumption in the standard coordination model by making authority (w_i) constant across the community⁵³.

Generalising over Reference Group Structures

In the previous section, I outlined how I generalised over different sources of social influence (i.e. from different kinds of reference group). In this section, I outline how I generalised over the *definition* of those reference groups themselves.

In the standard coordination model, there is only one reference group: the social reference group ($\{\text{social-ref}\}_i$). This consists of all other actors in the community; I referred to this as a ‘global’ reference group. In the general model, I defined actors’ social reference group as their connections within a social network. I used the spatial network algorithm outlined in Chapter 5, because of its plausible structural properties. However, the key point is that I parameterised the average connectivity in this ‘social reference network’ ($\mu_{Rsocial}$). As such, the global reference network in the standard coordination model can be recovered simply by setting $\mu_{Rsocial}$ to some arbitrarily large number (i.e. greater than the size of the population).

As previously, I defined household reference groups as network cliques ($\{\text{household}\}_i$). However, because the model includes a social reference network and household reference cliques, the overlap between the two becomes important. If one were to just randomly group actors into households, then household networks would ‘join’ otherwise disparate parts of

⁵³ In practice, one would set the variance of the distribution of w_i arbitrarily close to zero (see section: *Generalising over the Distribution of Agent-Characteristics*, below).

the social network. Households in real FGM communities may or may not ‘bridge’ parts of the social network. To allow flexibility in this part of the model, I changed the way households are constructed. In the general model, the population of agents are divided into $\eta_1 (> 1)$ partitions of the network space and then agents are randomly grouped into households of size $\mu_{hsize} (> 1)$ within each partition. As η_1 increases above 1, households will tend to contain agents who are *also* closely connected in the social network.

The final question is the definition of the decision-maker reference group ($\{\text{decision-maker}\}_i$). In the standard coordination model, this is all decision-makers in the population. In the general model, I used a network definition, based on the spatial algorithm. Specifically, $\{\text{decision-maker}\}_i$ is all decision-makers within a certain distance of actor i in network space. However, given the implication that the influence of decision-makers is not based on their *relationship* with actor i but their role as an indicator of the proportion of girls being cut, I allowed for the possibility that the set of actors included in the $\{\text{decision-maker}\}_i$ reference group was larger than the range included in the social reference group. So, where $\{\text{social-ref}\}_i$ was defined as all actors within distance r of actor i in network space, $\{\text{decision-maker}\}_i$ was defined as all decision-makers within distance $r \cdot s_8$ ($s_8 \geq 1$) of actor i .

Generalising over the Distributions of Agent-Characteristics

There were a range of important characteristics (a.k.a attributes) of agents considered in the previous chapter. Agent characteristics in the general model are defined here, along with the symbols used to represent them:

- Perception of the value of FGM ($H_i \in [-\hat{M}, \hat{M} \cdot s_2]$)
- Authority ($w_i \in [0,1]$)
- Autonomy ($\alpha_i \in [-1,1]$)
- Role as a decision-maker ($m_i = 0/m_i = 1$)
- Role as a pro-FGM norm enforcer ($\beta_1 = 0/\beta_1 = 1$)

In the general model, I also added a role as an *anti*-FGM norm enforcer ($\beta_2 = 0/\beta_2 = 1$).

Next, I discuss how I generalised over the possible distribution of these characteristics in the population of agents.

w_i , α_i and H_i are continuous variables with values bounded within a fixed interval. In Chapter 5 I assumed that w_i and α_i were uniformly distributed. I also assumed that H_i was distributed according to one of five distinct distribution shapes (see Chapter 5, Table 3). In the general model, I aimed to generalise over the possible continuous distributions of these characteristics.

To do this, I used the beta-distribution. The beta-distribution family is a highly flexible family of continuous probability distributions, which can accommodate a wide range of ‘shapes’. Importantly, the beta-distribution family can accommodate all of the five important distribution shapes discussed in Chapter 5⁵⁴. All five of these distribution types can be seen as special cases of the beta distribution with particular means and variances⁵⁵, as illustrated in Figures 29-39. These show an annotation of the parameter space of the beta-distribution (defined in terms of mean and variance) and provide examples of each distribution ‘type’, implemented using the beta-distribution.



Figure 29: The Parameter Space of the Beta-Distribution (with annotations showing parts of the parameter space at which the shape of the beta-distribution corresponds, roughly, to the five ‘shapes’ discussed in Chapter 5, Table 3)

⁵⁴ In fact, the beta-distribution was used to represent these shapes in the previous chapter.

⁵⁵ Note that the beta-distribution is typically defined in terms of arbitrary unbounded parameters α and β , see Appendix D3 for proofs related to its re-definition in terms of (bounded) mean and variance parameters. Note also that Efferson et al. (2019) use a beta-distribution to generalize over threshold distributions, but rely on parameterizing in terms of α and β .

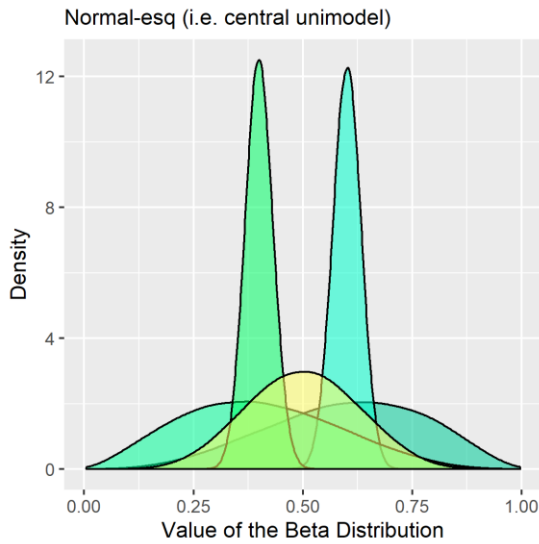


Figure 30: Beta-Distribution Density: 'Normal' Distribution (i.e. central and unimodal distribution using the beta-distribution family, 10^6 MC Samples w. kernel smoothing)

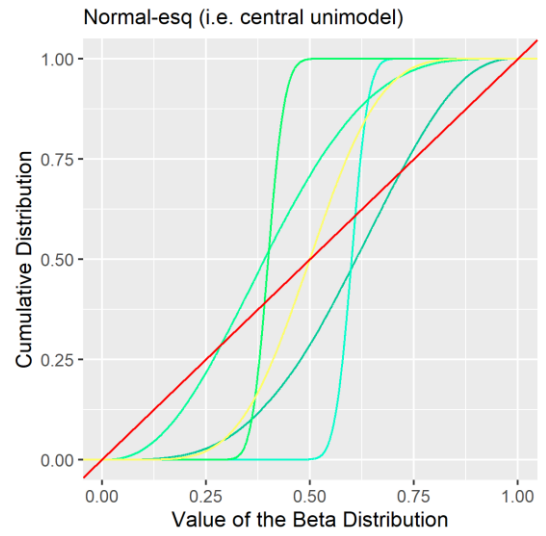


Figure 31: Beta-Distribution CDF: 'Normal' Distribution (i.e. central and unimodal distribution using the beta-distribution family, 10^6 MC Samples w. kernel smoothing)

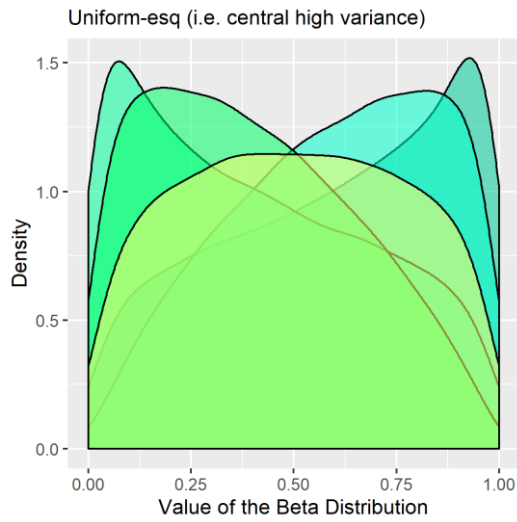


Figure 32: Beta-Distribution Density: 'Uniform' Distribution (i.e. central w. high variance distribution using the beta-distribution family, 10^6 MC Samples w. kernel smoothing)



Figure 33: Beta-Distribution CDF: 'Uniform' Distribution (i.e. central w. high variance distribution using the beta-distribution family, 10^6 MC Samples w. kernel smoothing)

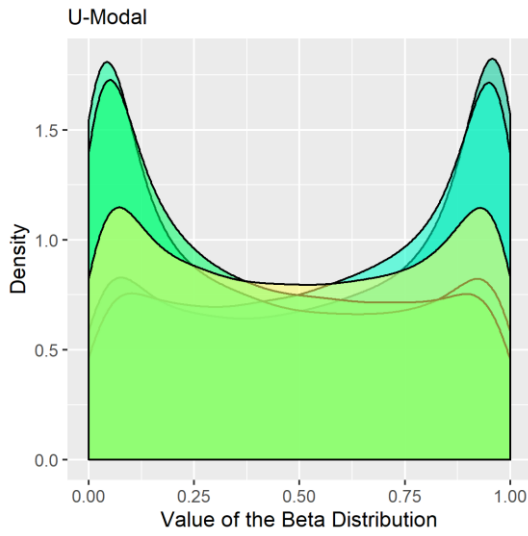


Figure 34: Beta-Distribution Density: 'U-Modal' Distribution (using the beta-distribution family 10^6 MC Samples w. kernel smoothing)

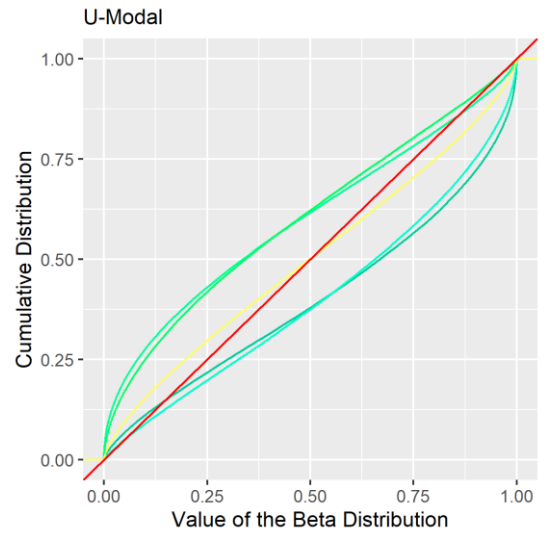


Figure 35: Beta-Distribution CDF: 'U-Modal' Distribution (using the beta-distribution family 10^6 MC Samples w. kernel smoothing)

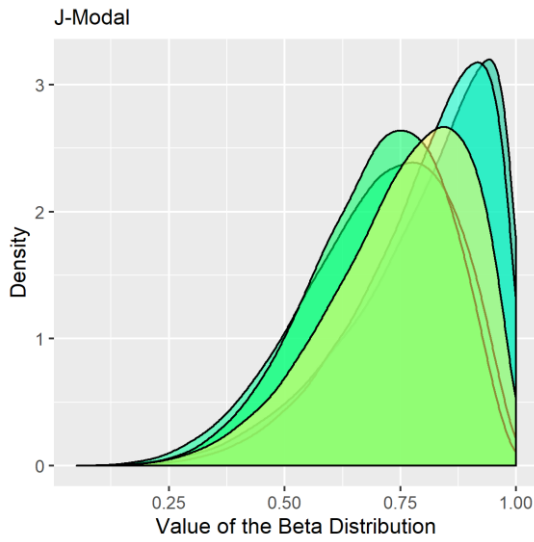


Figure 36: Beta-Distribution Density: 'J-Modal' Distribution (using the beta-distribution family 10^6 MC Samples w. kernel smoothing)

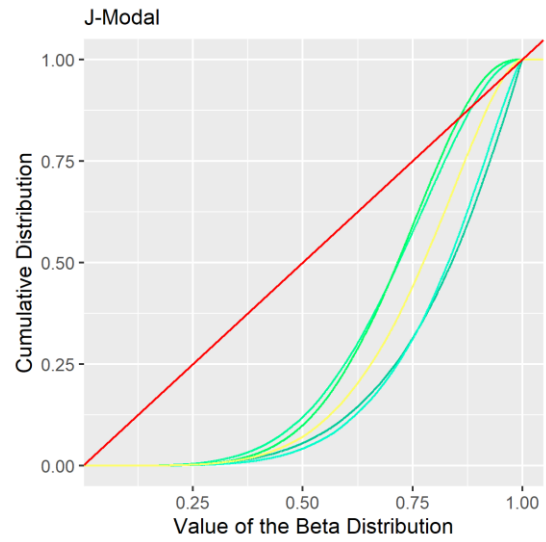


Figure 37: Beta-Distribution CDF: 'J-Modal' Distribution (using the beta-distribution family 10^6 MC Samples w. kernel smoothing)

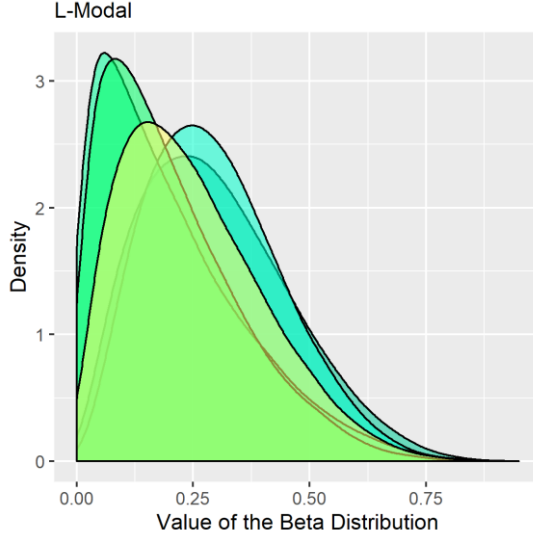


Figure 38: Beta-Distribution Density: 'L-Modal' Distribution (using the beta-distribution family 10^6 MC Samples w. kernel smoothing)

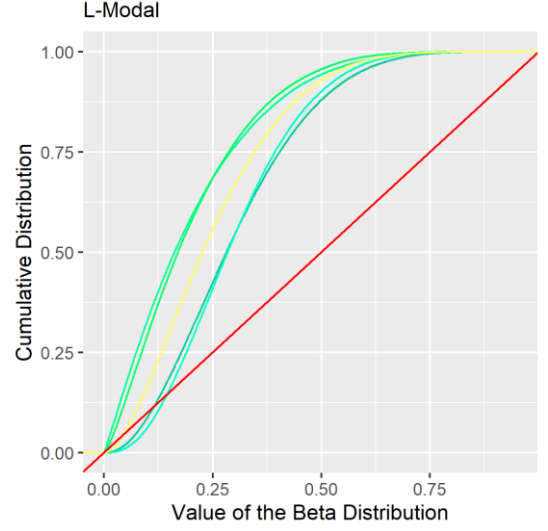


Figure 39: Beta-Distribution CDF: 'L-Modal' Distribution (using the beta-distribution family 10^6 MC Samples w. kernel smoothing)

In the general model, w_i , α_i and H_i were beta-distributed with a flexible mean and variance controlled by two parameters (Table 5). This allowed the model to be highly flexible with respect to the distribution of these features.

Table 5: Parameters Controlling the Marginal Distribution of Agent-Characteristics

Agent Attribute	Parameter for Mean	Parameter for Variance
Authority (w_i)	$\mu_w \in [0.2, 0.8]$	$\sigma_w^2 \in [0.001, 0.111]$
Autonomy (α_i)	$\mu_\alpha \in [0.2, 0.8]$	$\sigma_\alpha^2 \in [0.001, 0.111]$
Intrinsic Value of FGM (H_i)	$\mu_H \in [0.2, 0.8]$	$\sigma_H^2 \in [0.001, 0.111]$

Readers should note that while the three agent characteristics in Table 5 were beta-distributed, two of them were also scaled (so that their values were not restricted to the $[0, 1]$ range of the beta-distribution). Recall that the definition of H_i is

$$H_i = -\hat{M} + (q \cdot [\hat{M} + (\hat{M} \cdot s_2)])$$

where q is a beta-distributed random variable. Thus, μ_H is, strictly speaking, the mean of the random component q (i.e. μ_q). The actual mean of H_i is also determined by \hat{M} and s_2 .

A similar logic applies to α_i which is defined:

$$\alpha_i = -1 + q_2 \cdot 2$$

Where q_2 is the beta-distributed random component of α_i .

The key point is that controlling the marginal distribution of continuous agent attributes using beta-distributions (parameterised in terms of mean and variance) allowed the model to generalise over a very wide space of distributions of these characteristics, encompassing those discussed in earlier chapters and used in past modelling of FGM.

In Chapter 5, I also made assumptions about the relationship *between* attributes. I assumed, for example, that w_i has a perfect positive rank-correlation with H_i , such that actors who approve of FGM also have the most authority. In order to generalise over the possible relationships between continuous attributes of agents in the model, I added parameters to control the rank-correlations between them⁵⁶. This is summarised in Table 6.

In addition to considering relations between H_i and α_i/w_i , I added a new kind of possible relationship: between H_i and y_i , where y_i represents the y-axis position of the agent in network space (see spatial network algorithm, Appendix C3). Allowing for a relationship between these attributes made it possible to introduce homophily of preferences into the social network of the population⁵⁷. This feature wasn't tested in Chapter 5, yet modellers have found that it can have an important impact on the dynamics of coordination models of FGM (Efferson et al., 2019).

Table 6: Parameters Controlling the Joint Distribution of Agent-Characteristics

Relationship between attributes of agents	Parameter
Rank correlation between H_i and α_i	$\rho_{H\alpha} \in (0,1]$
Rank correlation between H_i and w_i	$\rho_{Hw} \in (0,1]$
Rank correlation between H_i and y_i	$\rho_{Hy} \in (0,1]$

Parameterising the correlation between these attributes allows the model to generalise over different relationships between them. Models in previous chapters assumed a perfect correlation (≈ 1). Crucially, the above parameters were implemented such that there was no effect on the marginal distribution of each characteristic (see Appendix D4 for details). Thus, the marginal and joint distributions of the agent attributes could be manipulated separately.

⁵⁶ I assumed that this correlation is either positive, or absent.

⁵⁷ Note that this also required removing the 'vertical' wrapping of the 2D network space, which (strictly speaking) undermines the proof in Appendix C3, nevertheless, the formula controlling average connectivity ($\mu_{Rsocial}$) still works very effectively for practical purposes.

This means, for example, that there was no effect on the spread of actors across network space. y_i was always random uniformly distributed, regardless of its relation to H_i .

Other important agent attributes in the model were discrete and binary, rather than continuous. As such, I used a different approach to generalising over their distributions.

The assignment of the role of ‘decision-maker’ to agents in the community had a simple implementation. $1 + x$ agents in each household (where x is a Poisson distributed random variable with mean 1) were assigned the role of decision-maker. Thus, the concentration of decision-makers in the community depended directly on the size of households (controlled by μ_{hsize} , see above).

Rules for assignment of the role of pro-FGM or anti-FGM norm enforcer were more sophisticated. I assumed that actors would only become pro-FGM enforcers if they actually believed it had intrinsic value. Likewise, I assumed that actors would only become anti-FGM enforcers if they believed that it was intrinsically costly. Then I defined two parameters: s_6 and s_7 , which control the proportion of FGM *supporters* (those believing it to have intrinsic value) and FGM *opposers* (those believing it to be intrinsically costly), respectively, who assumed the role of norm-enforcers. Using these parameters, it was possible to generalise over the different levels of pro-FGM and anti-FGM enforcement.

In Chapter 5, I assumed that pro-FGM norm-enforcers are those who attributed the highest value to FGM. In the general model, I allowed for one of three assignment rules for pro-FGM and anti-FGM enforcers. These were:

- *Random*: enforcers are chosen at random
- *Zealots*: norm-enforcement roles are assigned according to agents’ perception of the value of FGM (H_i). So, those with the highest perceived value of FGM (for pro-FGM enforcers) or lowest perceived value of (for anti-FGM enforcers) were chosen as enforcers.
- *High-authority*: Agents with the highest authority (among supporters and opposers of FGM) are chosen as enforcers.

These choices were controlled by categorical parameters a_1 and a_2 .

Generalising over the Process Flow of the Simulation

Broadly speaking, the flow of the processes in the general simulation model followed the pattern of the standard coordination model. Each time-step, all agents decide whether to

practice or abandon FGM. However, the model was designed to be flexible with respect to two aspects of this process.

First, as in Chapter 5, decision-maker agents only update their ‘final decisions’ about cutting girls in their household with a given probability in each time step⁵⁸. As discussed in Chapter 5, this probability was based on a number of objective features of real communities. It was also based on the largely arbitrary choice of the number of time-steps that occur ‘per-year’. In Chapter 5, I assumed there were 12 time-steps per year and calculated the probability of decision-making per-step accordingly. In the general model, this was controlled by a flexible parameter η_2 which controlled the number of time-steps per-year (with impact on the probability of decision-making updated accordingly).

Second, for all models in Chapter 5, I assumed that actors make decisions in random sequential order. This is the ‘move order’ of the model. In the general model, I allowed for up to five possible kinds of move order, as summarised in Table 7, and controlled by categorical parameter a_4 .

⁵⁸ Note that they still update their decision about FGM practice in each time-step, their decision about cutting a girl in their household is updated to match their decision about FGM practice, with a given probability.

Table 7: More-Order Rules Implemented in the Simulation

Move order rule (a_4)	In each time step...
a_4 = random-sequential	Agents update their decisions in a random-sequential order
a_4 = simultaneous	Agents update their decisions simultaneously, based on the previous time-step
a_4 = autonomy	Agents update their decisions in descending order of autonomy (α_i)
a_4 = intrinsic-value	Agents update their decision in descending order of their perception of the value of FGM (H_i)
a_4 = authority	Agents update their decisions in descending order of their authority (w_i)

Generalising over the Simulated Intervention Process

In Chapter 5, I defined a new ‘standard’ representation of the intervention process that included coalition formation. I also considered a range of elaborations of this representation including:

- Heterogeneity of responsiveness
- Organized Diffusion
- Building social relations between coalition members

The general model incorporates all of these elements but parameterises them so that the extent of their role in the intervention process is flexible, as follows.

Heterogeneity of Responsiveness

The probability that an agent is will agree to join the initial coalition of FGM abandonment is defined as follows.

$$\Pr(\text{join initial coalition})_i = \begin{cases} -\left(\frac{1-z_3}{\widehat{M}} \cdot H_i\right) + z_3 & \text{if } H_i < 0 \\ 0 & \text{if } H_i \geq 0 \end{cases}$$

This probability is a function of a parameter $z_3 \in [0,1]$ as well as the H_i attribute of agents. I assume that agents will only be willing to join the coalition if they have actually been persuaded to view FGM as intrinsically costly ($H_i < 0$). If z_3 is 0, the probability of joining the intervention strongly decreases with increasing H_i values. In this case, the probability will be $-\left(\frac{H_i}{\widehat{M}}\right)$ for H_i values below 0.⁵⁹ If z_3 is 1, then the probability becomes 1 for all actors for whom H_i is less than 0. As such, z_3 controls the extent of heterogeneity of responsiveness.

⁵⁹ Note that this value is always between 0 and 1 and is increasing as H_i becomes more negative.

It can be thought of as the *minimum* probability that an actor who opposes FGM will join the initial coalition.

Organised Diffusion

The organised diffusion process is implemented as described in Chapter 5, with the number of others that each agent attempts to recruit controlled by a parameter z_5 . When z_5 is 0, there is no organised diffusion.

Building relationships between coalition members

In Chapter 5, I implemented relationship-building as a feature of the simulated by connecting all initial coalition members in the social reference network. In the general model, the extent of this process is controlled by a parameter z_6 . Once the initial coalition is formed, each agent forms a connection with a randomly chosen z_6 proportion of other coalition members, and adds them to their social reference group. When z_6 is zero, no new relationships are created.

Other Features

The general model also introduces two other novel options into the representation of the intervention process. Since agents in the general model can be defined by a whole range of different attributes, the model also allows that there could be a whole range of different *targeting biases* in community-interventions. Interventions might (deliberately or otherwise) target supporters of FGM, or they might target those in the community with the most authority, and so on. The options supported in the model are summarised in Table 8 and controlled by categorical parameter z_6 . These options include those used in Chapter 5.

Table 8: Intervention Targeting Biases in the General Model

Targeting bias (z_6)	Targeting on...
Random	Randomly Chosen
Supporters (of FGM)	H_i descending order
Opposers (of FGM)	H_i ascending order
Autonomy	α_i descending order
Authority	w_i descending order
Pro-FGM enforcement	Pro-FGM enforcers $\beta_{1i} = 1$ selected
Anti-FGM enforcement	Anti-FGM enforcers $\beta_{2i} = 1$ selected
Social connectivity	Number of connections in the social network, in descending order
Household Size	Size of $\{household\}_i$ in descending order
Network Localisation	The x-axis position of the agent in network space (x_i) in descending order
By Household	A proportion z_1 of households are targeted, and all actors in those households participate in the intervention

The final feature included in the representation of the intervention process was the creation of *new* anti-FGM enforcers. This recognises the new role of anti-FGM enforcers in the general model. It also reflects observations of real anti-FGM interventions, which have sometimes included attempts to generate new practices of enforcing FGM abandonment (UNICEF, 2008; Diop et al., 2008). This was implemented in the simulation as follows.

Any actors in the initial coalition lose the role of pro-FGM enforcer⁶⁰. Then, with probability $z_7 \in [0,1]$, randomly chosen actors who are *also* not anti-FGM enforcers already, will become anti-FGM enforcers ($\beta_2 \rightarrow 1$). Thus, it is possible for the intervention to create new anti-FGM enforcers, although this will not occur when z_7 is set to 0. Also, it will only have an effect if explicit norm enforcement plays some role in coordination dynamics (i.e. $s_5 > 0$).

⁶⁰ Note that all actors who join the initial coalition have will come to view FGM as intrinsically costly.

Global Sensitivity Analysis

As noted at the beginning of this chapter, a considerable advantage of building a general model prior to empirical calibration is that this permits global sensitivity analysis of the space of possibilities captured by the general model. In the general model, different design possibilities have been parameterised (i.e. mapped to the parameter space of the model). Therefore, we can use a global sensitivity analysis of the parameters in the model to identify key sources of uncertainty in its design. If uncertainty about a particular parameter drives uncertainty about the predictions of the model, then it is clear that efforts should be focused on the calibration of this characteristic. The remainder of this chapter reports the global sensitivity analysis of the general model.

To perform a global sensitivity analysis of the general model, I used Sobol Sensitivity Indices (Saltelli et al., 2008; ten Broeke et al., 2016). Sobol Sensitivity Indices use a variance-based approach to estimate the importance of different parameters in a simulation model. In essence, uncertainty about the dynamics of a model is equated with the variability in one or more of its key outputs.

Sobol Sensitivity indices are something of a gold-standard in a sensitivity analysis. Most importantly, they are a model-free metric. This means that they require no assumptions about the form of the relationship between the input parameters and outputs of a model (Saltelli et al., 2008). The most important measures are the first-order and total-order sensitivity indices of a parameter. These have a straight forward interpretation. Also, in the case of total-order sensitivity, they can take into account *possible interactions* between parameters. The first-order sensitivity for parameter ' x '; is defined as:

“the reduction of the model variance that would occur, on average, if the parameter became exactly known” (ten Broeke et al., 2016: 3.6)

The total order sensitivity index takes into account all possible interactions between parameters and is defined as:

“the proportion of the variance that would remain, on average, when all other parameters are exactly known” (ten Broeke et al., 2016: 3.6)

Taken together, these metrics provide a good indication of the importance of a parameter in driving uncertainty about the outputs of a model. The total-order sensitivity has the particular advantage that it takes into account the *interactions between* some parameter ' x ' and other parameters in the model.

Simulated Experiment: Global Sensitivity Analysis

The key output of the model used to quantify uncertainty was the rate of FGM practice in the agent-population after a simulated intervention. This relates to the key policy *application* of the model, which is to explore intervention failure (see Chapter 8). To measure this, runs of the ‘Dynamics of Intervention’ experiment from chapter 5 were used. However, all parameters, including the size of the intervention (etc.) were chosen using Sobol sampling (Saltelli et al., 2010). This is a way of maximising the accuracy of the Sobol Sensitivity estimates using as few simulations as possible.

I used 400 samples of each parameter (based on a range of plausible values for each parameter), which, using the Sobol sampling scheme, corresponds to 13,200 simulation runs in total. The analysis was conducted using the GNU R ‘nlrx’ package (Salecker et al., 2019) to interface with the ABM model in Netlogo. Unfortunately, this package does not support Sobol sampling for categorical variables, so parameters a_1 , a_2 , a_4 and z_6 had to be excluded from the analysis⁶¹. These were randomised in each run of the simulation, so the results do not depend on particular values of them.

Although the sensitivity analysis included variables related to the intervention process, I did not consider these eligible for calibration. Reducing uncertainty about these parameters would conflict with the eventual application of the model to identifying possibilities of intervention failure. The aim of this later analysis was to identify possibilities (including in the structure of the intervention itself) that might undermine practitioners’ efforts, *not* to represent some particular intervention that occurred in the past.

Table 9 reports estimates of the Sobol Sensitivity indices for 25 key parameters in the simulation, including 95% confidence intervals based on bootstrap sampling (200 samples). Table 9 excludes the six parameters related to the intervention itself. These are reported in Appendix D5, which shows that all 6 had a meaningful effect on the outputs of the intervention.

⁶¹ These controlled assignment rules for norm-enforcement (a_1, a_2), the move-order of the simulation (a_4) and the targeting bias of the intervention (z_6).

Table 9: Global Sensitivity Analysis of the General Model

Parameter	Description	Parameter Range	Sobol Total Order Sensitivity [95% CIs]	Sobol First Order Sensitivity [95% CIs]
μ_H	Average perceived intrinsic value of FGM in the population (H_i)	[0,2,0.8]	0.496 [0.371,0.618]	0.143 [0,0.277]
s_2	Maximum perceived intrinsic value of FGM (relative to perceived cost)	[0,1]	0.348 [0.223,0.49]	0.036 [0,0.14]
δ_1	Relation between social pressure and cost to abandon FGM (higher values mean social costs increase faster)	[-1,1]	0.231 [0.131,0.331]	0.019 [0,0.098]
s_1	Maximum social cost to practice FGM (relative to maximum cost to abandon)	[0,1]	0.214 [0.107,0.309]	0 [0,0.089]
δ_2	Relation between social pressure and cost to practice FGM (higher values mean social costs increase faster)	[-1,1]	0.206 [0.101,0.296]	0 [0,0.044]
V	Maximum non-linearity of the relation between social pressure and social costs	[1,4]	0.18 [0.077,0.268]	0 [0,0.067]
η_1	Number of vertical partitions of the network space (limits the extent to which household networks 'span' the social reference network)	{1,2, ...,10}	0.178 [0.089,0.268]	0 [0,0.002]
μ_{social}	Average number of network connections of actors in the social reference network	[10,100]	0.17 [0.091,0.252]	0 [0,0.037]
s_3	The relative contributions of household versus wider social reference-group	[0,1]	0.168 [0.061,0.265]	0 [0,0.073]
s_4	The relative contributions of decision-maker reference-group influence, versus family or other social reference-group influence.	[0,1]	0.167 [0.08,0.249]	0.005 [0,0.088]
μ_α	Average autonomy in the population (α_i)	[0.2,0.8]	0.167 [0.075,0.248]	0 [0,0.065]
$\rho_{H\alpha}$	Correlation between autonomy (α_i) and perceived intrinsic value of FGM	[0.05,1]	0.164 [0.08,0.257]	0 [0,0.045]
σ_α^2	Variance of autonomy in the population (α_i)	[0.001,0.111]	0.164 [0.071,0.26]	0 [0,0.066]
ρ_{Hy}	Correlation between the network position of actors (α_i) and the perceived intrinsic value of FGM (H_i)	[0.05,1]	0.161 [0.066,0.248]	0 [0,0.057]
ρ_{Hw}	Correlation between actor's authority (w_i) and their perceived value of FGM (H_i)	[0.05,1]	0.154 [0.079,0.24]	0 [0,0.042]
μ_w	Average authority in the population (w_i)	[0.2,0.8]	0.15 [0.067,0.232]	0 [0,0.053]
n-actors	Number of agents in the simulation	{200, ...,2000}	0.148 [0.057,0.249]	0.006 [0,0.107]
σ_w^2	Variance of authority within the community (w_i)	[0.001,0.111]	0.147 [0.057,0.223]	0 [0,0.037]
s_5	The relative contribution of explicit (i.e norm-enforcement) versus implicit social influence, within the household and social reference groups.	[0,1]	0.146 [0.053,0.22]	0 [0,0.018]
η_2	Number of time-steps per 'year' (used to calculate probability that decision-maker agents will update their decision about cutting a girl in their household).	{1,2, ...,12}	0.13 [0.051,0.205]	0 [0,0.031]
s_8	Scaling factor for the network-reach of actors when constructing the decision-maker reference network.	[1,3]	0.116 [0.03,0.198]	0 [0,0.057]
s_6	The proportion of FGM-supporters who actively enforce FGM practice.	[0,1]	0.116 [0.027,0.196]	0 [0,0.059]
σ_H^2	Variance of perceived intrinsic value of FGM in the population (H_i).	[0.001,0.111]	0.114 [0.032,0.201]	0.015 [0,0.101]
s_7	The proportion of FGM-opposers who actively enforce abandonment.	[0,1]	0.112 [0.04,0.195]	0.021 [0,0.106]
μ_{hsize}	Average number of agents per household.	[1,10]	0.105 [0.029,0.18]	0.013 [0,0.081]

The results of the global sensitivity analysis suggest at least two important conclusions. First, two variables stand out as especially important in their impact on the dynamics of the model. These are μ_H and s_2 . Together, these control the average perceived value of FGM (H_i) for agents in the simulation. They had substantially higher total-sensitivity score than other

parameters. These two parameters were, therefore, a particular focus of empirical calibration efforts in the next chapter. Second, almost all parameters (except H_i) had low first-order sensitivity scores, relative to their total order score. This suggests that much of their effect on the simulation occurred through *interaction with* other parameters (Saltelli et al., 2008).

The results show that μ_H and s_2 play an important role in many aspects of the dynamics of the model. So, insights drawn from the model will have greatest credibility in the context of calibrated values of these parameters. This shouldn't be confused with the idea that other parameters (and model features) have only a limited role to play in the *application* of the model. As I show in Chapter 8, a number of important insights provided by the model depend on the effects of *other* parameters *in combination with* calibrated values of μ_H and s_2 .

Implications of the Results

The results of the sensitivity analysis present a clear agenda for empirical calibration. In the next chapter, I focus, in particular, on the empirical calibration of μ_H and s_2 parameters (along with σ_H^2 and ρ_{Hy} , for reasons discussed in that chapter). Constraining these parameters to empirically supported values helps the *possibilistic failure analysis* application of the model (in Chapter 8) to focus on the subset of failure scenarios which have the greatest credibility.

The approach taken in this chapter could be generally adopted in the development of complex models of social dynamics of FGM. The approach of combining and parameterising different design possibilities can be generally applied using simulation tools (it's always possible to map segments of code to particular parameter values)⁶². Having done this, sensitivity analysis provides a general approach to *prioritising* among these different areas of uncertainty. At a more conceptual level, the construction of a general model, such as this one, ensures that, to a certain extent, uncertainty about model design is *embodied* in the model itself (rather than hidden in arbitrary design assumptions). This makes uncertainty about the model design more transparent to the modeller and their audience. If the model is applied to making positive *predictions* (see Empirical Validation in Chapter 7) it also means that this uncertainty is apparent in the dynamics of the model itself.

⁶² Although I faced difficulties with categorical parameters, other methods, such as regression-based analogues of variance-based sensitivity analysis could be employed to include these.

Chapter 7: Calibration and Validation of an ABM of the Social Dynamics of FGM

Introduction

This chapter represents an important stage in the model development strategy of this thesis. It is also the final stage of model development prior to applying the model to the practical task of *possibilistic failure scenario analysis*. To briefly recap, starting with Chapter 4, I identified social coordination as the best supported ‘theoretical core’ with which to begin model development efforts. In Chapter 5, drawing on variant model design assumptions, theory and empirical research, I explored a range of credible elaborations of a ‘standard’ coordination model. I showed that various possible design choices were *important*. They were important because they had the potential to alter the key dynamics of the model. In Chapter 6, I drew together the variant design features into a single ‘general’ agent-based model. I also parameterised the various design options within the model, meaning that the original social coordination model, and the elaborations explored in Chapter 5, could be considered *special cases* (i.e. particular parameter settings) of this general model. I then used global sensitivity analysis to identify those aspects of the parameter space which contributed *most* to uncertainty about the key dynamics of the model.

The results of this global sensitivity analysis represented a set of *priorities* for reducing uncertainty about the model design. The predictions of the model will have the greatest credibility in the event that they *occur in the context of* values of important parameters that are empirically supported. That, in essence, is the primary goal of this chapter: to (where possible) identify an empirically supported subset of values for the key parameters of the model. This undertaking was meant to ensure that efforts in Chapter 8 to employ the model for policy-relevant analysis could concentrate on an empirically supported region of the model’s parameter space.

Before outlining the specific calibration and validation steps taken in this chapter, some of the key issues associated with these techniques are discussed, with examples from the ABM literature.

Calibration and Validation of ABMs

Issues of calibration and validation are central to the methodology of empirical ABM (Chattoe-Brown, 2014, 2017, 2020a; Edmonds and Meyer, 2013; Gilbert and Troitzsch,

2005; Railsback and Grimm, 2012; Squazzoni, 2012). Although, as noted below, the terminology around these issues is not always consistent. As discussed in Chapter 3, calibration and validation constitute the core of data-driven approaches to reducing uncertainty about the design of a model. However, while calibration and validation are widely recognised activities, there is not necessarily a consensus view of how they should proceed in practice – and different approaches may have different advantages and disadvantages.

Calibration of ABMs

The calibration of ABMs (also called ‘empirical specification’ Squazzoni, 2012) is generally understood as the process of using empirical data to refine and improve the design of the model, often through the selection of empirically supported parameter values, or refinement of other model design elements (Chattoe-Brown, 2017; Railsback and Grimm, 2012; Squazzoni, 2012).

A classic illustration of this kind of activity can be found in the iconic ‘Anasazi model’, which was an ABM of the population development and settlement patterns of the Kayenta Anasazi in the Long House Valley (Arizona, US) between 800AD and 1350AD (Axtell et al., 2006). This model was designed to simulate and study the patterns of settlement location and size within the valley over this historical period. It simulated households who consumed resources and could move to new locations to meet their consumption needs. The accuracy of the simulation was considered to depend, crucially, on an accurate representation of the agricultural productive capacity of different spatial areas in the valley. To *calibrate* this feature of their model, the authors used a range of environmental data sources to estimate parameters such as the annual production of maize that a given area of land within the valley could accommodate at a particular time during the period.

Another distinctive example can be found in Bravo et al. (2012) who conducted experiments in which human participants played economic games, then used this data to help calibrate parameters controlling the socially contingent investment decisions of agents in an ABM of trust and cooperation in economic exchange.

A general illustration of the concept of calibration as the use of empirical data to control model design can be found in Scott et al. (2016). The authors include a calibration table (Chattoe-Brown, 2017b) showing how different aspects of the specification of their ABM of binge-drinking were derived from empirical data (Scott et al., 2016: 10). Parameters in their model, including (for example): the probability of moving between bars, the distribution of starting times for ‘nights out’, and the distribution of 18-21 drinking rates, were estimated

directly from a survey of young people in Melbourne Australia (the Young Adults Alcohol Study, Dietze et al., 2014)

While references to calibration are often about the selection of numerical values, calibration can also be used to mean any attempt to use empirical data to refine the specification of an ABM, especially the low-level details of individual behaviour and interaction (Chattoe-Brown, 2017b). Railsback and Grimm (2012), for example, recommend a form of *indirect* (see below) calibration of ABMs that they refer to as theory development, which involves experimenting with different agent decision specifications to find examples that allow the model to fit macro-level empirical patterns. In general, Squazzoni (2012: 151) notes that calibration (he also refers to this as ‘empirical specification’) can be based on experimental methods, consultation with domain experts, qualitative data or quantitative data.

The key methodological questions in the *application* of calibration within ABM development are (a) whether the calibration should be ‘direct’ and (b) whether the calibration should be ‘independent’

To understand these distinctions, readers should first recall the distinction between micro-level and macro-levels of analysis in ABM research (also discussed in Chapter 1). The micro-level refers to the actions and interactions of agents (often within social structures like networks), the macro-level refers to the aggregate patterns of behaviour arising from this. Generally, the micro-level of an agent-based model is ‘designed’ whilst the macro-level *emerges* from the interaction of micro-level elements (see Epstein, 2006, also Chapter 3).

One popular understanding of calibration is that it involves the use of data (either qualitative or quantitative) *which is about* micro-level features of a real social system in order to specify the corresponding features of a model (also known as the model’s ‘micro-specification’, Chattoe-Brown, 2017b; Epstein, 2006; Squazzoni, 2012). We can think of this understanding of calibration as ‘direct’ because it generally implies the use of empirical data to *directly estimate* aspects of the micro-level specification of a model. Calibration of the Anasazi model (Axtell et al., 2006), noted above, is a clear example of direct calibration. An extreme form of direct calibration is exhibited by Cointet and Roth (2007), who imported the empirically observed co-attendance network of company directors from major US firms directly into an ABM, to control the interaction structure of their model. Direct calibration is also called *submodel parameterisation* (Railsback and Grimm, 2012), where a submodel is a recognisable element of an agent-based model, whose specification is chosen using empirical data about that element.

Like all forms of calibration, direct calibration is designed to reduce uncertainty about the design of specific elements of a model, and ultimately, to reduce uncertainty about the *aggregate dynamics of the target system*. As discussed in detail in Chapters 3 and 6, this reduction in uncertainty is maximised when the parameters targeted are those most responsible for driving the dynamics of the model (i.e. to which it is most sensitive, Chattoe-Brown, 2017; Saltelli et al., 2008). This issue is given formal treatment by Saltelli et al. (2008), under the concept of ‘factor prioritisation’.

The particular advantages of ‘direct’ calibration (discussed below) are (a) that it ensures calibration is ‘independent’ of validation and (b) that it reduces the burden (discussed below) associated with calibration based on fit to validation data, which I refer to here as ‘indirect’ calibration.

Calibration based on fit to validation data, referred to here as *indirect* calibration, occurs when *data associated with the macro-dynamics of interest of a model* are used to select values for specific parameters or other elements of a model’s design. This may occur through the use of an algorithm to discover parameters values which maximise the fit between the macro-dynamics of the model, and relevant macro patterns in the real-world social system (Thiele et al., 2014).

A recent example of this can be found in Brousmiche et al. (2016). The authors had a set of unknown and uncalibrated parameters associated with their model of opinion dynamics, which they sought to apply to a military scenario in Afghanistan. Among these were parameters relating to the internal mental processes of agents in their model. In the absence of data with which to *directly* estimate those values, Brousmiche et al. (2016) used an evolutionary optimisation algorithm to select values for the parameters that maximised the overall fit between the simulated macro-dynamics of their model, and two aggregate time series of public opinion. The characteristic that distinguishes this activity from direct calibration, is that the data (opinion time series) used were not ‘about’ these parameters. Rather, the implicit assumption was that the most likely values for the unknown parameters were those that maximised the overall fit of the model.

The above illustrates the pragmatic advantage of indirect calibration: it facilitates the empirical estimation of parameters (or other model features, see ‘theory development’ in Railsback and Grimm, 2012) that can’t be measured directly. However, therein also lies a *disadvantage* of the approach. The analyst needs to critically consider whether indirect calibration can provide a *credible* estimate of a given set of parameters. Brousmiche et al. (2016), for instance, used two time-series containing only 4 data-points each to estimate 8

model parameters. Moreover, one of these time-series was close to a linear trend. It is far from clear that this kind of macro-level data contains sufficient information to supply strong evidentiary support to the corresponding parameter estimates. Due to these kinds of issues of *underdetermination*, ABM methodologists who support indirect calibration (which is not all of them, see below) generally recommend that this is used sparingly and in combination with direct calibration (Railsback and Grimm, 2012: 257).

A second disadvantage of indirect calibration is that it can affect the modeller's capacity to convincingly validate their model (see below). Since indirect calibration, by definition, uses macro-level patterns in data that are related to the overall dynamics of the model, it uses patterns that could *also* be used for macro-level model validation (see below). The consequence of this will depend on whether the modeller is aiming for *independent* calibration and validation, or not.

Calibration is independent if the particular patterns in data that are used to calibrate the model are *not* the same as the patterns in data that are used to validate it. As I discuss further in the next section, model validation is generally understood as the comparison of macro-level data about a system to the macro-level dynamics of the model (Chattoe-Brown, 2014, 2017, 2019; Gilbert and Troitzsch, 2005, Squazzoni, Epstein, 2007)⁶³. Direct calibration is independent by definition since the data used are 'about' a particular element of the model (a 'submodel') not the overall dynamics that the model is being used to study. Whether indirect calibration is independent, depends on whether the same data are 're-used' later for validation. The indirect calibration used by Brousmiche et al. (2016) was non-independent in this sense because they used macro-level time series data to calibrate key parameters in their model, and then interpreted the fit of their model to this same data as an evaluation of its overall adequacy.

Independent calibration, where possible, is preferred by some ABM methodologists because it subsequently facilitates a stronger validation test that minimises the risk of poor models finding an arbitrary fit to data (see discussion of equifinality in 'Validation of ABMs' below). Chattoe-Brown (2017: 1) goes as far as to assert that "it is the independence of these two activities [calibration and validation] that provides Agent-Based Modelling with its distinctive claim to explanatory power" and later that this "distinctive methodology" is essential in

⁶³ Epstein (2007) doesn't use the term validation, but very clearly identifies the comparison of the macro behaviour of a model to macro data as central to testing its adequacy.

maintaining the “scientific quality of [ABM] research”. Other ABM methodologists, by contrast, support the use of non-independent (and therefore indirect) calibration of ABMs (Railsback and Grimm, 2012). However, they do so in the context of a very particular methodology tailored toward addressing the problem of equifinality in a different way (see discussion of multi-criteria approaches below).

Taking into account the potential advantages of independent calibration, the possible negative consequences of indirect calibration (as it relates to independence) are therefore either (a) that it ‘uses up’ macro-level data that cannot then be re-used whilst maintaining the independence of calibration and validation (see further discussion below), (b) that it prevents calibration from being independent because the macro-level data is then re-used for model validation.

Validation of ABMs

The validation of ABMs is generally understood (in the social sciences) as the process of assessing the adequacy of the model by comparing its outputs to corresponding data about the properties (e.g. aggregate dynamics) of the real-world system that it is being used to model (Chattoe-Brown, 2017; Gilbert and Troitzsch, 2005; Squazzoni, 2012).

A canonical example of an ABM validation test can be found, again, in the Anasazi model (Axtell et al., 2006). The authors compare, for instance, the time series of population size and settlement sizes simulated by their model to an empirical time series of the same quantities estimated for the period under study. There is a strong degree of qualitative correspondence between the two, which might be interpreted as an indication that their model has adequately captured at least some key features of the ‘real’ social system generating development in the Long Horn Valley during this period. Krebs (2017) conducted a similar kind of validation test of his ABM of green electricity adoption in Germany. He compared simulated time-series of the number of households buying green electricity with empirical estimates of the same quantities between 2005 and 2011.

Although comparison of empirical and simulated time-series data is a highly accessible example of macro-level validation, models can potentially be validated using any macro-level pattern that is thought to bear on the key dynamics of the system of interest (Grimm et al., 2005; Railsback and Grimm, 2012). Abdou and Gilbert (2009), for example, validated their model by (among other things) comparing single aggregate values simulated by their model of social and workplace segregation (including segregation indexes and unemployment levels) to comparable empirical estimates from a cross-sectional survey in Egypt.

The overarching issue in model validation is the *extent* to which it establishes the adequacy of a model, especially if it is required to make predictions in a policy context. The constituents of this issue are the threat of equifinality, and corresponding techniques designed to minimise that threat: independence of calibration and validation, and multi-criteria validation (also called pattern orientated modelling).

Sometimes, particularly in applications of relatively stylised ABMs, the comparison of models to data has followed a principle of ‘generative sufficiency’ (Epstein, 2006). This kind of validation assesses the adequacy of a model through testing whether it is *capable of* simulating (i.e. generating) some pattern observed empirically. Such patterns may be highly stylised, such as reproducing spatially segregated clusters (see discussion of Schelling below), or generating a particular general distribution of wealth inequality (Epstein, 2006; Epstein and Axtell, 1996). The logic of this form of validation is that if a model can be shown to reproduce a particular empirical pattern, then the model is *sufficient* to explain the pattern in question, and can be considered a ‘candidate explanation’ for that pattern. Here, no strong constraints are placed on the construction or calibration of the model, except that it should have a plausible specification incorporating the various signature elements of the ABM approach (such as local interaction, Epstein, 2006).

This form of validation makes only weak claims about model adequacy. A model is only considered a ‘candidate’ explanation for the pattern it reproduces – no substantive confirmation of the correctness of the model is established. This interpretation is appropriate because of the threat that equifinality poses for models of complex systems.

Equifinality, in applied ABM research, is the possibility that multiple, otherwise contradictory or substantively incorrect, models might reproduce the same aggregate dynamics (Evans et al., 2017; Poile and Safayeni, 2016; Richardson, 2002). This possibility arises when building agent-based models, due to their complexity and the flexibility available to the modeller in their construction. ABMs often contain large numbers of parameters, and modellers have considerable flexibility in their design choices when building the model (see Chapter 3, Poile and Safayeni, 2016). The concern is that this flexibility might permit the construction of fundamentally inadequate models which *nonetheless* can reproduce a simple aggregate pattern in data, despite containing inappropriate simplifications, unrealistic assumptions and so on (c.f. Polhill and Salt, 2017: Table 8.1). While, for example, Schelling’s famous segregation model reproduces a stylised regularity that is observed empirically (residential segregation based on social group, e.g. race), it is far from clear that the

specification of decision-processes in the model accurately reflects key contingencies of the decision-making of individuals in real US cities (Chattoe-Brown, 2019). The possibility that inadequate models might, nonetheless, pass validation tests, is of particular concern when a model is due to be applied to policy problems. Such an application will almost certainly involve the model making predictions about *novel* scenarios, and thus rely fundamentally on the adequacy (the ‘structural realism’ Railsback and Grimm, 2012) of its specification.

The practices of independent calibration and validation, as well as multi-criteria modelling (a.k.a pattern-orientated modelling), can therefore be seen as attempts to render the test provided by validation more rigorous, and in turn, convey greater assurance of the adequacy of a model.

To illustrate this argument, let us revisit the question of the independence of calibration and validation (see also ‘Calibration of ABMs’ above). The potential advantage of independent calibration and validation is that it minimises the likelihood of a model arbitrarily fitting data. This occurs because the approach provides an independent source of evidence for the micro-specification of a model (Chattoe-Brown, 2014, 2019; Squazzoni, 2012). In principle, an independently calibrated and validated model cannot be said to have been manipulated in order to arbitrarily fit aggregate data (i.e. ‘affirming the consequent’ Badham et al., 2018: 174), instead (if it does fit aggregate data) it can be more convincingly argued that this occurs *because* the model’s design has been calibrated independently on empirical data (e.g. micro-level data on agents’ behaviours, attributes, environment, and so on).

In other words, by clearly separating model design (informed by empirical calibration) and model validation (based on independent aggregate patterns in data), the latter arguably provides a more genuine test of the adequacy of the model – and conveys stronger confirmation of its design. Abdou & Gilbert (2009), introduced above, included elements of independent calibration and validation of an ABM (although they don’t use the term ‘calibration’ specifically). Features of the social network used in their model were, for example, based directly on data about social networks in their target population, and the model was then tested against aggregate figures from the target system (see above) without being ‘fit’ to these data – which was self-evidently different.

By contrast, when calibration and validation are *not* independent, such that a large number of model parameters are fit directly to validation data, the risk that this will permit an inadequate model to fit data may be maximised - since a model with a sufficient number of parameters may be able to arbitrarily fit a range of patterns in data (whilst still providing an

inadequate representation of the underlying processes). Brousmiche et al.'s (2016) model (noted above), for instance, is 'fit' to two simple time-series, each containing only four data-points. The authors suggest that their model's ability to reproduce these time series quite closely is encouraging. However, taking into account the relative banality of the data used (two time-series, one of which is nearly linear, each with only four data points), and the number of parameters fit to that same data (eight), it couldn't be claimed that any substantive confirmation is conveyed by this exercise.

By contrast, the multi-criteria approach (as known as 'Pattern Orientated Modelling' (Grimm et al., 2005; Railsback and Grimm, 2012) tackles the equifinality problem from a different direction. Rather than constraining the design of the model through independent calibration, it focuses on increasing the difficulty posed by the validation test. The basic idea is to assemble multiple distinct patterns in data (multiple criteria) and require that a model reproduce all of them adequately in order to pass validation. The underlying assumption is that using multiple patterns in this way will filter out inadequate model designs. In other words, whilst an inadequate model might be able to reproduce a single pattern, the assumption is that only an adequate 'structurally realistic' model will be able to reproduce multiple (sufficiently distinctive) patterns (Grimm et al., 2005; Railsback and Grimm, 2012: 227).

In theory, multi-criteria approaches have some practical advantages over an approach based on independent calibration and validation because they permit a 'safer' use of indirect non-independent calibration (i.e. fitting a model specification directly to validation data). This potentially allows the modeller to tailor difficult-to-observe parameters (or other model elements) of their model to aggregate data, without estimating these directly, and without engendering concerns about equifinality (see discussion of indirect calibration, above).

The potential disadvantage of a multi-criteria approach is its dependence on the assumption that the set of empirical patterns available for the validation test are sufficient to filter out inadequate models, especially since indirect fit-to-data' calibration is employed. Whilst, other things being equal, a multi-criteria approach almost certainly provides a stronger validation test than validation based on a single stylised regularity, there is no formal way to establish that a given set of patterns are sufficient to ensure that a model that reproduces them will provide an adequate representation of the underlying system.

There is, of course, nothing to prevent modellers from combining independent validation and a multi-criteria approach, such that an independently calibrated model is then tested

against multiple macro-level patterns in data. This would provide a particularly strong test of a model. Railsback and Grimm (2012: 238) acknowledge the special role that independence plays in model validation by suggesting that true validation status is conferred when a model makes ‘independent’ predictions not considered in the model development process (including testing against known patterns), and these predictions are then confirmed by subsequent empirical observation⁶⁴.

Realities in Practice

It is important to contextualise these high-level discussions of the principles of calibration and validation against the reality of common practice in the ABM literature, as well as broader issues of model design uncertainty.

As established through surveys of the literature, and noted by methodologists, any attempt to validate ABMs against data remains the exception rather than the norm (Chattoe-Brown, 2020a). This lack of substantive validation can also be seen among models of the social dynamics of FGM (see my reviews in Chapters 2 and 3). This makes *any* attempt to undertake formal validation using detailed empirical data an improvement on current practice in this field, and among ABMs in general.

Moreover, in real-world projects, choices of different strategies for calibration or validation may be heavily constrained by available data. Independent calibration can only be applied, for example, to aspects of a model for which pertinent data is available (or can be collected by the modeller). Likewise, multi-criteria validation is only possible if multiple distinct and relevant empirical patterns at the macro-level are available for use in testing the model.

Furthermore, the kinds of calibration strategies that are appropriate may depend on the intended application of the model. If a narrow range of predictions is required (in the style of social ‘engineering’, Edmonds and Aodha, 2019) then a modeller may be inclined to employ indirect calibration (i.e. based on fit to data) in order to select plausible values for all key parameters in the model. However, modellers following a possibilistic approach (as I do in this thesis, see Chapter 3) are unlikely to wish to calibrate parameters indirectly. This is

⁶⁴ Note, Railsback & Grimm’s (2012) use of terminology related to calibration and validation is somewhat different from the rest of the ABM literature. The terms ‘validation’ and ‘testing’ are separated (the former reserved for novel predictions only), and ‘verification’ is used to refer to what would typically be called ‘validation’ (testing against known empirical patterns). Whereas, ‘verification’, typically refers to assessment of the quality of the underlying computer code (Gilbert and Troitzsch, 2005). Nevertheless, the methodological concepts, and the authors’ distinctive ‘pattern orientated’ approach to them, are clearly recognisable.

because of the increased risk that the estimated values will be arbitrary or uncertain. ‘Fixing’ their values using indirect calibration would therefore obscure (rather than legitimately remove) uncertainty about the possible dynamics of the target system.

Finally, it is important to recognise the limits of calibration and validation in removing uncertainty about an ABM’s design. This issue is discussed further in Chapter 3. Briefly summarised: calibration of any kind may struggle to remove uncertainty arising from prior choices about simplifications and idealisations in model design. Elements of the real social system abstracted out of a model cannot be calibrated⁶⁵. Moreover, there is never any way to establish definitively that a validation test is sufficient to confirm the adequacy of a model (Polhill and Salt, 2017), especially if it is to be applied to policy analysis involving predictions about novel scenarios (that, by definition, were not part of the validation test).

Approach to Calibration and Validation Taken in this Project

The approach that I adopt in this chapter is based on independent calibration and validation, and (to a limited extent) on multiple-criteria validation. I employed direct and independent calibration of a number of the key parameters in the ABM developed in Chapter 6. I chose independent calibration in order to ensure that subsequent validation tests provided a genuine test of the model. I only used direct (rather than indirect) calibration because the available macro-level data was reserved for the validation test, and because of the risk that a fitting process would generate arbitrary parameter values that would obscure, rather than remove, uncertainty about the dynamics of the model. The modelling strategy employed in the thesis benefits from the fact that the intended application: possibilistic scenario analysis (see Chapters 3 and 8) can accommodate uncertainty about the parameters in the model. Thus, it was not necessary to arbitrarily constrain parameters for which direct estimates were not available.

Direct calibration was undertaken by estimating the values of key model parameters (and the distribution of those values) using national household survey data from Senegal (see details below). Validation of the general ABM was undertaken as follows: estimated ‘parameters’ about the characteristics of individual communities in Senegal (*not* including their rate of FGM) were ‘input’ to the ABM, which then ‘output’ a prediction about the rate of FGM practice in those communities. The ‘fit’ between the simulated rates of FGM in the ABM

⁶⁵ The example that I use in Chapter 3 is that we cannot calibrate the representation of a social network in a model if we don’t choose to include one in the first place!

and the observed rates of FGM in the ‘input’ communities, was the primary validation test. Next, using calibrated parameters, the capacity of the ABM to reproduce empirical patterns in the marginal and joint distributions of characteristics of FGM practising communities was tested. This can be considered a (limited) example of a multi-criteria validation test since it involved multiple tests of the ability of the model to reproduce separate macro-level patterns.

As noted in the previous section (and Chapter 3), calibration and validation do not eliminate uncertainty about the adequacy of model design, although they increase our confidence in it. As such, I do not claim that either the calibration or validation procedures employed in this chapter convey ‘final’ confirmation of the correctness of the model. Quite apart from the uncertainties present in these procedures themselves (see below), they both rely to a significant extent on assuming that the model is sufficiently realistic that the features of the target (real communities practising FGM in Senegal) can be construed in terms of the parameters and outputs of the model. The entrenched questions of idealisation and simplification (in this and other models, see the previous section, and Chapters 1-3) mean that such a premise is difficult to firmly establish (although it is necessary for the modelling exercise). Instead, I argue that both calibration and validation *contribute* to improving the model and to establishing the basic credibility of the failure scenarios identified in the next chapter (8), with the *extent* of this contribution remaining difficult to quantify.

Past Calibration of Models of the Social Dynamics of FGM

To the best of my knowledge, only one previous attempt to empirically calibrate the parameters of a model of the social dynamics of FGM has been undertaken. This is reported in Novak (2016). She began with a heterogeneous threshold model, then used survey data from Burkina-Faso to try to estimate upper and lower bounds for the distributions of thresholds in different provinces in Burkina-Faso (direct calibration). Unfortunately, however, the assumptions of her empirical approach excluded important model design possibilities identified in previous chapters. The key issue is that she operationalised actors’ ‘thresholds’ in terms of the proportion of girls being cut (historically) in the same province (or same ethnic group and same province). This makes the strong assumption that the reference group of all actors is all other cutting families in the same province. It, therefore, excludes the possibility that social interaction is localised within social networks. It also assumes that social pressure stems only from decisions about cutting, rather than other kinds of interaction, such as explicit norm-enforcement. Empirical research suggests that the notion of localised interaction (i.e. within local social networks) and explicit norm

enforcement are highly credible (Mackie and LeJeune, 2009; Shell-Duncan et al., 2011; Shell-Duncan and Hernlund, 2006).

The further problem is that the concept of a singular decision threshold is meaningful for all actors. As I noted in Chapter 6, this will not necessarily be the case in all plausible ‘social norm of coordination’ models of FGM.

Instead, the assumptions necessary for the empirical estimation used in this chapter do not impose strong assumptions about the form of social influence. This is primarily because I focus on measuring the *components* of individuals’ utility functions (e.g. their perception of the value of FGM: H_i), rather than the more abstract notion of a ‘threshold’⁶⁶.

Targeting Parameters for Calibration

The foremost question in developing a calibration strategy for the general ABM, was *which* parameters should be calibrated. The previous chapter provided a set of priorities for calibration. In particular, the global sensitivity analysis highlighted parameters related to the distribution of the perceived intrinsic value of FGM (H_i and s_2) as having a key impact on the dynamics of the model. Ideally, all of the key parameters in the model would be calibrated. However, the scope of independent calibration that could be undertaken depended on the availability of suitable empirical data.

The primary source of data used in the project was a publicly-accessible cross-sectional survey, undertaken in 2005 as part of the Demographic and Health Survey (DHS) program, of communities in Senegal (see description of the survey in Chapter 4). This survey included random household sampling within localised clusters (census enumeration areas – typically a large village or city-block) and a number of measures of beliefs and outcomes related to FGM. The local-cluster design of the survey had a considerable advantage that community-level properties could (potentially) be estimated from it.

Across the ten parameters generating greatest uncertainty about the outputs of the ABM (see Table 9, Chapter 6) the available data potentially allowed for quantitative calibration of the two *most important* parameters. There were s_2 (the ratio of the maximum perceived intrinsic

⁶⁶ This also underscores my argument (Chapter 5), that it is unhelpful to try to abstract decision-making away by relying solely on thresholds.

value of FGM, relative to the maximum perceived intrinsic cost) and μ_H (mean perceived intrinsic value of FGM – expressed in the unit interval).

Although *not* estimated to be among the most important parameters in the simulation, I also undertook empirical calibration of the σ_H^2 and ρ_{Hy} parameters. σ_H^2 controls the variance of the distribution of the intrinsic value of H_i in the agent population. Thus, taken together, μ_H , σ_H^2 and s_2 fully define the distribution of H_i in the simulation. Together, these deal with an aspect of coordination models of FGM (the distribution of preferences) which has been the focus of the most attention in past modelling. I also undertook calibration of ρ_{Hy} (which determines the homophily of preferences in the social reference network of agents), because, despite the results of the sensitivity analysis, homophily has been found to have an important impact on intervention in other recent analyses (Efferson et al., 2019).

μ_H , σ_H^2 , ρ_{Hy} and s_2 could be estimated using the available data because they could be estimated, in principle, from cross-sectional data, *and* relevant measures existed in the Senegal 2005 DHS for that estimation. Other key parameters could not be measured because they required either longitudinal data, or additional measures (e.g. introspection of actors' decision-processes), or both. Measuring, for instance, the relationship between social pressures and social costs (see Chapter 6), would require either (a) longitudinal data on the social environment and decision-making of multiple individuals (not available), or (b) some way of eliciting this information directly from social actors (i.e. introspection).

At first glance, the availability of calibration data for only two out of the ten most important parameters may seem to undermine the value of the exercise. However, it is important to recognise that the two *most important* parameter *could* be calibrated, and that the contribution of these two parameters to uncertainty in the outputs of the model was disproportionately large. The combined total-order sensitivity indices (see Chapter 6) of the two most important parameters: μ_H and s_2 was 0.844. This is more than the combined total-order sensitivity of the next four most important variables combined.

The key intellectual challenge in calibrating the parameters in the ABM was to develop *estimation procedures* (a.k.a estimators) that could be applied to the available empirical data. Since the empirical estimation of ABM-specific parameters is not a 'standard' problem, bespoke strategies needed to be employed. In developing such strategies, I considered the following four questions:

1. Given cross-sectional data from the ABM (i.e. an ABM simulated population) could these parameters be estimated, and if so, how?
2. Were suitable measures in the empirical data available to create analogous *empirical* estimators for these parameters?
3. What assumptions were necessary to facilitate empirical estimation?
4. Given the necessary assumptions, what level of uncertainty in the empirical estimators could be anticipated and quantified?

Framing the problem in this way reveals an otherwise implicit premise in the calibration exercise – that the target population is ‘like’ the model (in a broad sense). Put another way; it was necessary to assume that the parameters ‘exist’ and are measurable (even if they weren’t directly observable). Although not necessarily always framed in the same terms, this is a conventional assumption. It is the same assumption that, for instance, past attempts to calibrate the similarity-preference parameter of the Schelling Segregation model have made (Clark, 1991). Such attempts rest on a prior theoretical commitment that this is a meaningful quantity than could, in principle, be measured. To assume otherwise at this stage, would be to re-litigate the core design of the model (i.e. to return to the previous three chapters)⁶⁷.

In the following sections, I discuss strategies for estimating μ_H and σ_H^2 , as well as s_2 , and ρ_{Hy} , addressing each of the above four questions. The subsequent parts of this chapter detail the actual empirical estimation of the parameters, the generation of a calibrated parameter space and then empirical validation of the ABM using these calibrated parameters.

Strategy for Quantitative Estimation of μ_H and σ_H^2

μ_H and σ_H^2 correspond to two aspects of the distribution of the same characteristic of the simulated population: q_i , representing the underlying intrinsic value that the actor i attributes to FGM, scaled to the unit interval (recall that the distribution of q_i is scaled to the interval $[-\hat{M}, \hat{M} \cdot s_2]$ to generate the distribution of H_i , so strictly speaking μ_H represents μ_q , with the true mean of H_i depending on another parameter: s_2).

⁶⁷ Of course, measurable in principle does not mean measurable in practice, and certainly does not mean estimated without uncertainty

Estimation from ABM Simulated Data

In principle, μ_H and σ_H^2 can easily be estimated from a cross-section of the simulated population itself.

The estimator for μ_H ($\widehat{\mu}_H$) is simply:

$$\widehat{\mu}_H = \sum_{i \in n} \frac{q_i}{n}$$

Where $i \in n$ indexes the simulated population. This estimator is unbiased and has a known standard error (based on the community size and the observed standard deviation of q_i).

Similarly, the estimator for σ_H^2 is the variance of q_i in the simulated population, for which a standard error formula is also available.

Analogous Empirical Measures

The difficulty in converting this to an *empirical* measure is that H_i and q_i are *not* observed directly for actors in the real population. Instead, H_i is a *latent* construct (representing the overall perceived intrinsic value of FGM). Thus far we have not discussed the underlying process which generates H_i in actors (the model focuses on the implications of such preferences for social dynamics, not where the preferences come from originally), except to assume that it depends on actors' beliefs about FGM, and that H_i is positive when FGM has a perceived positive intrinsic value, and negative when FGM has a perceived negative intrinsic value.

Fortunately, however, the 2005 Senegalese DHS survey contains *both* (a) detailed measures of actors' beliefs about FGM (as well as potential mediating factors such their level of education) *and* also (b) measures of actors' beliefs about whether FGM should continue, or stop (which can be interpreted as an indicator of whether H_i is positive). As such, it may be possible to estimate H_i *indirectly*, using a latent variable model.

The argument underlying this approach proceeds as follows. There is an underlying latent variable H_i , which represents actor i 's perception of the intrinsic value of FGM. H_i is not observed, however, we observe an indicator, which is 1 when $H_i \geq 0$, and 0 otherwise. H_i itself depends on a linear combination of a number of determinants (e.g. beliefs) which *are* observed. As such, we define H_i as

$$H_i = \beta_0 + \beta_1 x_{1i}, \dots, \beta_p x_{pi} + \epsilon_i$$

where x_{pi} are observed beliefs (and other relevant characteristics) of actors, and ϵ_i is unmodelled error. To estimate H_i , we need to estimate the beta-coefficients of the model. However, we cannot do this directly, since H_i is unobserved. Instead, the challenge is to model H_{obs_i} (which is an indicator that $H_i \geq 0$), as a function of this linear combination:

$$H_{obs_i} = f(\beta_0 + \beta_1 x_{1i}, \dots, \beta_p x_{pi}) + \epsilon_i$$

Assuming that ϵ_i has a logistic distribution leads us naturally to logistic regression as an estimation strategy, in which values for the beta coefficients are selected which maximise the likelihood of the model (Glasgow and Alvarez, 2009). This is sometimes called the latent-variable interpretation of logistic regression. The sum of terms (a.k.a. the linear predictor) is then our estimate of H_i is

$$\widehat{H}_i = b_0 + b_1 x_{1i}, \dots, b_p x_{pi}$$

where b_p are the estimated coefficients.

As such, a potentially viable procedure for estimating H_i in individual actors in the 2005 Senegalese survey (and subsequently, μ_H and σ_H^2 for individual communities), is to regress individual attitudes about whether FGM should continue (H_{obs_i}) on variables determining H_i (e.g. beliefs about FGM), and to take the linear predictor for each case as an estimate of H_i . This estimate will, of course, not necessarily be on the same scale as H_i in the agent-based model (recall that all values in the model are scaled relative to the arbitrary constant \widehat{M} and the parameter s_2). However, it will be proportional to H_i , and thus can be arbitrarily re-scaled to the values used in the model. In this case, since we are interested in estimating the underlying mean of q_i , we will want to scale estimates of H_i to the unit interval $[0,1]$.

Readers interested in the underlying theory should consult Glasgow and Alvarez (2009) for a concise explanation of the latent-variable interpretation of logistic regression. In Appendix E1, I provide an idealised demonstration of the application of this technique to a hypothetical sample. In this demonstration, logistic regression is shown to recover the underlying latent variable almost perfectly (see below for discussion of the question of error when employing the method in a real population).

Necessary Assumptions

A range of assumptions are required to employ this approach to estimate H_i for participants in the 2005 Senegal DHS survey. These include the notions that:

- There are strong regularities in the relations between the actors' beliefs (and other characteristics) and their perception of the intrinsic value of FGM
- The characteristics of the community-level samples of the DHS survey, involving interviews of all women within a random sample of households, support estimates for the local community as a whole.
- Estimates are not significantly undermined by survey non-response
- The variables assumed to underly actors' perception of the intrinsic value of FGM (e.g. beliefs about the practice) are actually related to this perception.

A detailed discussion of these assumptions is included in Appendix E2. Each is defensible to some extent. However, my intention is *not* to dismiss out-of-hand the possibility that these assumptions are violated, or that this violation will undermine the analysis. However, these assumptions are sufficiently credible that the analysis is worth undertaking – and there is a real possibility that it will enhance the usefulness of the ABM. Moreover, the success of the empirical analysis is *itself* tested by the independent validation of the model. As I show later in this chapter, use of the parameter estimates in this chapter is sufficient to create a strong positive correlation between the predictions of the ABM model, and observed levels of FGM in communities in Senegal (as well as allowing the model to reproduce two detailed macro-level empirical patterns). It is also worth noting, for readers sceptical of the notion of 'latent' variables, that we can arrive at a very similar estimation strategy through a pragmatic approach to 'ranking' survey respondents in terms of their level of support for FGM (see Appendix E3 for a detailed argument).

Anticipating and Quantifying Uncertainty in Estimates of μ_H and σ_H^2

While estimation of μ_H and σ_H^2 at the community level from the 2005 Senegal dataset may be a viable proposition, it is clearly *not* the case that these estimates can be made without uncertainty. As part of the calibration procedure, I tried to anticipate the joint contribution of two kinds of error that might arise in the estimation process, and to incorporate these expected errors into the generation of a 'calibrated' parameter set (see below). These two sources of error of key interest were *model-based errors* due to omitted variables in the logistic regression analysis, and *error arising from data-collection* (primarily sampling errors).

The discussion above of the use of logistic regression to estimate a latent variable is somewhat idealised. It assumes, for example, that all meaningful determinants of that variable are measured and included in the analysis (this is the case for the demonstration in

Appendix E1, where the R^2 of the model is close to 1). Yet this is unlikely to be the case using real data (and indeed the estimated model, below, does not explain the data perfectly). This introduces the potential for (perhaps substantive) errors in the model-based estimates of individual H_i scores.

Also, in the idealised use case for latent variable estimation, we have a large simple random sample from a relatively homogenous population. However, the real scenario is one in which the available data is in a set of intra-correlated clusters, with potentially very different average perceptions of FGM (e.g. in FGM practicing and non-practicing communities), with the model to be estimated from the collection of samples across all of these clusters, and estimates of μ_H then to be disaggregated by taking the mean of the estimated H_i scores at the community level.

It might be straight forward to quantify the estimation errors associated with either of these issues *considered in isolation*. If H_i were measured directly, for instance, then the expected degree of sampling error associated with the community-level averages (μ_H) could be calculated using standard formulas (e.g. standard-error of the mean).

However, anticipating the degree of error arising from the *combination* of model uncertainty and the structure of the available data, is much more challenging – I am not aware of any straightforward technique to calculate it analytically. To develop an approximate estimate of the error (formally, Root Mean Squared Error, which is the combination of variability and bias in the estimator) associated with community-level estimates of μ_H and σ_H^2 , I employed a Monte Carlo (MC) simulation approach.

The MC simulation involved stochastic generation of populations representing the 2005 Senegalese DHS sampling frame, with characteristics mimicking those of the Senegalese population (e.g. clustering of correlated ‘beliefs’ about FGM within communities), with sampling from localised clusters, and using logit regression with (varied numbers of) omitted variables. The full details of the simulation are given in Appendix E4. The results suggest that the RMSE of the estimator for community-level μ_H and σ_H^2 remains low enough to be useful, even under conditions similar to those in the real data, and that this expected error is related to the generalised R^2 of the model. Given the R^2 of the model used in the real empirical calibration (below), the MC simulation suggests an expected RMSE of around 0.059 (Table E4.1) for community-level estimates of μ_H and 0.0045 (Table E4.3) for community-level estimates of σ_H^2 , after H_i is scaled to the unit interval.

Strategy for Quantitative Estimation of s_2

The s_2 parameter of the model represents the ratio of the maximum perceived intrinsic benefit for practicing FGM to the maximum perceived intrinsic cost of the practice. Formally, the model represents H_i as being in the interval $[-\hat{M}, \hat{M} \cdot s_2]$, where \hat{M} is an arbitrary constant relative to which incentives in the model are scaled. In executing the simulation, \hat{M} is usually (arbitrarily) given a value of 10, such that the H_i attribute of all actors will be in the interval $[-10, 10 \cdot s_2]$. The role of the s_2 parameter in the model is to create flexibility with respect to whether the maximum value that actors attribute to FGM is as great as the maximum cost that others attribute to it. As discussed in Chapter 5, modellers have typically assumed that s_2 is 0, such that they assume that actors only treat FGM as an intrinsic cost, with incentives to engage in the practice purely social. Yet, as evident in empirical analysis (e.g. see Chapter 4) actors may have powerful reasons for viewing FGM as *intrinsically* valuable, including because they view it as a religious requirement.

Estimation from ABM Simulated Data

At first glance, it might appear to be straight forward to estimate S_2 from a cross-section of a given run of the ABM simulation. After all, for a given community, one could examine the ratio of the absolute value of the maximum and minimum value of H_i in a simulated community. Where s_2 is equal to 0.5, for instance, we might expect the maximum H_i value to be ' x ', and the minimum to be ' $-2x$ '. However, $\hat{M} \cdot s_2$ is a *theoretical* maximum, it only occurs as a value of H_i in the ABM simulated population in the event that the 'extreme' upper end of the distribution ($q_i = 1$) appears in a given run of the ABM simulation. If the variance of the distribution of q_i (and consequently H_i) is too low (such as if everyone in the community dislikes FGM), $\hat{M} \cdot s_2$ will not necessarily be observed in a given community.

However, in the event that we have access to *multiple* cross-sections from multiple ABM simulation runs, we can potentially estimate s_2 more reliably. By using the maximum and minimum value of H_i across multiple simulations with the same S_2 parameter, we will correctly estimate the S_2 parameter as long as H_i values of $-\hat{M}$ and $\hat{M} \cdot s_2$ have occurred in at least one of the ABM simulations.

As such, under ABM simulated data, the estimator for S_2 is defined as follows, where $\max[\cdot]$ and $\min[\cdot]$ operate across multiple cross-sections of ABM simulations, with a shared s_2 parameter.

$$\hat{s}_2 = \frac{\max[H_i]}{|\min[H_i]|}$$

Analogous Empirical Measures

The natural translation of this estimation procedure into an empirical estimator is to use the H_i estimates from the available survey sample, and take the ratio between the maximum and (absolute value of) the minimum H_i in the whole surveyed population. In this sense, the task of estimating s_2 is just an extension of the task of estimating H_i (discussed above).

Necessary Assumptions

The key assumption necessary to estimate s_2 in the manner described above, is that s_2 is homogeneous across the different communities in the 2005 DHS survey. Abstractly, this is the assumption that the maximum ‘possible’ perceived intrinsic value of FGM is the same in all communities – irrespective of whether that perception is present in a given community.

At first glance, this assumption seems poorly defined. However, we can make it much more concrete, and in so doing we find that it is *entailed by* assumptions that we have already made. We have operationalised H_i as a latent variable determined by a number of ‘determinants’, primarily, beliefs about FGM. Given this interpretation, the concept of the ‘maximum possible’ intrinsic value that an individual could attribute to FGM, can be defined in terms of the pro-FGM beliefs that they could hold, the anti-FGM beliefs they could fail to hold, and so on. Moreover, since we have assumed that the relation between these beliefs and the perceived value of FGM in the population is homogeneous (i.e. the beta-coefficients are homogeneous) this ‘maximum value’ is *definitionally* the same for all actors. Since we are primarily interested in the maximum perceived intrinsic value that *can* actually occur, and assuming that the DHS survey captures the full variation in possible perceptions of the value of FGM, the maximum (and minimum) possible perceived value of the FGM is properly defined as the highest (and lowest) H_i score among survey participants.

Anticipating and Quantifying Uncertainty in Estimates of S_2

Estimates of the uncertainty in estimating S_2 were calculated as part of the Monte Carlo experiment described in Appendix E4 (and above). The expected RMSE for the estimator is also understood relative to the generalised R^2 of the model used in estimating H_i . At an R^2 of 69%, the expected RMSE of the estimator is approximately 0.2 (Table E4.2). This

expected error of the estimate of S_2 is incorporated into the generation of a calibrated parameter set (see below).

Strategy for Qualitative Estimation of ρ_{Hy}

The ρ_{Hy} parameter of the ABM ($\in [0,1]$) can be understood as the strength of the positive relationship between the H_i attribute of agents in the ABM and their position in the y dimension of network space. When ρ_{Hy} is 1, there is a perfect rank-correlation between actor's y -coordination position in network space and their H_i attribute. When it is 0, there is no relationship between this coordinate and H_i .

In the ABM, network space is a 2D space used as part of the social network generation algorithm of the model. Actors' co-location in the network space determines whether they fall into one another's social reference groups (see Chapters 5 and 6). The function of ρ_{Hy} is to introduce *homophily of preferences* into these networks, such that, on average, actors who are socially connected tend to have more similar views about the intrinsic value of FGM, relative to actors who are not connected.

Estimation from ABM Simulated Data

Estimation of ρ_{Hy} from a cross-section of a simulated population is not straight forward. It is an abstract property of the underlying network generating algorithm of the model, rather than something that can be read from the properties of the agents in the simulation. However, since the only function of ρ_{Hy} is to introduce homophily of preferences into the social network of the population (and this doesn't arise otherwise) we can infer that ρ_{Hy} is positive if there is homophily of the social network(s) of the simulated population.

Analogous Empirical Measures

Inferring empirically whether ρ_{Hy} is meaningfully greater than zero is complicated by the fact that the social network of actors, including the relationships between surveyed households within individual communities in the survey, is not measured directly. It is also important to note that, in this context, we are referring to homophily of preferences *within* the social network of individual communities, *not* to homophily of preferences with respect to community membership (which is certainly present).

However, we may be able to use proxy measures of social group membership within communities as an indicator of whether two households are likely to be connected (or at least whether they are likely to be part of a shared network 'cluster' within the community).

In Senegal, shared ethnicity is considered to be a key vector of social influence, including with respect to FGM (UNICEF, 2013). We may, therefore, expect (at least typically) those survey participants from the same residential community and the same ethnic group, will be more closely connected in the social network of the community than participants from the same residential community and different ethnic groups.

This suggests the following as a test for the presence of homophily of preferences with respect to FGM. If homophily of preferences is present, we should expect to find that actors' preferences with respect to FGM (measured by H_i) are more strongly correlated with other actors who are in the same ethnic group *and* community, than with other actors from a *different* ethnic group (but the same community).

Anticipating and Quantifying Uncertainty in Estimates

Calibration of ρ_{Hy} is primarily qualitative, in the sense that evidence of homophily of preferences will be found, or it won't. There is no clear way to map this back to a particular quantity in the parameter, or to calculate error-bounds for that parameter. Instead, I adopted a simple qualitative scheme for calibration. If there is no clear evidence of homophily of preferences I map this to 'low' values of ρ_{Hy} (between 0 and 0.2), if there is clear evidence of homophily of preferences I map this to 'high' values of ρ_{Hy} (between 0.2 and 1).

Empirical Calibration

Modelling H_i/q_i as a Latent Variable

H_i was estimated for adult women who responded to the 2005 Senegalese DHS survey (n=14,602). Of these women 1,984 (in total) were excluded from the analysis because they had missing or uninterpretable data on the dependent or independent variables (see discussion of estimation assumptions in Appendix E2). H_i was modelled as a latent variable, and taken to be the linear predictor of a logistic regression model. The dependent variable was a binary indicator of whether the respondent stated that FGM should continue (1) or be stopped (0). (respondents that answered ‘depends’ or ‘don’t know’ were excluded). The independent variables were twelve nominal variables related to participants stated beliefs about FGM, one nominal variable corresponding to the wealth quartile of participants’ household, and one interval variable measuring the number of years of education of the participant (this variable was entered as a linear term *and* a second-order exponential term).

Details of the model fit, which had a generalised R^2 metric of 69% (Zhang, 2017), are given in Table 10.

Table 10: Logistic Regression of Support for FGM on Beliefs, Education and Wealth (Senegal 2005 DHS)

Latent Variable Model of H_i/q_i using Logistic Regression	
Independent Variable	Dependent Variable: FGM should continue (ref: should be stopped) (Logit Beta Coefficient)
Advantage FGM: Improves Women's Hygiene (Ref: Not Stated)	2.006***
Advantage FGM: Improves Social Acceptance (Ref: Not Stated)	0.700***
Advantage FGM: Improves Sexual Pleasure for Men (Ref: Not Stated)	0.939*
Advantage FGM: is a Religious Necessity (Ref: Not Stated)	1.073***
Advantage <i>not</i> Practicing FGM: Avoids health problems (Ref: Not Stated)	-1.176***
Advantage <i>not</i> Practicing FGM: Avoids suffering (Ref: Not Stated)	-0.342**
Advantage <i>not</i> Practicing FGM: Improved Sexual Pleasure for Women (Ref: Not Stated)	-0.329
Advantage <i>not</i> Practicing FGM: Improved Sexual Pleasure for Men (Ref: Not Stated)	-0.276
Advantage <i>not</i> Practicing FGM: (Non-Practice) Accords with Religion (Ref: Not Stated)	-0.969**
Years of Education	0.003
Years of Education (Squared Term)	-0.002
Wealth Quintile: Poorer (Ref: Poorest)	0.198
Wealth Quintile: Middle (Ref: Poorest)	-0.110
Wealth Quintile: Richer (Ref: Poorest)	-0.402**
Wealth Quintile: Richest (Ref: Poorest)	-0.895***
Circumcision reduces premarital sex? Yes (Ref: No)	0.206*
Circumcision reduces premarital sex? Don't Know (Ref: No)	-0.334***
Circumcision required by religion? Yes (Ref: No)	1.742***
Circumcision required by religion? Don't Know (Ref: No)	0.560***
Men want circumcision to continue? No (discontinue – Ref: Yes, continue)	-4.768***
Men want circumcision to continue? Depends (Ref: Yes, continue)	-2.681***
Men want circumcision to continue? Don't Know (Ref: Yes, continue)	-2.647***
Constant	1.165***
N	12,618
Log Likelihood	-2,539.514
Generalised R^2 (Zhang, 2017)	0.6931

*p < .05; **p < .01; ***p < .001

The values of the linear predictor for the model (the log-odds scores) were scaled to the unit-interval ([0,1]) as an estimate of q_i for each respondent. Figure 40 illustrates the range of q_i values estimated for survey participants across all communities.

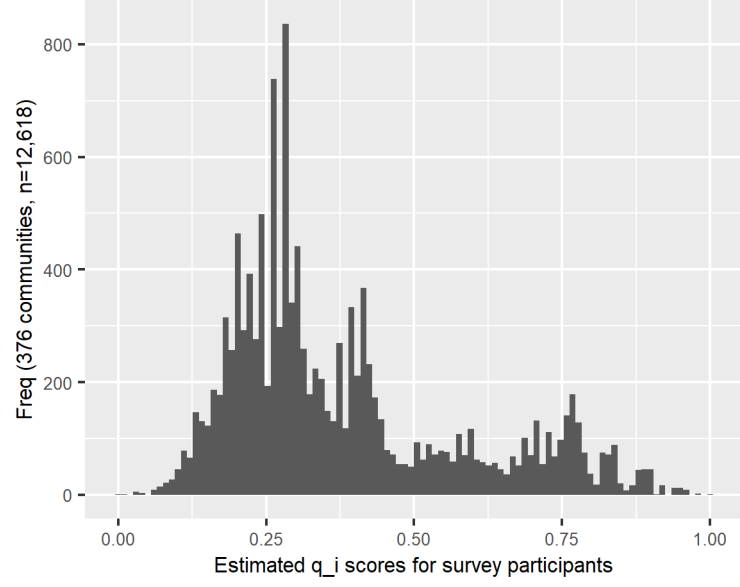


Figure 40: q_i Estimates Across Participants in the 2005 Senegalese DHS Survey

Estimating μ_H and σ_H^2 at the Community Level

μ_{H_j} , where j indexes the 376 communities in the Senegal DHS survey, was estimated as the arithmetic mean of q_i in community j . $\sigma_{H_j}^2$ was estimated as the sample variance of q_i in community j . The estimated empirical marginal distributions for μ_{H_j} and $\sigma_{H_j}^2$ are given in Figures 41 and 42 (respectively).

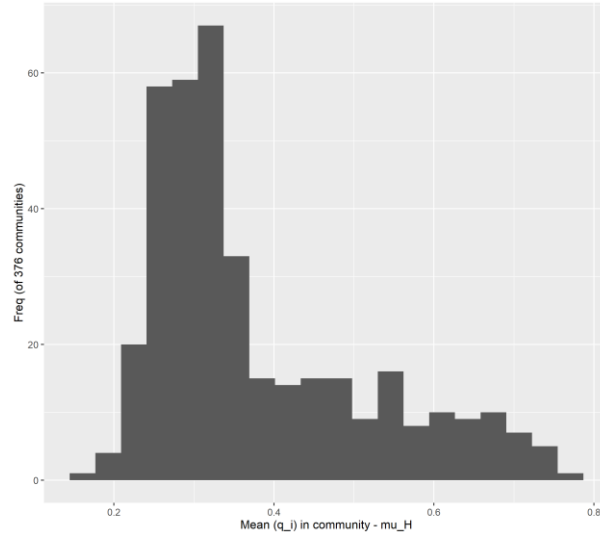


Figure 41: Marginal Empirical Distribution of μ_H Parameter

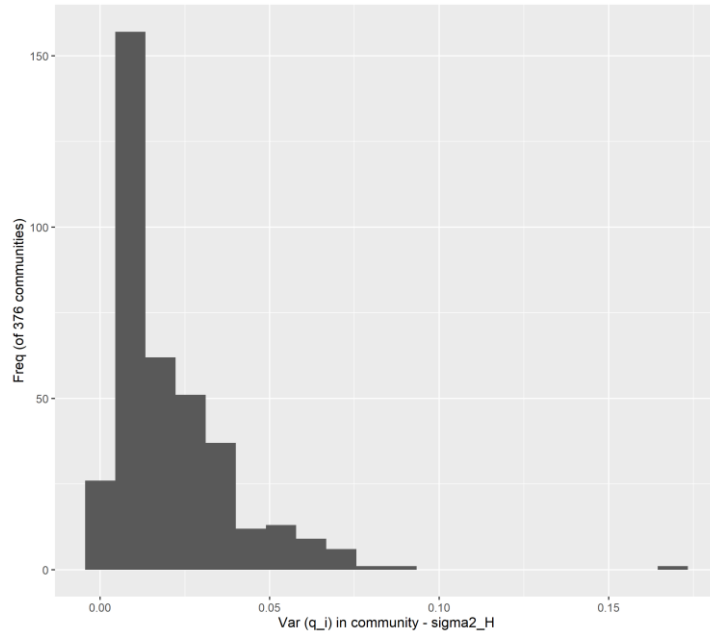


Figure 42: Marginal Empirical Distribution of σ_H^2 Parameter

Of particular interest is the observation that μ_H is generally bounded in an interval between 0.2 and 0.8 (which corresponds to the arbitrary bounds used in the uncalibrated model), while σ_H^2 rarely exceeds 0.075 (in the uncalibrated model this value can be as high as 0.111).

Figure 43 shows the joint relationship between μ_H and σ_H^2 . This is of interest from the perspective of calibration, because together these parameters control the shape the distribution of preferences in the ABM simulated population. Communities are coloured according to tertiles (three equal-sized groups) of the proportion of women who have been cut in that community. This is important in Chapter 8, where I restrict the parameters used for policy-analysis to the range of values of μ_H and σ_H^2 that are observed for communities where FGM is highly prevalent (i.e. the blue points), since these are the communities likely to be the target of intervention efforts.

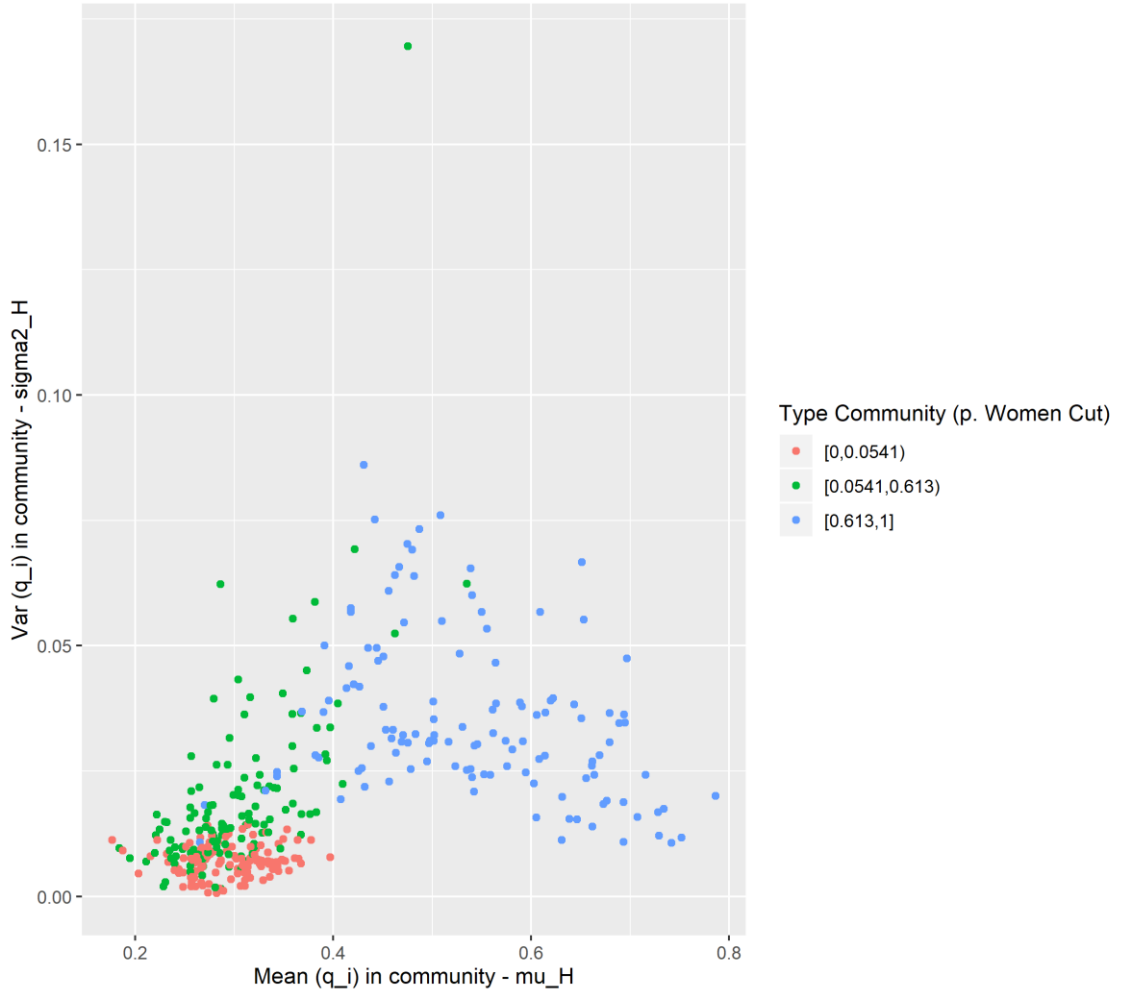


Figure 43: Empirical Relationship Between μ_H and σ_H^2 in Communities in the 2005 Senegalese DHS Survey

Estimating s_2 at the Community Level

Following the strategy discussed earlier in the thesis, s_2 was estimated as the maximum linear predictor of the fitted logit scores for survey participants, divided by the absolute value of the minimum fitted logit score. This yielded an estimated s_2 of 0.944 for the surveyed population.

Assessing Evidence of Homophily of Preferences (ρ_{Hy})

To assess evidence of homophily of preferences in FGM practicing communities, two measures were calculated for all survey participants. The first measure was the average (estimated) q_i value of *other* actors in the same community, and of the same ethnic group. The second measure was the average q_i value of other actors who were in the same community, but of a *different* ethnic group. The test for homophily of preferences was whether the correlation between self and (average) other q_i in the community was stronger when ‘other’ was defined as others in the same ethnic group, than when it was defined as

others in a different ethnic group. Scatter-plots for both measures are shown in Figures 44 and 45.

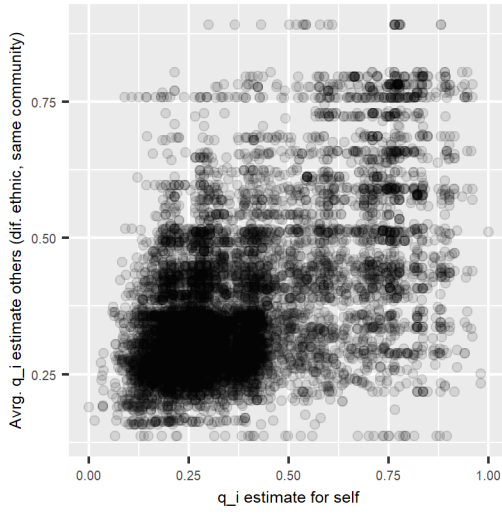


Figure 44: Joint Relation: Avg. q_i of Others with a Different Ethnicity in Same Community by q_i of Self (1272 cases removed due to non-availability of ‘others’ from different ethnic group. Points are shown with 10% opacity.)

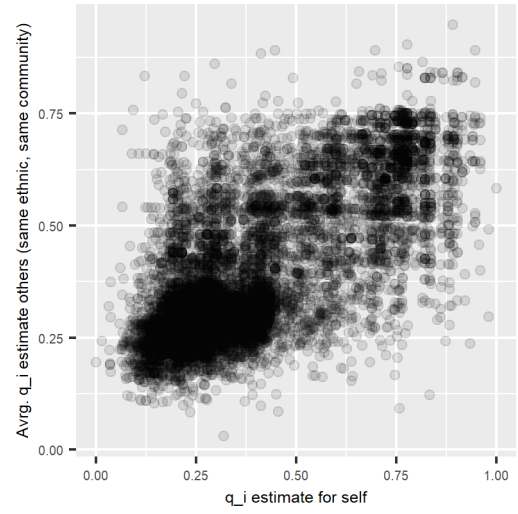


Figure 45: Joint Relation: Avg. q_i of Others with the Same Ethnicity in Same Community, by q_i of Self (350 cases removed due to non-availability of ‘others’ from same ethnic group. Points are shown with 10% opacity.)

Visual inspection of both figures appears to show a closer (i.e. tighter) linear relation between the q_i attributes of actors, and those of others in the community who share their ethnicity (relative to those in the community who don’t share their ethnicity). However, assessment is made difficult by the volume of data (which creates visual noise). The stronger association in the case of ‘same’ ethnic group, is confirmed by correlational analysis, which shows a Pearson’s r correlation coefficient of 0.701 (95% CIs [0.692, 0.710]), versus 0.543 (95% CIs [0.530, 0.557]) for the ‘other’ ethnic group association. Accordingly, in the calibrated parameter space, ρ_{Hy} is assumed to have a value of 0.2, or above (see discussion above).

Generating a Calibrated Parameter Space

Based on estimates for μ_H and σ_H^2 at the community level, and estimates for s_2 and ρ_{Hy} at the population level, I generated a ‘calibrated’ parameter space for the ABM. The key aim for this parameter space was that it would (a) capture the marginal and joint distributions for μ_H and σ_H^2 at the community level (i.e. per-ABM simulation), (b) control s_2 and ρ_{Hy} (across all runs of the simulation) and (c) take account of uncertainty in each of the estimates.

In order to do this, I adopted a Monte Carlo approach, in which I generated a set of 1000 credible parameterizations of the ABM for *each* of the communities measured in the survey, and combined these together into an overall calibrated parameter ‘space’ for the ABM of

376,000 empirically supported sets of parameter values. This parameter space then represented the set of possible inputs (on the calibrated variables) for the simulation, with other possible combinations of inputs excluded as lacking empirical support. By associating each community with an equal number of parameterizations, it was also possible for inputs in the parameter space to be sampled randomly, whilst maintaining (on average) the marginal and joint distributions of the characteristics of the surveyed communities.

In order to generate a set of 1000 credible parameter-sets for each community $j \in \{1, 2, \dots, 375, 376\}$. I defined a data generating mechanism for each parameter $\{s_2, \rho_{HY}, \mu_H, \sigma_H^2\}$ involving a random variable centred on the empirical estimate of that parameter and with a standard deviation equal to the expected RMSE for that parameter (see Appendix E4, Tables E4.1-3). I then took 1000 samples from each data generating mechanism (for each community). Formal definitions are as follows:

The set of credible parameter values (indexed $i \in \{1, 2, \dots, 1000\}$) for community j on parameter s_2 was defined as:

$$\{s_{2_{1j}}, s_{2_{2j}}, \dots, s_{2_{999j}}, s_{2_{1000j}}\}, s_{2_{ik}} \sim \bar{N}(\widehat{S_{2j}}, RMSE_{s_2}, 0, 1)$$

Where $s_{2_{ij}}$ are independent random variables with a truncated normal distribution (\bar{N}) of mean $\widehat{S_{2j}}$ (the empirical estimate for community j) and standard-deviation $RMSE_{s_2}$ (the expected average error for the estimator), with lower limit 0 and upper limit 1.

The $\mu_{H_{ij}}$ and $\sigma_{H_{ij}}^2$ parameter sets for community j were similarly defined as follows:

$$\{\mu_{H_{1j}}, \mu_{H_{2j}}, \dots, \mu_{H_{999j}}, \mu_{H_{1000j}}\}, \mu_{H_{ij}} \sim \bar{N}(\widehat{\mu_{Hj}}, RMSE_{\mu_H}, 0, 1)$$

$$\{\sigma_{H_{1j}}^2, \sigma_{H_{2j}}^2, \dots, \sigma_{H_{999j}}^2, \sigma_{H_{1000j}}^2\}, \sigma_{H_{ij}}^2 \sim \bar{N}(\widehat{\sigma_{Hj}^2}, RMSE_{\sigma_2}, 0, 0.25)$$

The ρ_{Hy_i} parameters set had the same definition for all communities:

$$\{\rho_{Hy_1}, \rho_{Hy_2}, \dots, \rho_{Hy_{999}}, \rho_{Hy_{1000}}\}, \rho_{Hy_i} \sim UNIF(0.2, 1)$$

Where $UNIF(0.2, 1)$ is the continuous uniform distribution in the interval $[0.2, 1]$. The empirically supported parameter sets for each community (C_j) were then defined as rows of the 1000 x 4 matrix:

$$C_j = \begin{bmatrix} s_{2_{1j}} & \mu_{H_{1j}} & \sigma_{H_{1j}}^2 & \rho_{Hy_1} \\ \dots & \dots & \dots & \dots \\ s_{2_{1000j}} & \mu_{H_{1000j}} & \sigma_{H_{1000j}}^2 & \rho_{Hy_{1000}} \end{bmatrix}$$

Empirical Validation

Having defined a set of independent empirical calibration strategies and executed them in order to generate a calibrated parameter space for the ABM, we now turn to the question of empirical validation. As noted above, empirical validation provides a check against the adequacy of calibration efforts, as well as the overall adequacy of the model. With respect to the former, the primary test is whether or not calibration induces a positive association between the social dynamics produced by the ABM and those observed in the real communities from which parameter sets were selected. With respect to the latter, the primary test was whether or not the model could reproduce key empirical patterns observed in the empirical data *under independently calibrated parameter settings*.

For both tests, a procedure needed to be defined which allowed for a comparison between characteristics of the ABM simulation and the observed communities. The procedure I adopted was to repeatedly run the ABM simulation using the calibrated parameter space as ‘input’ and to compare the rate of FGM practice simulated by the ABM (the simulated output) to the observed rate of FGM in the community from which the ‘input’ was taken. Further details about the empirical comparison are given below. First, I outline the procedure used to run the ABM simulations.

A set of parameters were chosen for the ABM (including an initial rate of FGM practice), then the simulation was run for between 12 and 120 time-steps (depending on the random η_2 parameter defined in Chapter 6 and corresponding to ‘10 years’ of real-time). Subsequently, the rate of FGM practice in the population (proportion of actors participating) was measured, along with the proportion of actors who preferred FGM in the population (i.e. proportion for whom $H_i \geq 0$).

Unlike in the global sensitivity analysis (Chapter 6), not all of the parameter values were selected uniformly at random within their defined intervals. Instead, a parameter set $(s_2, \mu_H, \sigma_H^2, \rho_{Hy})$ was chosen (with equal probability and without replacement) from the calibrated parameter space (376,000 empirically supported parameter sets), for each

simulation run (all other parameters were still randomised as in the sensitivity analysis of Chapter 6).

For this set of simulations, the *initialised* rate of FGM practice (i.e. rate of FGM prior to beginning the simulation) was not straightforward to define. For the purposes of exploring simulated interventions in Chapter 6, the initial rate of FGM was always set to 1 (all actors initially practice FGM). However, here I was interested in the capacity of the ABM to reproduce rates of FGM observed in real communities. Real communities may not have arrived at a rate of FGM from a ‘historical’ rate that was 1. Some communities may not have practiced FGM in the past, or rates of practice may have remained at a stable interior rate for a long time.

To get a better intuition for the implications of this issue for the assessment of the predictions of the ABM, consider the following heuristic characterisation of the model. We can think of the model as an ‘equilibrium finding’ machine, in the sense that, given an initial rate of FGM practice in a local population, as well as range of other information (including the distribution of preferences) about that local population, the ABM often (although not always) arrives at a ‘nearby’ equilibrium, representing a stable rate of FGM within the population (including everyone practicing FGM or no-one practicing).

This is only a heuristic, but it helps to clarify the problem. Even if the model was extremely accurate, if it was initialised with a rate of FGM practice, at, say 5%, but with population characteristics that otherwise correspond to an observed community with a rate of FGM practice at 100%, it might ‘correctly’ predict that this population *from that initial point* would not arrive at a rate of FGM of 100%. Yet clearly this prediction would not match the observed rate in the community (whose history potentially never involved a rate of FGM of 5%).

As such, running the ABM simulation with a randomised initial rate of FGM would represent a highly conservative test of its ability to reproduce observed rates of the practice. On the other hand, starting the model with an initial rate equal to the observed rate would rather nullify the validation test (since a perfect model in this scenario would be one in which nothing happened).

In running the ABM simulations, I employed two alternative assumptions about the initial ‘historical’ rate of FGM. The first set of simulations used a completely randomised initial rate of FGM, therefore providing a highly ‘conservative’ test of calibration and of the model.

The second set of simulations used the *popularity* of FGM (defined as the proportion of women who say it should continue) in the community from which the parameter set was taken, as the initial rate of FGM in the simulation. We would expect popularity to be correlated with the rate of FGM in the community in the past, however (as noted below) the *popularity* of FGM at the community level does *not* mimic the key empirical patterns of interest in the *rate* of FGM within communities – thus allowing these to ‘emerge’ from the ABM simulation.

Summary of Model Test Procedure

For each initialisation condition (randomised initialisation and popularity as initialisation, see above) the simulation was run 2000 times. For both sets of simulations, a random selection of 2000 parameter sets was chosen (chosen from the calibrated parameter space of 367,000 parameter sets) and used to set the value of the s_2 , ρ_{Hy} , μ_H and σ_H^2 parameters in the 2000 simulated runs. In the case of the ‘popularity as initialisation’ condition, the initial rate of FGM participation in the population was also set equal to the popularity of FGM in the community from which the parameter-set was taken. The model was then run for between 12 and 120 time-steps (with no simulated intervention). Subsequently the popularity and practice-rates of FGM in the ABM simulations were measured and compared to those observed in the communities from which the parameter set was taken. The same 2000 parameter sets were used for both initialisation conditions (to enhance comparability).

Does Calibration Induce Positive Associations between Simulated Rates of FGM and Observed Rates in Real Communities?

The key test of the success of empirical calibration was its capacity to induce a positive relationship between the predictions of the ABM about the rate of practice for particular communities (based on parameter sets derived from those communities) and the real rates of FGM observed in those communities.

There are two available measures which may be reasonable proxies of the level of FGM practice within communities in the 2005 Senegalese DHS survey. The first is the proportion of mothers who say that their children have been cut (or that they intended to cut them). The second is the proportion of adult women who have themselves been cut.

The first has the advantage that it is a relatively contemporary measure of FGM activity in the community (i.e. cutting happened within the lifetime of the child). However, it has the disadvantage that it may be subject to under-reporting bias, given that FGM is illegal in Senegal. By contrast, studies have found that there is little under-reporting of adult women's own FGM status (Mackie, 2017). However, this measure has the disadvantage that it is historical (women may have been cut years earlier - Yoder et al., 2004), and some women may have joined the community after being cut. Since neither measure is ideal, I use both and report all comparisons with the ABM simulation using both measures.

Without using the calibrated parameter set (which introduces information about real communities into the simulation), there *can be no* meaningful correlation between the predictions of the ABM and observations in practicing communities. If calibration failed entirely (the measures were meaningless), or if the model were entirely inadequate, we would *still* expect there to be no meaningful correlation after using community parameters as input data. Correlation is, therefore, the primary test of independent calibration.

Figures 46-49 show scatter plots of the relationship between the rate of FGM observed in the ABM simulation under randomised initialisation (Figures 46 and 47) and popularity-based initialisation (Figures 48 and 49), and the observed rate of FGM in the real communities from which simulation parameters were taken (based on either the % of women cut, Figures 46 and 48, or the % of mothers with cut daughters, Figures 47 and 49). All figures are shown with Loess curves fitted to the relation between the predicted and observed data.

Table 11 reports performance measures based on Pearson's r and Spearman's ρ metrics of correlation. It includes 95% confidence intervals for each measure. It also includes a 'null model' estimate, which represents the correlation (with 95% confidence intervals) expected if there were no relationship between the predictions of the model and observed rates in practicing communities.



Figure 46: Bivariate Relation, Simulated Rates FGM vs Observed % Women Cut (by community) (ABM w. Random Initialisation)

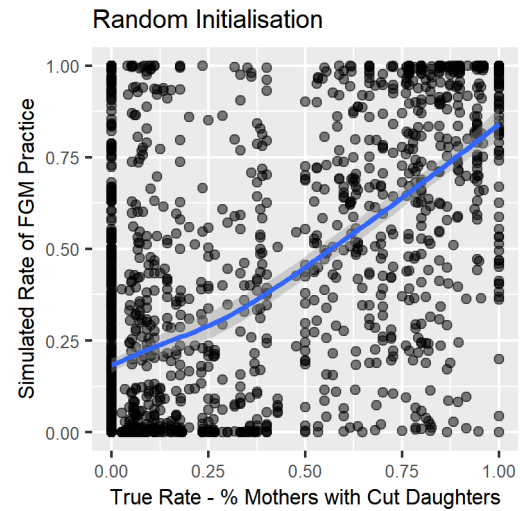


Figure 47: Bivariate Relation, Simulated Rates FGM vs Observed % Mothers with cut daughters (by community) (ABM w. Random Initialisation)

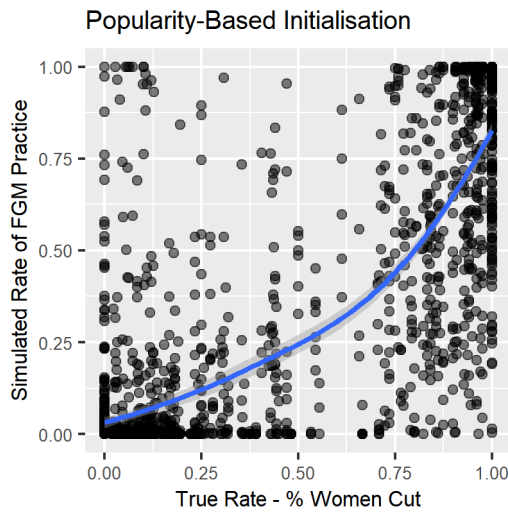


Figure 48: Bivariate Relation, Simulated Rates FGM vs Observed % Women Cut (by community) (ABM w. Popularity-Based Initialisation)

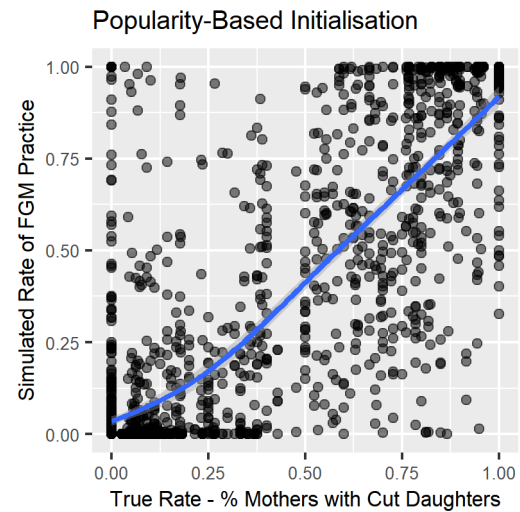


Figure 49: Bivariate Relation, Simulated Rates FGM vs Observed % Mothers with cut daughters (by community) (ABM w. Popularity-Based Initialisation)

Table 11: Validation Performance of the ABM (Correlation Metrics)

Test / Estimate Used	Null Model (expected)	ABM (with randomised initialisation)	ABM (with popularity-based initialisation)
Pearson's r (% Mothers Cutting)	$E[r] = 0, 95\% \text{ Cls}[-0.043, 0.043]$	$r = 0.602, 95\% \text{ Cls}[0.573, 0.629]$	$r = 0.838, 95\% \text{ Cls}[0.825, 0.851]$
Spearman's ρ (% Mothers Cutting)	$E[\rho] = 0, 95\% \text{ Cls}[-0.043, 0.043]$	$\rho = 0.566, 95\% \text{ Cls}[0.535, 0.595]$	$\rho = 0.783, 95\% \text{ Cls}[0.766, 0.8]$
Pearson's r (% Cut Women)	$E[r] = 0, 95\% \text{ Cls}[-0.043, 0.043]$	$r = 0.574, 95\% \text{ Cls}[0.544, 0.603]$	$r = 0.813, 95\% \text{ Cls}[0.798, 0.827]$
Spearman's ρ (% Cut Women)	$E[\rho] = 0, 95\% \text{ Cls}[-0.043, 0.043]$	$\rho = 0.538, 95\% \text{ Cls}[0.506, 0.568]$	$\rho = 0.760, 95\% \text{ Cls}[0.741, 0.778]$

The results of the correlational analysis are encouraging. Under both of the initialisation conditions, considerable ‘noise’ in the predictions of the ABM simulations is evident (more so under randomised initialisation). However, there is a clear positive relationship between the predictions of the simulation and the observed practice of FGM in communities used as parameter inputs. This relationship is substantive (approaching 0.6 correlation) even where the initial rate of FGM in the ABM simulation was randomised (so that *only* information about the distribution of preferences was input into the simulation). Where the popularity of FGM in the community was used as the initial rate of FGM in the simulation, this association rose to (approximately, see Table 11) an 80% correlation. In both cases, this positive association greatly exceeded the association to be expected if there were no underlying relation between predicted and observed rates by an order-of-magnitude (we would expect no more than $r = 0.043$ in 95% of simulation tests of this kind, given no underlying relationship).

While these results don’t confirm that the model or the calibration exercise is ‘correct’, it implies some minimal level of adequacy of both the estimation procedure and of the underlying simulation as a model of the social dynamics of FGM in real communities. This interpretation is supported by the fact that no ‘model fitting’ has occurred – any reproduction of the observed data is by virtue of the structure of the model’s design, and the independent estimation of its parameter values.

Can the Model Reproduce Key Empirical Patterns under Calibrated Parameter Settings?

While correlational analysis provides some indication that the model and associated calibration has ‘captured’ some aspects of the social dynamics of the target, the adequacy of

the model can be *further* assessed through examination of the ability of the model to reproduce distinctive *empirical patterns*, also known as *stylised facts*, in the practice of FGM in communities in Senegal (cf Chapter 3). This is sometimes called multi-criteria assessment or Pattern Orientated Modelling (Grimm et al., 2005; Railsback and Grimm, 2012). As discussed previously, multi-criteria validation has a number of conceptual advantages, chief of which is that it helps to maximise the strength of the validation test provided by limited data by distilling that data into multiple distinctive patterns, each of which provides its own test of the model. The strength of the test provided by multi-criteria validation is greatest when this done in the context of independent empirical calibration.

The key question in implementing multi-criteria validation is the selection of patterns to be used as the test. The patterns selected should have a bearing on core questions about the design of the model, or its applications (Railsback and Grimm, 2012). In an ideal situation, sufficient secondary data would be available to identify distinctive patterns in *the outcomes of anti-FGM interventions* (given the final intended application of the model to policy failure analysis, Chapter 8). However, such patterns do not appear to have been clearly identified for FGM. While, for example, while evaluations of the Tostan program in Senegal indicate real successes in discouraging FGM, they do not clearly reveal distinctive ‘regularities’ that could be used to test a model. Furthermore, to the best of my knowledge, there is also no available raw secondary data on intervention outcomes through which such patterns might be discovered post-hoc.

However, there are more general patterns in the persistence of FGM that are of clear theoretical relevance to the model, and that are visible in the 2005 Senegal survey. The first of these relates to the *marginal* distribution of rates of FGM at the community level. Readers will recall that one of the touted explanatory virtues of the original social convention model was that it predicted ‘local universality’, i.e. that FGM tends to be universal or absent within communities (Mackie, 1996, 2017). This has been considered something of a signature of the social coordination theory of FGM (although see Chapter 2 for a critical discussion of this form of argument).

Early evidence for the local universality regularity appears to have been primarily anecdotal, or based on data at the regional and ethnic-group level. However, two recent studies (including a preliminary output of this project, see Appendix G) have provided empirical evidence at the *community level* that rates of FGM tend to be concentrated near to universal abandonment or near-universal practice (Droy, 2018; Mackie, 2017). Analysing combined

community-ethnicity clusters across DHS surveys in 25 countries, Mackie argued that 93% of communities have rates of FGM practice below 10% or above 90% (based on the % of women cut) (Powell, 2017).

This pattern is visible in the rate of FGM in communities in Senegal. Figures 50 and 51 show a probability kernel density estimate (smoothed histogram) for the rate of FGM across the 376 communities measured in the 2005 DHS survey, with the rate of FGM defined as either the proportion of women cut (Figure 50), or the proportion of mothers with cut daughters (Figure 51). For both measures, we see some evidence of the polarisation of the rate of FGM, with communities less likely to exhibit intermediate rates of FGM, and more likely to exhibit rates of FGM close to universal practice or abandonment.

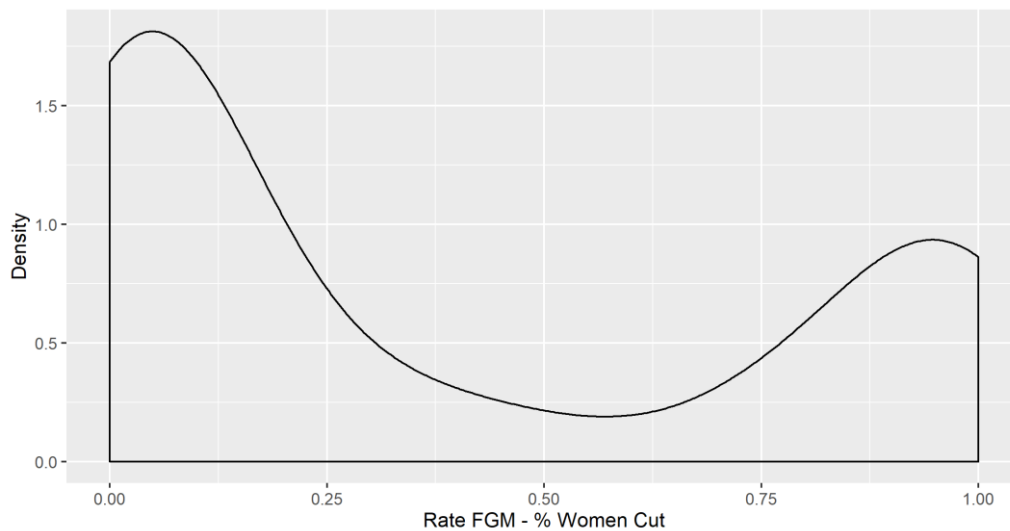


Figure 50: Marginal Distribution (Density) of % Women Cut (by Community)

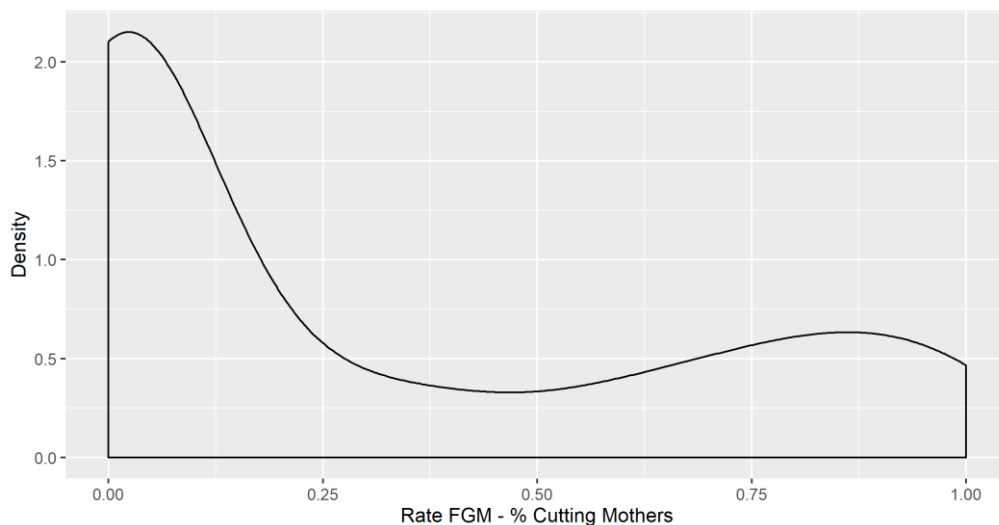


Figure 51: Marginal Distribution (Density) of % Mothers Cutting (by Community)

This pattern in the marginal distribution of rates of FGM at the community level was used as the first pattern to test the ABM simulation. The test was whether the ABM would reproduce this U-shaped curve pattern under empirically-supported parameter settings, using either randomised initialisation or popularity-based initialisation. For this test, the results of the previous simulations were used, with the marginal distribution of simulated FGM rates acting as the output of interest. Since the outputs of the simulations were based on sampling 2000 parameter inputs (each corresponding to a community) the appropriate comparison was with the marginal distribution of FGM practice among those 2000 (re)-samples of communities (since this distribution might fluctuate slightly from the empirical picture above).

Results for the simulations, along with the marginal distributions of the rate of FGM (using both measures) in the 2000 community re-samples to which these simulations corresponded, are shown in Figures 52-56. As reassurance that the popularity-based initialisation condition did not ‘impose’ the pattern of interest on simulated outputs, the marginal distribution of the *popularity* of FGM in the 2000 re-sampled communities is also shown (Figure 56)⁶⁸.

Although neither set of simulations exactly reproduced the marginal distribution of rates of FGM in the communities used as input, both clearly captured the key U-shaped pattern of interest. Rates of FGM in the simulations were concentrated near to universal practice or abandonment. It is particularly encouraging that the simulations reproduced this pattern under randomised initialisation conditions.

⁶⁸ Readers will observe that this does not show the distinctive concentration of rates of FGM near to universal practice.

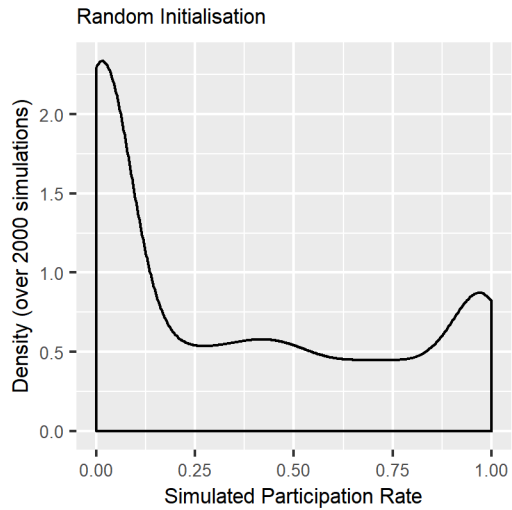


Figure 52: Simulated Marginal Rate FGM (ABM w. Random Initialisation)

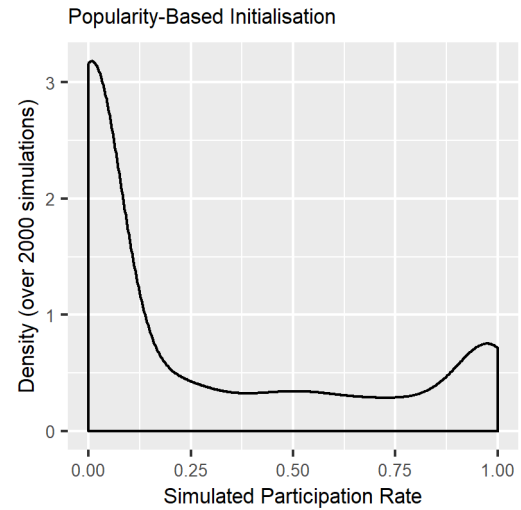


Figure 53: Simulated Marginal Rate FGM (ABM w. Popularity-Based Initialisation)

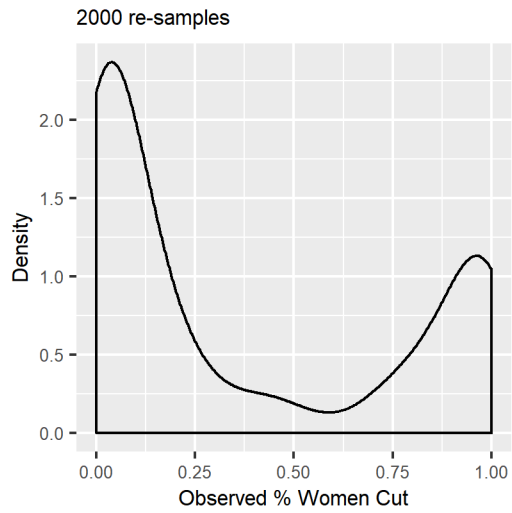


Figure 54: Empirical Marginal Distribution % Women Cut (2000 re-samples from 376 communities)

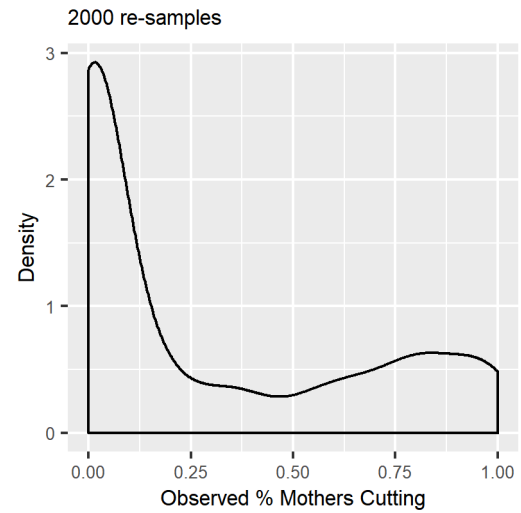


Figure 55: Empirical Marginal Distribution % Mother's Cutting (2000 re-samples from 376 communities)

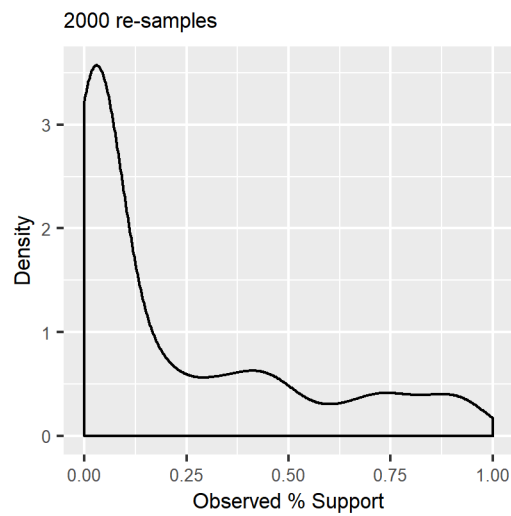


Figure 56: Empirical Marginal Distribution % Women Supporting FGM (2000 re-samples from 376 communities)

A second empirical pattern of key theoretical interest is the bivariate relationship between the popularity and prevalence of FGM in practicing communities. A related regularity was noted in defence of the social convention model by Mackie (1996), who suggested that FGM may be practiced by those who oppose it. Similarly, in UNICEF's (2013) review of national surveys related to FGM, the authors noted that the prevalence of FGM typically exceeds its popularity at the national level. However, to my knowledge, no publication prior to Droy (2018 and related conference presentations, which were an early output of this project, see Appendix G) has examined the bivariate empirical relationship between the popularity and prevalence of FGM at the community level.

Figures 57 and 58 plot the bivariate relationship between the popularity of FGM (x-axis) and the rate of FGM (y-axis), measured as the percentage of women cut, or the percentage of mothers with cut daughters, with Loess curves fitted for both figures. For both measures, the observed pattern is quite striking. We see that in the vast majority of communities, the practice of FGM exceeds its popularity. There are also other interesting features in the trend of the relation between popularity and prevalence. In both cases, but especially in Figure 58, we see that the rate of FGM remains high while support for the practice is above 25%, declining more rapidly below this point, with instances of the rate of FGM being *below* the popularity of the practice concentrated in this lower quarter (below 25%). Although I am not aware that this pattern has been identified empirically previously (except in my own outputs, see above), it bears a resemblance to a 'hypothetical' pattern of social norm change suggested in (Mackie et al., 2015: 18)⁶⁹.

⁶⁹ Although these authors refer to social norm adoption rather than social norm abandonment.

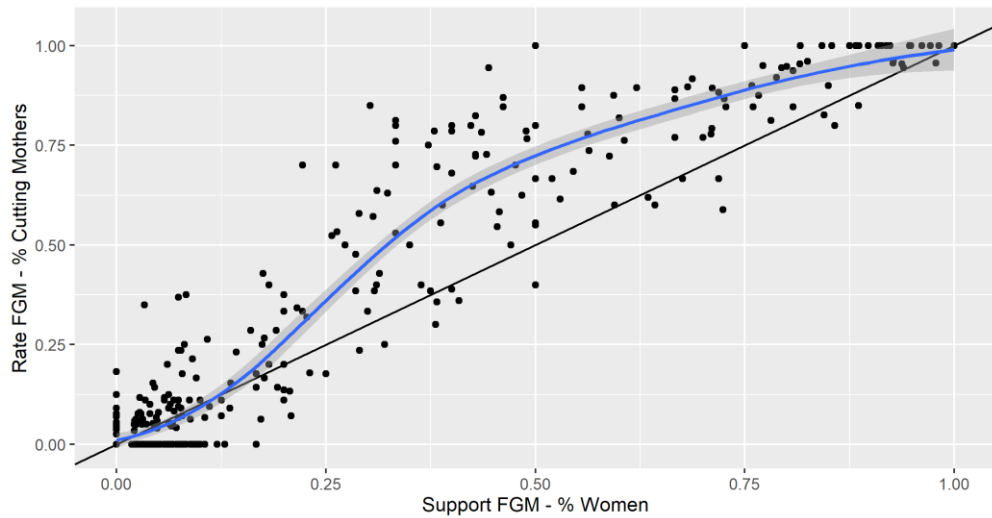


Figure 57: Empirical Bivariate Relation: % Mothers Cutting by % Women Supporting FGM

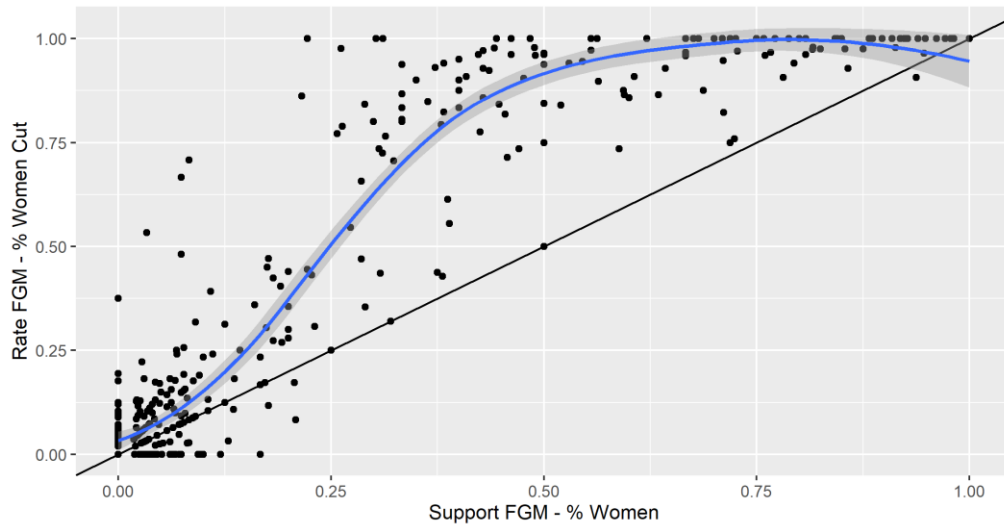


Figure 58: Empirical Bivariate Relation: % Women Cut by % Women Supporting FGM

The test of whether the ABM simulation reproduced these patterns in the relation between popularity and practice, followed the structure of the previous test (and used the same two sets of 2000 simulated runs). In Figures 59-63, the outputs of the simulations (under randomised and popularity-based initialisation) are compared to the bivariate relationship between popularity and the rate of FGM (based on the two definitions) in the 2000 (re) sampled communities used as input to the simulations. For the simulated data, the plot shows the relationship between the *simulated* rate of FGM (proportion agents participating) and the *simulated* popularity of FGM (the proportion of agents for whom $H_i \geq 0$). Figure 63 provides a hypothetical comparison in which the popularity of FGM at the community level is plotted against itself, with random uniform ‘jitter’ of 30% added to both axes. The point of this latter figure is to provide an indication of the popularity and prevalence relationship expected if the two quantities are essentially the same (but measured with significant error).

The analysis shows that this is not the case and demonstrates that the patterns of interest (with respect to the popularity and prevalence relationship) are not ‘imposed’ on the results of simulations under popularity-based initialisation.

Once again, neither set of simulations reproduced the observed pattern exactly. However, the results are clearly encouraging, especially under popularity-based initialisation. Both sets of simulations reproduce the observed pattern that rates of FGM consistently exceed the popularity of the practice. Also, in both simulations, rates of FGM *below* the popularity of the practice were concentrated in simulations in which the popularity of the practice was below 25%. However, the simulations using randomised initialisation frequently predicted high rates of FGM in communities where the popularity of the practice was well below 25%. This may be attributed to scenarios in which the simulation was run with a high initial level of FGM practice, despite low levels of popularity. By contrast, under the popularity-based initialisation, a rapid decline in the predicted rate of FGM is clearly evident, and (appropriately) begins at levels of popularity around 25%.

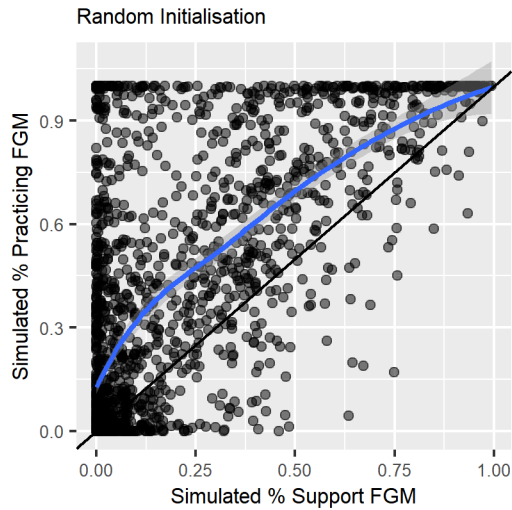


Figure 59: Simulated Bivariate Relation: % Agents Practicing FGM by % Agents Supporting FGM (Random Initialisation)

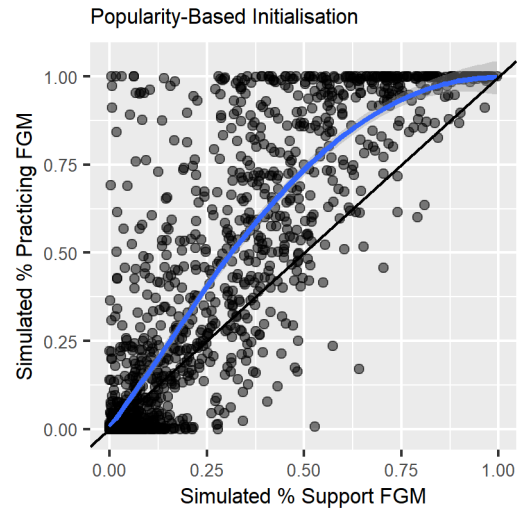


Figure 60: Simulated Bivariate Relation: % Agents Practicing FGM by % Agents Supporting FGM (Popularity-Based Initialisation)

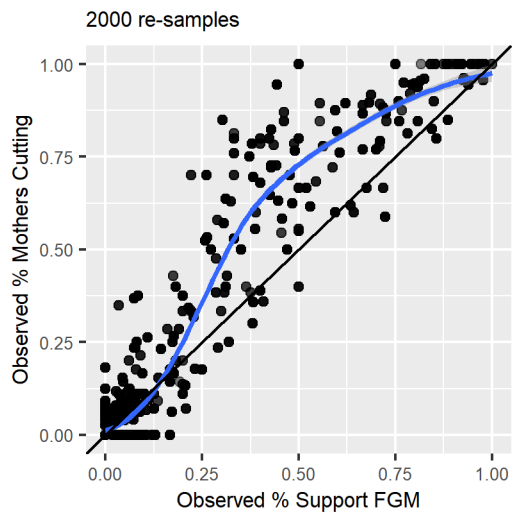


Figure 61: Empirical Bivariate Relation: % Mothers Cutting by % Women Support (based on 2000 re-samples from 376 communities)

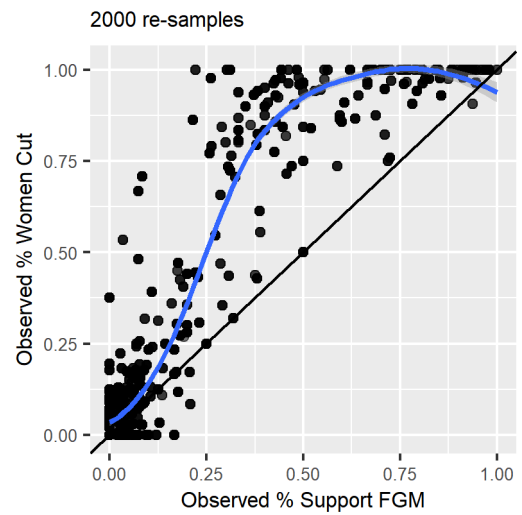


Figure 62: Empirical Bivariate Relation: % Women Cut by % Women Support (based on 2000 re-samples from 376 communities)

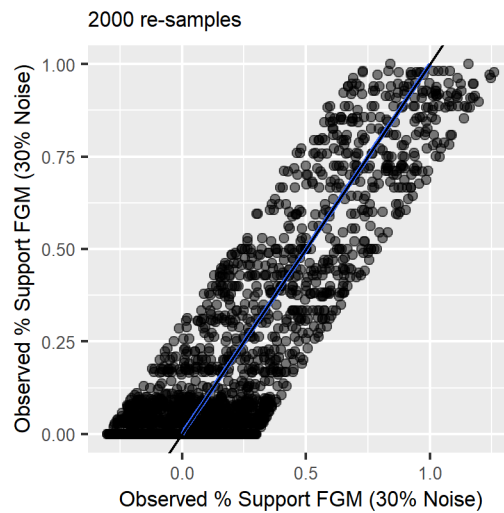


Figure 63: Bivariate Relation: % Support FGM by % Support FGM (based on 2000 re-samples from 376 communities with 30% uniform noise added)

Conclusions

This chapter developed and implemented empirical strategies for estimating parameters in the general coordination ABM of FGM described in Chapter 6. Calibration efforts prioritised those parameters revealed to have the highest importance under the Global Sensitivity Analysis (see Chapter 6). Empirical validation tests indicate this calibration effort had some success. Using community-level calibrated parameter estimates as ‘input’ induced a strong positive association between the predictions of the ABM and observed rates of FGM practice in real communities. Moreover, the ABM was able to reproduce key empirical patterns in the prevalence of FGM, observed in the 2005 Senegalese DHS survey, under calibrated parameter settings.

These results establish a certain degree of credibility for both the ABM and the calibrated parameter space. They suggest that the ABM is sufficiently credible that its continued exploration as part of the *possibilistic failure scenario analysis* of the next chapter is worthwhile, and may yield valuable insights. Moreover, the results support the notion that particular attention should be paid to the failure scenarios anticipated by the model under calibrated parameter settings.

Chapter 8: Possibilistic Failure Scenario Analysis

Introduction

This chapter uses a strategy of *possibilistic failure scenario analysis* (PFSA) to apply the general ABM model developed in previous chapters to policy problems. PFSA represents an alternative to ‘social engineering’ applications of formal models to policy. In particular, PFSA is an alternative to using models to make predictions about ‘optimal’ or ‘successful’ interventions designs. Social engineering approaches may be untenable when there is substantive uncertainty about model design (see Chapter 3). PFSA, on the other hand, is compatible with uncertainty about model adequacy.

The core strategy in PFSA is to use formal models to identify *possible scenarios* in which intervention efforts could be disrupted by social dynamics. Unlike social engineering, PFSA construes these scenarios as credible possibilities, rather than predictions. PFSA is also distinct in its emphasis on helping practitioners to *recognise* possible sources of failure, rather than instructing practitioners to take particular actions. As such, the information provided by PFSA is deliberately structured to play a subordinate role to the practical knowledge and experience of practitioners.

As discussed in Chapter 3, PSFA builds directly on the intellectual tradition of scenario planning and on the ideas expressed in Aodha and Edmonds (2017) and Edmonds and Aodha (2019). Edmonds and Aodha (2019) argue in favour of a re-orientation of the role of models of complex real-world systems (including social systems) toward *possibilistic* analysis. As an example, they present a model of marine ecology, specifically, the process of fish population replenishment. This model suggests that, in some scenarios, relatively conservative fishing practices could lead to a sudden collapse of marine populations. The model is un-tested and such scenarios are only ‘possibilities. Yet, a recognition that fishing practices *might* have this impact may have important implications for fisheries management. The authors point out that a past overreliance on optimistic formal predictions about fish stock replenishment contributed to a collapse of coastal fish stocks in Newfoundland, Canada.

In this chapter, I report a PFSA of the general ABM model developed in Chapter 6 and calibrated and tested in Chapter 7. The aim of my analysis was to identify ‘possible scenarios’ in which coordination dynamics might lead to a failure of community-level anti-FGM

intervention programs. Seven distinct scenarios are reported, with demonstrations through simulated experiments, including simulated interventions.

Intervention Failure and Past Experiences

There have been attempts to develop an understanding of possible sources of failure of anti-FGM interventions, based on past intervention experiences. Here I focus on a recent meta-review by Johansen et al. (2013) of ‘What works and What Does Not’ in interventions aimed at promoting the abandonment of FGM. The authors’ review, which draws on past reviews, and the authors’ own fieldwork experience, provides a useful point of contrast between current insights into intervention failure that have accumulated through fieldwork experience, and issues of intervention failure that can be addressed through formal modelling.

Many of the insights provided by Johansen et al. (2013) review, while undoubtedly valuable, are outside of the scope of coordination dynamics. Much of the discussion focuses on the failure to *persuade* intervention participants to oppose FGM. In this vein, Johansen lists ‘Risks and Disadvantages’ including poor quality information, defensive reactions, inadequate content and training, materials poorly adapted to the local culture, lack of collaborative approach to education and perceptions that anti-FGM advocates are insincere.

However, Johansen et al. (2013) *also* recognise social dynamics as a key source of intervention failure. They note that changes in attitudes among community members may not necessarily translate into decisions to abandon the practice, owing to social pressures.

“A major reason for this apparent contradiction between attitude and behaviour is a social and cultural pressure to uphold the tradition.” (p.2)

They note also, consistent with the emphasis of social norm theorists, that social norms may ‘overrule’ (p.3) the concerns of individuals about health-risks information. In discussing intervention activities that might be hoped to actually address social pressures to practice FGM, such as Alternative Rites of Passage (ARP)⁷⁰, community-led activities, and public declarations of abandonment, Johansen et al. (2013) describe a number of ways in which social dynamics might upset these efforts. ARP interventions in which there is limited integration of the ‘whole community’ in the alternative celebrations, and which were

⁷⁰ These aim to retain the cultural ceremonies surrounding cutting, while dropping the ‘cutting component’ (cf Droy et al., 2018).

insufficiently proceeded by attitude change, experienced reduced efficacy (p.5). They note that community-led intervention, such as Tostan, have achieved significant success in some communities in Senegal (p.6, see also, Diop et al., 2008). Yet, it achieved less success in other communities, and in other national contexts, including Burkina Faso and Somalia. As such, they suggest that ‘*Success Varies between Communities*’, which they attribute to social and religious factors (p.6). They acknowledge that ‘public statements’ (i.e. public declarations) aim to ‘express and facilitate’ (p.6) a shift away from the social convention of FGM. However, they also suggest that one of the ‘risks’ with this approach is that if only certain sub-groups of the community are involved in the declarations, this may fail to result in widespread abandonment, especially if the sub-group is insufficiently influential over FGM decision-makers⁷¹.

There are clear connections between these issues experienced by fieldworkers, and the issues addressed through modelling dynamics of intervention. Modelled dynamics address *who* in the community is targeted by the intervention, the extent of attitude change, and whether this will lead to stable or widespread change through coordination dynamics. The potential contribution of formal analyses to the informal insights of practitioners is also clear. Formal modelling can be used to explore a range of scenarios under which a sub-group of participants may be insufficient to generate stable or widespread change, or when the social circumstances of a community might make change difficult. In particular, formal modelling can be used to explore the potentially complex social dynamics that may link particular circumstances to particular ‘failure’ outcomes.

Intervention Failure in Past Coordination Models of FGM

The primary contribution of recent formal analyses of coordination models, to a project of FGM intervention ‘failure analysis’, has been through the discussion of the possible effects of different distributions of thresholds in the community. As discussed extensively in this thesis (Chapters 2 and 5), some distributions of ‘thresholds’ in simple coordination models (such as L-shaped distributions) lead to a prediction that public declarations *without preference change* will fail to generate stable or widespread abandonment of FGM. Moreover, small interventions with extreme preference change (targeted actors unconditionally abandon FGM), may fail to generate widespread abandonment under certain threshold distributions

⁷¹ They suggest, for example, that declarations primarily involving men may be problematic in cases where FGM is seen as the responsibility of women.

(Efferson et al., 2019). These issues interact with questions of *targeting bias*, whereby interventions targeting supporters of FGM may have different results to interventions targeting opposers of FGM (see Chapter 5).

Efferson et al. (2019) have recently extended this analysis, in particular, by exploring the possible effects of social-network structure (in combination with different targeting biases) on *positive-spillovers* from intervention efforts. They found that any homophily within the network (actors who support FGM are more likely to be connected to others who support FGM, and vice-versa) can reduce positive spillovers when targeting supporters of FGM. Conversely, when targeting actors at random, or targeting agents opposed to FGM using a small intervention (e.g. 10% of actors) homophily could improve spillovers.

These scenarios are treated more like predictions than ‘possibilities’ in existing analyses. However, they can be thought of as failure ‘possibilities’ for the purposes of PFSA. The analysis in this chapter strengthens and extends these insights in a number of respects.

The ABM representation of the intervention incorporates novel features, including a *coalition formation process* and *intermediate* levels of preference change. This allows the analysis to consider new kinds of intervention failure outcomes such as coalition collapse, as well as different kinds of intervention scenarios (e.g. related to moderate changes in preferences). This was impossible under Efferson et al.’s (2019) analysis – where preference change was extreme, and change therefore always stable. My analysis reveals, among other things, potentially strong contradictory effects of social dynamics which could be ‘good for’ creating stable change but bad for creating widespread change (i.e. positive spillovers).

The ABM model also contains other new, influential and credible design elements (see Chapters 5 and 6), which further extend the range of scenarios that can be considered through formal analysis. These include issues related to explicit norm-enforcement and different distributions of authority within the community.

Finally, the credibility of the scenarios considered through the ABM benefit from the methodology of its design and its engagement with empirical data. All of the scenarios considered in this chapter can be demonstrated to occur within independently *calibrated* values of key parameters (μ_H , σ_H^2 , s_2 and ρ_{Hy}), and using plausible values of un-calibrated parameters. Also, the model itself is the first in the field to have been independently validated against multiple detailed patterns in community-level data with some success (see Chapter 7).

Approach to PFSA

The task for PSFA, in this context, was to identify possible scenarios under which social dynamics might lead to anti-FGM intervention failure. As detailed in Chapters 5 and 6, interventions are simulated directly within the ABM model. The outcomes of those simulated interventions can be classified as ‘successes’ or ‘failures’ based on observing the simulated population after the intervention.

The experiments used to explore failure scenarios in the ABM were very similar to the ‘Dynamics of Intervention’ experiment described in Chapter 5. In every case, a set of parameter values, related to the scenario in question, is chosen. Then, the simulation is ‘initialised’ with all actors practicing FGM. Next, a simulated intervention is undertaken. The size of this intervention and its other features (incl. educational strength, organised diffusions, etc.) are all determined by chosen parameter values (see Chapter 6). The simulated intervention always includes a ‘targeting’ stage, where a subset of actors in the population are targeted by the intervention. These actors receive the educational effects of the intervention, and then decide whether to join an *initial coalition* of actors conditionally prepared to abandon FGM. Members of this initial coalition then decide if they are prepared to abandon FGM, conditional on others doing so too, this results in a *coalition formation* process, with some agents potentially leaving the initial coalition, until it either stabilises into a *final coalition* or collapses entirely (see Chapter 5 and 6). After the simulated intervention ‘ordinary’ coordination dynamics occur across the whole population of agents for 100 time-steps. 2000 agents were used in all experiments.

While conducting these simulated interventions was technically straightforward, the challenge for the analysis was to gather information about the *conditions* under which these simulated interventions succeed or fail and to *organise* this information in a way which could be of use to a wider audience – especially practitioners involved in anti-FGM interventions.

One might imagine that the best way to identify failure scenarios would be to list parameter settings where this occurs. However, this would be difficult in principle and not in itself a useful PFSA result. The ABM model contains 36 independent parameters that could affect simulated intervention failure. Even if we assumed (falsely) that each parameter had only 2 possible values – this would correspond to a set of 68,719,476,736 possible parameter settings. Simulated interventions in the ABM model can be completed at a rate of about 1 every 6 seconds of real-time when parallelising simulations across 8 cores of a modern CPU.

As such, conducting simulated interventions covering all parameter setting could take over 13,000 years⁷².

Even if we did have access to all of these results, this would not, on its own, represent a useful PFSA. Simply presenting combinations of parameters and noting that they correspond to intervention failure will be of little use to anyone. We want to know *why* failure occurs under certain parameters' settings, what changes in parameter values would reverse the change, and so on. Most importantly, we want to better understand the *dynamic processes* that cause intervention failures in the model. It is this kind of insight that may help to expand practitioners' knowledge of how social dynamics could obstruct their efforts. As such, dynamic processes are the appropriate 'unit of analysis' for PFSA.

I aimed to provide an account of how different 'failure processes' in the model work and the circumstances under which they could occur. In the interest of making the results accessible and memorable to a wider audience, I refer to these processes as 'effects' and suggest relevant labels for them. A failure scenario was therefore defined as a scenario in which one (or more) of these 'effects' occurs.

To the best of my knowledge, there are no automated techniques to discover all of the different processes that can operate within a complex computer simulation. As such, I took *an open-ended exploratory approach*. I began with my own intuitions about the processes operating in the model, including insights developed in Chapter 5. I also undertook a battery of exploratory analyses of the model. To do this, I conducted 24,000 simulations, including a simulated intervention at the beginning of the simulation. These interventions used randomly chosen parameter values, with the exception of the μ_H , σ_H^2 , s_2 and ρ_{Hy} parameters, which were chosen randomly from rows of the calibrated parameter space developed in Chapter 7. I analysed the results of these simulation runs using Monte Carlo filtering (Saltelli et al., 2008), regression-based approaches and decision-tree models (see overview in Appendix F1). I aimed to identify patterns in the relationship between model parameters and the outcomes of the simulations (which included a simulated intervention). Reports for each parameter, and other outputs of these explorations, are included in the DVD-ROM attached to the thesis.

⁷² That's if we took only a single sample for each unique parameter combination. On the bright side, K, the world's fastest super computer has 88,128 8-core CPUs. Using all of them, it might only take 54 days! (<https://www.extremetech.com/extreme/122159-what-can-you-do-with-a-supercomputer>)

These initial intuitions, combined with the exploratory analyses, provided a set of ‘hypotheses’ about the social dynamics leading to intervention failure in the model. I then confirmed and refined these hypotheses through an interactive exploration of the ABM model within the NetLogo modelling environment, as well as through the simulated experiments reported in this chapter.

I distinguished between three kinds of intervention failure. First, the failure to produce any reduction in the rate of FGM. Second, ‘negative spillover’. I defined ‘negative spillover’ as occurring if the number of actors abandoning FGM was *smaller than* the number of actors targeted by the intervention. This is the opposite of the lauded positive ‘spillover’ in which a small intervention leads to widespread change (Efferson et al., 2015, 2019). Third, partial change. ‘Partial change’ was defined as when some actors in the population keep practicing FGM despite the intervention. These different kinds of failure can all occur within the ABM simulated interventions.

The results of the PFSA are presented as a set of seven failure scenarios plus a re-examination of the role of heterogeneity of preferences under calibrated parameter values. To avoid excessive repetition, I provide two ‘worked’ examples, with a summary of the essentials of other failure scenarios presented in Table 13. Detailed discussions of the scenarios presented in Table 13 are provided in Appendices F2-F7, which include outputs of simulated experiments and other demonstrative outputs of the simulations.

Partly due to space constraints, the full details of the (thirty or more) parameters held constant in the simulated experiments associated with each failure scenario are not discussed. In fact, discussing these would be somewhat tangential to the PFSA approach of this chapter, which is to demonstrate *possible* failure scenarios, not to make claims about the generality of these effects (which would, in any case, depend on speculative assumptions about the ‘true’ distributions of uncalibrated parameters of the model). Nevertheless, details needed to replicate each of the experiments are included with the source code of the general ABM (within the ‘Behaviour Space’ module). The parameter values used in each experiment were drawn from the range of ‘calibrated’ values of key parameters, where available (see Chapter 7). When using uncalibrated parameters, values were chosen from ‘plausible’ ranges (see Table 9, Chapter 6) to illustrate the effect in question.

Failure Scenarios

Revisiting Heterogeneity of Preferences

The empirical analysis undertaken in Chapter 7, presents an opportunity to revisit the question of heterogeneity of preferences using calibrated values for certain ‘preference’ parameters: H_i and σ_H^2 . Typically, past analyses showing the sensitivity of coordination models of FGM to different preference distributions have relied on speculation about possible distributions⁷³ (Efferson et al., 2015, 2019; Platteau et al., 2017). Yet, the empirical measures from Chapter 7 allow for an assessment of the possible effects of the *empirically observed* variation in distributions of preferences.

Empirical estimates from Chapter 7 (see Figure 43) found that the average perceived value of FGM at the community level (μ_H), expressed in the unit interval, was typically between 0.4 and 0.8 in communities where the practice was widespread. The important question that then arises is whether this degree of ‘naturally occurring’ variation in preferences might be sufficient to affect the success of intervention efforts.

Explorations of the ABM using this calibrated range of parameter values suggest that it definitely *could*.

Figure 64 and 65 display the outcome of two kinds of simulation ‘scenario’ (including an intervention), one with ‘high’ perceived value of FGM ($\mu_H = 0.8$). The other with low average perceived-value of FGM ($\mu_H = 0.4$). Both targeted half the population and were repeated 10 times.

These figures are designed to summarise each stage of the simulation in a single graphic. The design, which is somewhat unusual, is repeated in this chapter (and in Appendices F2-F7). So, readers may wish to familiarise themselves with it at this stage. The figures display the progress of each simulation run (of 10) as a line running from left to right (representing the intervention from start to finish). An initial heuristic for each line is that when it goes up, this is ‘bad’ (from the point of view of FGM abandonment), and when it goes down, this is ‘good’. The lines are divided into coloured segments, indicating the stage of the intervention that the line refers to:

⁷³ The exception to this is Novak (2016), however she only estimated upper and lower bounds for threshold distributions, not the properties of preference distributions. Moreover, her empirical method had a number of limitations, see Chapter 7.

- First, where the line is red, it represents the *targeting* stage of the intervention. It literally represents the proportion of agents *not* targeted by the intervention. So, the lower the end-point of the line, the larger the intervention.
- Second, where the line is green, it represents the formation of an initial coalition among targeted actors. It literally represents the proportion of agents (in the population) who did *not* join the initial coalition. So, the lower the end-point of the line, the larger the initial coalition.
- Third, where the line is blue, it represents the formation of a final stable coalition, from among the initial coalition. It literally represents the proportion of agents (in the population) who did *not* join the final stable coalition. So, the lower the end-point of the line, the larger the final stable coalition.
- Fourth, where the line is purple, it represents the process of coordination after the intervention. It literally represents the proportion of actors practicing FGM (or *not* abandoning). So, as the line drops, this indicates that more actors are abandoning FGM, as a result of coordinating with intervention participants.

Therefore, the x-axis of each figure *first* represents the stage of the intervention (which has no explicit ‘time-steps’), *then* represents the time-steps of the simulation after the intervention. Likewise, the meaning of the y-axis depends on the colour of the line segment; first it represents the proportion of agents outside the coalition formation process, *then* it represents the proportion of agents practicing FGM. The horizontal line represents the size of the intervention. Literally, it is the proportion of actors *not* targeted by the intervention. If the final rate of FGM (purple line segment) goes below this line, the intervention has ‘positive’ spillover. If it is above the line, it has ‘negative’ spillover (fewer actors abandoned FGM than were targeted).

Returning to the question of variation in average preferences, Figure 64 demonstrates, intervention outcomes where the average approval of FGM is toward the lower end of the naturally occurring range ($\mu_H = 0.4$), and the intervention size was 50%. We see that a small final stable coalition formed in all simulations and the influence of this coalition spread to other members of the population through coordination dynamics. Here, this led to complete abandonment of FGM in all simulations. By contrast, when approval was set at the upper end of the observed range (Figure 65, $\mu_H = 0.8$), and all other parameters were held constant, coalition collapse occurred (no final stable coalition), leading to no reduction in FGM practice from the initial rate of 1.

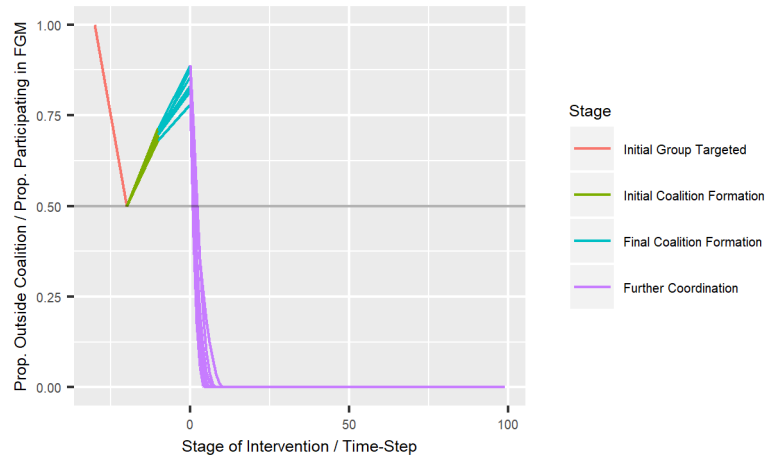


Figure 64: Simulated Intervention with a Low Average Approval of FGM ($\mu_H = 0.4$)

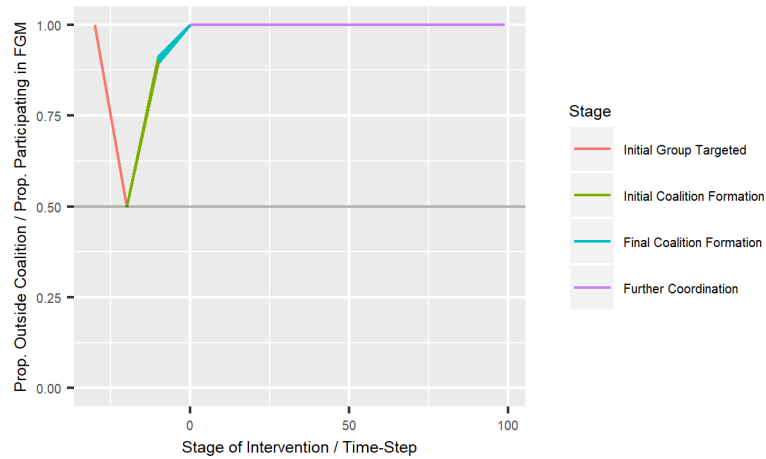


Figure 65: Simulated Intervention with a High Average Approval of FGM ($\mu_H = 0.8$)

Readers of previous chapters should have an intuition for the dynamic underlying this effect. Differences in the average support for FGM in the community correspond to differences in the distribution of willingness of actors to abandon the practice. However, the simulated experiments (Figures 64 and 65, above) show that ‘naturally occurring’ community-level differences in this distribution are sufficient to generate complete failure, or success, of relatively strong intervention efforts, in different contexts.

These experiments addressed differences in the ‘average’ approval of FGM in the community (μ_H). As part of my discussion of ‘Social Hump’ effects (Table 13 and Appendix F5), I show that ‘naturally occurring’ differences in the *variance* of preferences at the community level (σ_H^2) could also have substantive effects on intervention outcomes.

‘Worked Example’ Failure Scenario 1: Social Cost Imbalance Effects

In Chapters 5 and 6, I introduced the formal representation of features *other than* the approval of FGM that might influence an individual actor’s willingness to practice FGM⁷⁴.

The first was different degrees of symmetry in the coordination incentives that actors face for practicing, as opposed to abandoning, FGM. In the model, this is controlled by the s_1 parameter ($\in [0,1]$) which scales the maximum social cost for *practicing* FGM. The maximum social cost for *abandoning* FGM is an arbitrary constant: \hat{M} . The maximum social cost for *practicing* FGM in the model is $\hat{M} \cdot s_1$. So, when $s_1 = 0$, there is no social cost for FGM practice. When $s_1 = 1$, the maximum costs are symmetric for practice or abandonment.

The second was possible asymmetries in the *relation between* social pressure and social costs for practicing or abandoning FGM. In the model, this is controlled by three global parameters: $V \geq 1$, $\delta_1 \in [-1,1]$ and $\delta_2 \in [-1,1]$. To understand these parameters’ effects, readers should refer to Chapter 6. Put simply, as long as V is greater than 1, increasing δ_1 means that the social costs to *abandon* FGM rise more rapidly as a function of increases in *pro-FGM activities* (i.e. in a non-linear way). Conversely, increases in δ_2 mean that costs to *practice* FGM rise more quickly as a function of increases in *anti-FGM activities*.

These possible *asymmetries* in the social costs of FGM practice versus abandonment created by these parameters are referred to as *social cost imbalance effects*. The easiest way to gain an intuition for these effects is to think of them as impacting the willingness of actors to practice or abandon FGM.

Social cost imbalance effects are of particular interest in this chapter because they could occur *independently* of changes in the *approval* of FGM (H_i) within a population. Thus, they could ‘make the difference’ between intervention failure or success in communities with the same average view of the intrinsic value of FGM.

Figures 66 and 67 illustrate the potential for social cost imbalance effects to produce intervention failure. Two (otherwise identical) conditions were used. In the first, social costs were balanced, so $\delta_1 = 0$, $\delta_2 = 0$ and $s_1 = 1$. In the second condition, social costs were

⁷⁴ I also argued that these issues cannot be simply obfuscated to a question of ‘variation in thresholds’. Focusing only on threshold variation obscures the sources of that variation and is hard to operationalise. Also, thresholds lose their meaning under certain model designs (see Chapter 5).

heavily unbalanced in favour of continued FGM practice. In this second condition, $\delta_1 = 1$ and $\delta_2 = -1$ (s_1 remained at 1). Both conditions were repeated 10 times.

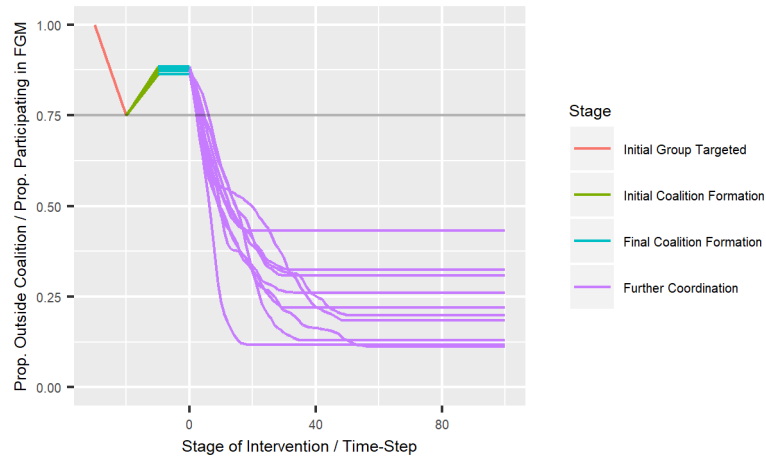


Figure 66: Simulated Intervention with Balanced Social Costs ($\delta_1=0$, $\delta_2=0$)

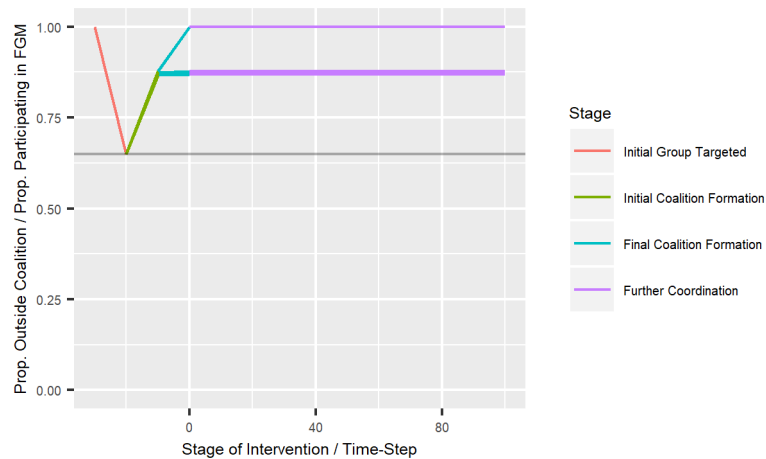


Figure 67: Simulated Interventions with Unbalanced Costs ($\delta_1=1$, $\delta_2=-1$)

We can see that when the costs were balanced (Figure 66), the intervention was able to produce significant reductions in the rate of FGM in the community. Where the costs were heavily imbalanced (Figure 67) in favour of FGM practice, then there were two distinct outcomes. In some cases, a stable coalition failed to form, meaning zero reduction in FGM practice. In other cases, a small stable coalition formed, but this influence failed to spread to the rest of the population.

So, social cost imbalance effects could result in intervention failure when there are limits to the extent of social costs that actors could pay of practicing FGM, or when social costs for *abandoning* FGM strongly increase as a function of *pro-FGM* activity, or when social costs for *practicing* FGM increase slowly as a function of *anti-FGM activity*. Practitioners may wish to

pay attention to how FGM activity is related to social costs in the villages they target. If for example, a small number of actors practicing FGM are able to monopolise important community activities, such as the education of girls or decision-making in the community, then social costs for abandoning might remain high with only low levels of FGM practice. Conversely, FGM practitioners' ability to monopolise these activities (perhaps because of the activities' cultural association with FGM), might imply that the social costs actors face for *practicing* FGM are low even when most others abandon.

Summary of Failure Scenarios 2-7

Failure scenarios 2-7 are summarised in Table 13, with further details of each scenario, as well as demonstrative simulation outputs given in Appendices F2-F7. The table summarises the circumstances under which the failure scenarios may occur, the process associated with each, the failure outcomes these can create, and other relevant observations. The latter observations include potential contradictory effects, where the effects in question could both help and hurt the intervention in different ways.

Failure Scenario	Circumstances	Dynamic	Failure Outcome(s)	Other Observations
Out of Reach Effects (#2)	The 'educational effect' (i.e. persuasiveness) of the intervention is weak (z_2), many actors are reluctant to join the initial coalition (z_3) and only a subset of actors are 'susceptible' to the influence of a small coalition	Fewer actors are persuaded to join the coalition of abandoners, and the influence of small coalitions only spreads to a 'susceptible' subset of actors	Increasing the intervention size, as a strategy to improve outcomes becomes highly inefficient, leading to negative spillovers and only partial change.	-
Supporters Gambit Effects (#3)	<p><i>Circumstance 1:</i> The 'educational effect' (i.e. persuasiveness) of the intervention is <i>weak</i> (z_2) <i>AND</i> the intervention targets <i>supporters</i> of FGM.</p> <p>OR</p> <p><i>Circumstance 2:</i> The 'educational effect' (i.e. persuasiveness) of the intervention is strong (z_2) <i>AND</i> the intervention targets <i>opposers</i> of FGM.</p>	<p><i>Circumstance 1:</i> Few actors agree to join the initial coalition, and the initial coalition is liable to collapse</p> <p><i>Circumstance 2:</i> A large stable coalition is formed, but its effect fails to spread through the whole community, because supporters of FGM have been unaffected.</p>	<p><i>Circumstance 1:</i> Potential for coalition collapse (i.e. no reduction in FGM), which would not have occurred if opposers of FGM had been targeted</p> <p><i>Circumstance 2:</i> Potential for only partial change, whereas change would have been universal if supporters of FGM had been targeted.</p>	<i>Contradictory Effects:</i> Targeting supporters of FGM can help (by promoting widespread change) or hurt (by leading to coalition collapse) the intervention.

Localised Deviance Effects (#4)	Social interaction is highly ‘localised’ as a result of low network connectivity, coordination within households and/or network homophily (cumulative). This leads to variation in the <i>social pressures</i> facing actors.	A coalition of abandoners stabilises easily within a ‘section’ of the network where actors oppose FGM, but their influence cannot spread to pockets of the network in which actors are resistant to change.	When the intervention is weak (small, low educational effect), coalitions can still form when, otherwise, coalition collapse would have occurred. When the intervention is strong (large, strong educational effects), only partial change occurs.	<i>Contradictory Effects:</i> Localised deviance can help (by preventing coalition collapse) or hurt (by preventing widespread change) the intervention.
Social Hump Effects (#5)	<p><i>Circumstance 1:</i> The distribution of approval of FGM is roughly symmetric, and there is low variation (within empirically calibrated values of σ_H^2) in preferences (strong social hump effect)</p> <p><i>Circumstance 2:</i> The distribution of approval of FGM is roughly symmetric and there is high variation (within empirically calibrated values of σ_H^2) in preferences (weak social hump effect)</p>	<p><i>Circumstance 1:</i> Most actors require (roughly) a majority to abandon before they will, so small interventions struggle to form a stable coalition.</p> <p><i>Circumstance 2:</i> Many actors require few others to abandon before they will, and many actors require <i>almost everyone</i> to abandon before they will. So small stable coalitions can form, but their influence doesn’t spread effectively</p>	<p><i>Circumstance 1:</i> Small interventions fail to form a stable coalition, leading to no reduction in FGM.</p> <p><i>Circumstance 2:</i> Small interventions can form coalitions but, this doesn’t lead to widespread abandonment.</p>	<i>Contradictory Effects:</i> Social Hump effects can help (by allowing social influence to be widespread) and hurt (by leading to coalition collapse) the intervention.
Power Imbalance Effects (#6)	<p>Those with more authority tend to support FGM.</p> <p><i>Strongly exacerbated by:</i> High variation in authority and L-shaped distribution in authority (low average authority).</p>	The more that the ‘weight’ of authority is concentrated among supporters of FGM (through the correlation of authority with the approval of FGM, and effects of the distribution of authority), the harder it is for large stable coalitions to form, and the weaker their influence on others in the community.	The potential for coalition collapse is greatly increased (leading to partial abandonment of FGM), the influence of the coalition on others is greatly weakened (leading to partial abandonment of FGM)	<p><i>Intuitive ‘solutions’ may fail:</i></p> <p>The intuitive solution: to target those with high authority, may fail, because it <i>also</i> means targeting supporters of FGM – leading to coalition collapse (see ‘Supporters Gambit Effects’)</p>

Enforcement Imbalance Effects (#7)	Many actors are willing to enforce FGM practice, but few are willing to enforce FGM abandonment	Intervention participation will only relieve pro-FGM pressures when norm enforcers are involved. Coalitions may relieve <i>pro-FGM</i> social pressure, but will not <i>create</i> anti-FGM social pressure.	The effects of stable coalitions will not spread into the population, leading to only partial change.	<i>Different kinds of 'solutions':</i> Targeting pro-FGM enforcers exclusively may help, but it may be more important to create <i>new</i> anti-FGM enforcement through the intervention.
--	---	--	---	--

Table 12: Summary of Failure Scenarios 2-7

As Table 12 shows, there were a range of circumstances under which social dynamics could lead to intervention failure within the simulation model. These included aspects of the social situation in the community itself (e.g. localised deviance effects), and aspects of the intervention (e.g. supporters-gambit effects). These circumstances could also *interact*, such as through the combined effects of weak educational strength and large intervention size, leading to negative spillovers (e.g. out-of-reach effects). Failure scenarios also, often, had contradictory effects, with some scenarios both precipitating coalition failure (under some kinds of intervention) and precipitating universal abandonment (under others)

In some scenarios, intuitive solutions to avoiding failure were shown to have their own failure possibilities, such as the potential difficulties presented in trying to recruit authoritative actors under power-imbalance effects, or the potentially limited effectiveness of focusing on *preventing* pro-FGM enforcement, rather than *creating* anti-FGM enforcement.

Practitioners attention should be drawn to this range of scenarios, and the dynamic processes associated with them. While I contend that practitioners themselves are best placed to judge the relevance of these scenarios in the communities where they operate, the scenarios summarized in Table 13 and above may help them to recognise these scenarios when and if they occur, and to consider what they might do to try to mitigate them. The scenarios do this by pointing to key concepts (such as localised interaction), and providing a model-based analysis of the connection between these concepts and intervention failure possibilities.

Discussion and Conclusions

The failure scenario analysis undertaken in this chapter was designed to highlight a range of ways in which coordination dynamics might precipitate the failure of anti-FGM interventions. It points to a range of such scenarios and touches on different kinds of intervention failure.

As with any complex simulation model with a large parameter space, the full range of possible dynamics of the model are not automatically known to the modeller. They ‘emerge’ from its many interacting parts, and must be discovered through exploratory simulation. Further analysis might discover other interesting failure scenarios that can occur within the simulation. Further applications of the model should also further explore other potential interactions between the dynamic processes underlying the different failure scenarios. As with other parts of the modelling strategy of the thesis, the effort is meant to be cumulative with future work (see also Chapter 9).

There are a number of reasons why a model-based failure scenario analysis of this kind should be taken seriously by policy-makers and others interested in anti-FGM interventions.

First, a number of the insights generated by the analysis are non-intuitive, or only partially intuitive. There has been, for example, little recognition in past discussions of FGM of the possibility that social dynamics can create trade-offs between different kinds of intervention failure, and no recognition outside of formal analyses (that I am aware of, beyond the standard implication that there is ‘tipping-point’ that needs to be reached). Instead, features of interventions and social dynamics are typically represented as either supporting or harming intervention efforts. Yet, my analyses illustrate a range of ways that coordination dynamics *could* help *and* harm intervention outcomes in different ways. I also show that these effects can interact with different aspects of intervention design.

Second, there is a substantial difference between post-hoc intuition and formal demonstration. The failure scenarios demonstrated here are shown to be logically coherent (in that they can be produced by a plausible formal model). Also, they are shown to occur in the context of a social simulation that itself has certain degree of credibility. The ABM model used to facilitate the simulations is built on a core theory which is recognised in existing literature, and has empirical support (see Chapter 4). The model includes features selected with reference to existing empirical and theoretical literature and has been subjected to some independent validation against observed empirical data (Chapter 7). Furthermore, all

simulations in this chapter were conducted using values of key parameters that were restricted to an empirically supported range (see Chapter 7). All of these factors lend some credibility to the failure scenarios discussed here, that would not exist if they were presented merely as informal theoretical speculation.

Certainly, if we put aside questions of credibility, it is true that one could speculate about coordination dynamics and intervention failure without a model. But in this case, how would one check that a given speculative scenario actually *could* occur as a result of coordination dynamics? Intuitions about complex social dynamics can easily be faulty and incoherent. We would need an act of formalisation to provide a basic check against the coherence of such intuitions. To adapt a phrase from Epstein's (2006: xii) famous discussion of social simulation, if you didn't simulate it (in a formal model) you didn't show it!

While the analysis here is meant to help enhance practitioners' awareness of different *possible* effects of social dynamics on their efforts, I would not, recommend that readers attempt to convert the results of this analysis into a set of positive recommendations. I do not think that practitioners (or other modellers) should try to identify an 'optimal' strategy that avoids all of these scenarios. Nor should they try to predict which scenarios will occur in a given intervention and alter their strategy to offset them. Put simply, the analysis provided here is designed to help practitioners recognise these scenarios when and if they occur. It is not intended to promote the view that they can be predicted before the fact. There is simply too much remaining uncertainty about the model to support this kind of application.

Nevertheless, it would be somewhat obtuse not to acknowledge some of the remedies to these failure scenarios that *are* prescribed in the model. These are noted here, with the proviso that there is no guarantee that their effectiveness in the model will translate into effectiveness in a real population.

Putting aside questions of efficiency, simulated interventions which are very large ($z_1 \approx 1$) and which have very strong educational effects ($z_2 \approx 1$), are likely to succeed in the model (i.e. produce large or complete reductions in the rate of FGM) irrespective of other factors. Practitioners should, especially, not be discouraged from trying to improve the educational effect of their interventions. This has a very positive effect on the outcomes of simulated interventions in the model.

In certain situations where it is difficult to form a stable abandonment coalition (e.g. when social hump effects are present), stability can be enhanced by building social relationships

(social network connections) *between* coalition members ($z_5 > 0$). This improves the balance of social costs and benefits that actors face if they remain in the coalition (because unconnected coalition members join one another's social reference groups). These effects were already observed in Chapter 5 in relation to Model Variant 11 "Building Social Relationships Between Intervention Participants".

In some cases, out-of-reach effects can be alleviated if actors targeted by the intervention *also* influence others and recruit them to join the initial coalition ($z_4 > 0$) (see discussion of Organised Diffusion in Chapter 5). In effect, this transfers the 'costs' of increasing the size of the intervention, from the practitioners to the participants. As such, the intervention can target a small number of actors, but then these actors recruit others, creating the equivalent of a 'large' intervention.

As noted in Chapter 5, some community-level interventions aspire to have this effect. A process of 'organised diffusion' beginning with a small group, is meant to spread the initial coalition through the social network of the population (Mackie and LeJeune, 2009). Clearly, if this *does* occur, then it would be very helpful to the intervention. Whether this actually *can* be achieved is a question for practitioners. Some organisations have certainly suggested that it might (UNICEF, 2007). The out-of-reach effect simply means that *if* increasing the intervention size is (in some sense) costly for practitioners, it could end up being a very inefficient way to try to generate change.

Chapter 9: Final Conclusions and Avenues for Further Development

Introduction

The later chapters of this thesis (4-8) have implemented the modelling strategy developed and articulated in Chapter 3. The principal aim of this activity has been to develop, demonstrate (and advocate for) a new methodological framework and strategy for model development and application when studying the social dynamics of FGM (hereafter, “the field”). In particular, a strategy has been developed (and successfully implemented) which helps to address the fundamental problem of model design uncertainty (as detailed in Chapter 1, and shown to be of critical importance to the field in Chapter 2).

This has been achieved through the systematic application (and combination) of model-based (sensitivity and robustness analysis), empirically driven (calibration and validation) and application-focused (possibilistic failure scenario analysis) methodological techniques.

The implementation of this strategy is detailed in previous chapters. In this chapter, I begin by focusing on the key contributions of the thesis, and the way that these have moved the field forward. These contributions include ‘global’ contributions that set new directions for the field as a whole (such as setting new standards of engagement with empirical data) and ‘local’ contributions that progress specific issues of importance in the field (such as providing new quantitative evidence favouring *social norm of coordination* as a theory of decision-making). Contributions span both methodological, as well as applied issues in the field.

The first two sections of this chapter summarise these contributions. The first focuses on methodological contributions. The second addresses other contributions. Both sections are organised to roughly follow the chronology of the thesis – highlighting the global and local contributions of each thesis chapter.

The third section considers the limitations of the research undertaken in the thesis and refers to the methodological strategy employed by the thesis to identify productive avenues for future work. An important theme of this section is that having followed an *explicit* modelling strategy, future avenues for progressive model development, critique and adjudication can be more easily identified.

The fourth section offers a final thought on the current popularity of social norm theory and its applications to solving social problems. It highlights the very broad implications of model design uncertainty for the applications of the theory.

Methodological Contributions

The study of the social dynamics of FGM (hereafter, ‘the field’) has long interested both researchers and development practitioners (Efferson et al., 2019; Mackie, 1996; UNICEF, 2010). Moreover, insights into these dynamics have had a substantive impact on policy thinking (Efferson et al., 2015; Mackie, 2017) around FGM – which remains an urgent global social problem (UNFPA & UNICEF, 2019).

Formal modelling has been applied by a range of researchers to try to understand the dynamics of FGM (see Chapter 2). An initial contribution of the thesis (Chapter 1) was to establish a formal rationale for this kind of activity and to identify the foundational methodological challenges associated with it. Previously there has been little general discussion of the rationale for using formal modelling in this field, nor recognition of the fundamental methodological challenges associated with doing so. Yet consideration of these issues is essential for understanding and evaluating the role that modelling plays in efforts to understand (and ultimately, prevent) FGM.

First, an explicit rationale was presented for the use of formal modelling in this field, based on the capacity of formal models to explore the macro-level implications of (i.e. the social dynamics arising from) micro-level processes. This is an activity that cannot reliably be undertaken through other approaches to FGM, such as narrative qualitative theory (e.g. Hernlund and Shell-Duncan, 2007; Shell-Duncan et al., 2011) or quantitative regression analysis (Hayford, 2005; Kandala et al., 2009; Kandala and Shell-Duncan, 2019; Modrek and Liu, 2013). Second, a clear statement of the fundamental challenges associated with modelling social systems is provided, based on the concept of model design uncertainty. Model design uncertainty is a fundamental uncertainty about which features of a social system need to be included a model of that system (versus simplified or abstracted away), in order to furnish an adequate representation that will provide reliable insights into the system’s dynamics.

Recognising the fundamental challenge that model design uncertainty poses for the field, strongly implies the need to evaluate the success of the field in meeting this challenge to date. Yet, such an evaluation has been largely absent. While critical discussions of particular

models, especially Mackie’s social convention model (Mackie, 1996; UNICEF, 2007) have taken place in the literature (Efferson et al., 2015; Novak, 2016; Platteau et al., 2017), there have been no general reviews of formal modelling in the field (with the partial exception of Platteau et al., 2017) and in particular, none which address the fundamental problem of model design uncertainty, and the way that it might be affecting the progress in the field as a whole. Chapter 2 undertakes this evaluation and has important implications for the field.

First, a new and productive understanding of the criticisms and controversies surrounding Mackie’s social convention model of FGM (Efferson et al., 2015; Mackie, 1996; UNICEF, 2007) is provided. The problems associated with the model are problems of model idealisation, *not* problems with the underlying ‘social norm of coordination’ account of decision-making. This points to the investigation of idealisations in existing model designs as an important new avenue for progressive research in the field. For example, findings that the social convention model is not robust to the idealisation of ‘homogeneity of preferences’ (Efferson et al., 2015; Novak, 2016; Platteau et al., 2017) creates a new standard for the field – that heterogeneity of preferences must be included in future models. Moreover, my review identified idealisations in existing model designs in the field as a largely unacknowledged, but substantive, *general* threat to the validity of their use as policy tools. For instance, de-idealisations of the social convention model (e.g. Novak, 2016) still contain a range of simplifying assumptions (e.g. global interaction) whose potential impact hasn’t been explored by their authors, nor shown to be ‘safe’ for policy applications of the model.

Second, *diverging* (and largely subjective) choices about idealisations in the designs of different models of FGM in the field (such as those in Chesnokova and Vaithianathan, 2010 vs. Mackie, 1996) are shown to present a significant barrier to adjudicating between, or accumulating knowledge from, existing models of FGM.

My review also establishes that these problems have not been resolved through attempts to validate models empirically (where this occurs at all). Indeed, validation in the field has largely been restricted to comparing models to vague stylised regularities that lend support to some of the basic motivating assumptions of the model (rather than lending meaningful support to a particular model design).

Ultimately, Chapter 2’s evaluation of the field reveals a stark gap between the current focus of activity and the fundamental challenges associated with building models that can inform policy. While a considerable amount of research effort in the field has focused on arguments about the minutiae of a particular assumption in the social convention model (Efferson et

al., 2015; Novak, 2016; Platteau et al., 2017), the field faces much deeper and more entrenched challenges associated with the fundamental problems of model design uncertainty. There is currently little evidence of widespread or systematic engagement with methodologies or data that might address or manage this uncertainty. Yet without a resolution, it is unclear how the field can progress, or justify the application of its models to FGM policy.

Chapter 3 takes up the challenge of identifying individual methodological techniques, and an explicit *modelling strategy* that combines them, to allow the field to move towards: managing uncertainty in model design, the responsible application of uncertain models to policy, and (ultimately) progressive modelling.

Drawing on ideas from social simulation and the philosophy of model-based science (including Aodha and Edmonds, 2017; Edmonds and Aodha, 2019; Kuorikoski et al., 2010; Levins, 1993; Poile and Safayeni, 2016; Railsback and Grimm, 2012; Saltelli et al., 2008; Saltelli and Funtowicz, 2014; Schindler, 2013; Thiele et al., 2014 and Weisberg, 2006, 2012) the methodological elements proposed are: *sensitivity and robustness analysis* (model-based), and *targeted empirical calibration and empirical validation* (model-based).

Each of these techniques has seen some partial application in the field, yet none have been widely or systematically applied within a modelling project in the field⁷⁵, in particular, modellers have not found a way to combine them together into a holistic modelling strategy. Yet, used in combination these have the potential to greatly advance the management of model design uncertainty in the field through:

- Early identification of problematic idealisations (Robustness Analysis)
- Prioritisation of empirical calibration toward sensitive model design elements and parameters (Sensitivity Analysis / Empirical Calibration)
- A rigorous test of overall model adequacy (through independent and/or multi-criterion validation) *as well as a test of the adequacy of the previous model development steps*

⁷⁵ With the possible exception of Efferson et al. (2019), which was published toward the end of the work on this thesis and (arguably) included meaningful robustness and sensitivity analysis, but still without detailed empirical calibration, or any empirical validation.

In addition to managing uncertainty as part of *model development*, my modelling strategy addresses the management of uncertainty as part of *model application to policy*. The application of models to policy in the field has largely followed an ‘engineering’ paradigm (Aodha and Edmonds, 2017), in which models are constructed, and then predictions about optimal policies are derived from them (e.g. Chesnokova and Vaithianathan, 2010; Coyne and Coyne, 2014; UNICEF, 2007) in the form of claims about the kinds of policy actions that are likely to succeed. This paradigm fails to recognise the risk that model design uncertainty poses for model application in this field. The new approach proposed in Chapter 3 (PFSA, see below) is compatible with uncertainty in model design, and therefore offers the field a more responsible strategy for using imperfect models to inform policy efforts.

Drawing on recent arguments in the social simulation literature (Aodha and Edmonds, 2017; Edmonds and Aodha, 2019) as well as the intellectual tradition of scenario planning (Amer et al., 2013; Nair and Howlett, 2017; Schoemaker, 1995; Stirling, 2010; Volkery and Ribeiro, 2009), Chapter 3 proposes possibilistic failure scenario analysis (PFSA) as a novel alternative to the social engineering paradigm of model application in the field. PFSA has the potential to retain part of the key contributions of existing models, such as warning policy-makers of the potential difficulties of creating change purely through health information (Mackie, 1996), without making unsubstantiated claims about which interventions are ‘likely’ to succeed (either generally, or under specific conditions).

This new methodological agenda requires new tools. With the recent exception of Efferson et al., (2019), modelling in the field has relied on analytical methods, especially n -person game theory. Yet, I show that if the problem of model design uncertainty is to be taken seriously, these techniques cannot meet the needs of the field. In particular, the need to investigate a range of idealisations in existing models designs, and to address the low-level details of social interaction that frequently ‘matter’ in complex social systems (Chattoe-Brown, 2020b; Squazzoni et al., 2013). I recommend ABM as a novel alternative whose improved representational capacity will allow the field to tackle these issues directly.

The field has recently been mired in controversy surrounding the social convention model of FGM (Mackie, 1996; UNICEF, 2007), and Efferson et al.'s (2015) critique of it (see also, Mackie, 2017, Platteau et al., 2017, Novak, 2016). Chapter 2 traced this controversy to a failure (of parties on both sides) to distinguish clearly between idealisations in the design of this model, and the underlying theory of FGM as a social norm of coordination, which motivated the model.

As such, Chapter 3 (drawing on discussions in model-based science, Kuorikoski et al., 2010; Levins, 1993) introduces a methodological principle to correct this problem. Specifically, modellers of FGM should distinguish between core theories of decision-making that inform the foundations of their model designs and the models themselves which formalise and extrapolate from these theories to model their implications for the social dynamics of the practice. This distinction has the potential to render more productive adjudication of models and theory in the field, by separating evaluation of the latter, from the issue of idealisations (and other distorting assumptions) in the former.

The *explicit* logic of the model development strategy proposed in Chapter 3 also has advantages for the challenges of *model adjudication* and *progressive modelling* in the field. By associating different model design decisions with specific steps of the modelling strategy, such as using empirical analysis to initially select a core theory, or prioritising empirical calibration according to prior global sensitivity analysis, the strategy offers a framework for the productive critique of the resulting model. It can also facilitate the accumulation of insight through repeated application of the strategy. This is discussed in greater detail in a later section, which addresses limitations of the thesis, and future avenues for research.

The modelling strategy proposed as the conclusion of Chapter 3 is the only model development strategy proposed in the field, and, to the best of my knowledge, one of the few proposed for modelling social dynamics that takes model design uncertainty as its central problem. The closest equivalent is the Pattern Orientated Modelling framework (Grimm et al., 2005; Railsback and Grimm, 2012). However, as discussed in Chapter 7, this approach critically depends on the availability of a sufficient number of relevant macro-level patterns in data, to resolve questions about model adequacy in an instrumental fashion. This is not a situation in which modellers of FGM find themselves. In general, the modelling strategy proposed may prove useful to modellers of social dynamics in a range of fields where uncertainty about model design is high, models may have major impacts on policy, and detailed and relevant macro-level data is relatively sparse.

The strength of contribution provided by the methodological proposals in this thesis doesn't only lie in the various problems in the field for which they offer some resolution. It also lies in the *demonstration*, in Chapters 4-8, that these proposals can be applied successfully and fruitfully to improve modelling standards in the field. The details of the 'local' contributions of this applied work are given in the next section. However, their core 'global' contribution to modelling standards in the field can be stated as follows.

Later chapters of the thesis demonstrate that (following the proposed strategy) it *is* possible to meaningfully investigate a range of uncertain idealisations as part of model development (Chapters 5 & 6), that this information can be used to inform targeted empirical calibration (Chapter 7), resulting in models that are successfully validated against detailed macro-level data (Chapter 7). Furthermore, the developed model can be successfully applied to generate policy-relevant insights, without inappropriately obscuring uncertainty about the model's ability to make accurate novel predictions (Chapter 8).

Previous attempts to model the social dynamics of FGM can certainly not be claimed to have met all (or in some cases, any) of these standards (see below). In fact, in the field of ABM as a whole, detailed empirical validation (let alone in combination with independent empirical calibration, robustness analysis and a policy application that considers uncertainty) is rare (Chattoe-Brown, 2020a). Yet, I have argued throughout this thesis that none of these issues (addressing idealisations, engaging with detailed empirical data etc.) can be neglected if the problem of model design uncertainty is to be taken seriously (which it has to be if models are to be used for policy).

Theoretical, Empirical and Applied Modelling Contributions

Modellers and theorists interested in FGM have adopted distinctive core ideas about decision-making in FGM practicing communities. Some have emphasised the role of coordination incentives arising from social pressures (Efferson et al., 2019; Mackie, 1996, 2017; Novak, 2016; UNICEF, 2007), others have emphasised the role of competition incentives arising from the marriage market (Chesnokova and Vaithianathan, 2010; Ross et al., 2016), still others have considered the idea of downplaying the role of interdependence in decision-making, implying that social influence (where it occurs) may be informational (see Chapter 4). While these theoretical positions are identifiable in the respective works, the distinctions between them, especially between coordination and competition, have not always been clearly identified or articulated as issues of contention in the field. Chesnokova and Vaithianathan (2010) for instance, don't acknowledge that their assumption of marriage *competition* as the primary incentive facing individuals in an FGM marriage market is fundamentally contradictory to the social *coordination* incentives emphasised by proponents of social norm theory.

Chapter 4 takes important steps toward articulating and adjudicating these theoretical positions. In addition to articulating their differences, it provides a first comparative review of the evidence base underlying each. Focusing on the Senegalese context, I found that

available evidence from qualitative research (e.g. Shell-Duncan et al., 2011) favoured the social coordination hypothesis, over a hypothesis of informational influence or social competition. However, I also found that corresponding evidence from quantitative research (e.g. Hayford, 2005; Kandala et al., 2009; Kandala and Shell-Duncan, 2019; Modrek and Liu, 2013) failed to clearly differentiate between the three hypotheses about decision-making. Chapter 4 addressed this gap by developing and implementing an innovative approach, based on statistical modelling of FGM decision-making as a function of different stated beliefs about the practice. The results contribute crucial quantitative evidence that supports the social coordination hypothesis, in favour of alternatives.

Beyond a narrow exploration of the question of preference heterogeneity (and with the recent exception of Efferson et al. 2019), modelling in the field has done little to explore the robustness of established models of FGM to alternative design choices, such as the relaxation of idealisations (like global interaction), or the inclusion of previously omitted features (like norm enforcement). Chapter 5 (the ‘Exploratory Robustness’ stage of the modelling strategy) demonstrated that a wide range of such features ‘matter’ (Chattoe-Brown, 2020b), in that they disrupt the predictions of the social convention model of FGM about intervention outcomes, or about the persistence of the practice (or both). The results clearly establish that the features explored cannot be omitted or idealised away in future models of FGM as a social norm of coordination, and thus mark out a new modelling agenda for the field.

Perhaps the most notable discoveries in Chapter 5 were (a) that claims about the advantages of targeting ‘resistant’ social actors in anti-FGM interventions were significantly disrupted (and even reversed) by the inclusion of a more realistic coalition formation process, and (b) that networked interaction can lead to the persistence of FGM at interior levels (between 0 and 1) when the distribution of preferences in the population is one which would preclude this possibility under assumptions of global interaction. The former calls into question existing claims in the literature about the advantages of targeting different kinds of social actors in interventions (Efferson et al., 2015, 2019) by showing that such claims may be an artefact of idealisations in the way the intervention process is represented. The latter provides yet further evidence that models of FGM based on the social norm of coordination theory, can comfortably accommodate empirical findings regarding interior levels of the practice – without this violating the core tenets of the theory (see further discussion under ‘Critiques of the Social Convention Model’ in Chapter 2).

Efferson et al. (2015) tested the predictions of the social convention model against the distribution of rates of FGM across communities in Sudan. With the exception of this clear attempt at validation (which the model failed, though see Mackie, 2017), direct comparison of the dynamics of models of FGM to detailed macro-level data has been largely absent. Instead, we find vague references to stylised facts (such as that older women tend to defend the practice of FGM, Coyne and Coyne, 2014, or that cut women have better marriage outcome, Chesnokova and Vaithianathan, 2010). Likewise, with the exception of Novak (2016), who attempted to directly calibrate the distribution of preferences in her threshold model of FGM using survey data, the calibration of existing models in the field has been limited to relatively vague references to published research as motivating the broad assumptions of the model (Chesnokova and Vaithianathan, 2010), rather than the specific details of its implementation. Yet, model-based analysis (including that with an aspiration to policy relevance), continues to be published in the field without detailed calibration or validation of the underlying model against hard data (e.g. Efferson et al. 2019). Unfortunately, this situation matches the state of social simulation research in general, where the direct comparison of model outputs to detailed data is the exception, rather than the rule (Chattoe-Brown, 2020a).

Chapter 7 takes important steps toward moving FGM modelling beyond this situation. Most importantly, it presents the first example with direct (and independent) calibration *and* validation of a model of the dynamics of FGM using *detailed empirical data*. This demonstration establishes that a more intensive engagement with data is possible (and indeed, necessary) for the field. Additionally, it offers a performance benchmark that could be compared to future modelling attempts. In achieving this calibration and validation, Chapter 7 develops and applies a number of tools and results that may help to facilitate future model validation and calibration in the field. A novel strategy for estimating community-level preference distributions from DHS data is proposed and implemented (with apparent success, given the successful validation). Moreover, a new macro-level empirical regularity is extracted from DHS data as a model validation test, using the joint distribution of the popularity and prevalence of FGM at the community level. Both the estimation strategy used in calibration and the macro-level empirical data pattern used in validation could be utilised by future modellers of FGM wishing to validate their model's dynamics against secondary data.

As noted in the previous section, one of the methodological contributions of the thesis was the identification and development of Possibilistic Failure Scenario Analysis (PFSA) as a new paradigm for model application in the field. The existing employment of social engineering

approaches in the field is incompatible with a frank acknowledgement of uncertainty about the adequacy of existing models. Instead, the PFSA approach focuses on raising practitioners' awareness of how social dynamics might undermine intervention efforts, without requiring that undue trust is placed in particular models to guide activities. This approach is compatible with reviews of past intervention efforts (Johansen et al., 2013), which clearly identify social dynamics as a complex source of intervention failure that can undermine current activities.

The application of PFSA in Chapter 8 (using the ABM developed in previous chapters), demonstrates the fruitfulness of the approach and makes important contributions to the current understanding of the potential role of social dynamics in FGM interventions. The analysis takes advantage of the rich set of model design features developed and calibrated in the previous chapters to explore new scenarios of intervention 'failure' (no change, negative spillover, and partial change), and new possible antecedents of these scenarios in the social dynamics of communities.

As a result of this richness in model design, in combination with PFSA's focus on exploring a range of *possible* scenarios for intervention failure (rather trying to identify 'robust' or optimal intervention designs, Efferson et al. 2019: 1, 5, 7, 9, 11), Chapter 8 identifies a range of policy-relevant dynamic processes that are either (a) not acknowledged in existing modelling of FGM, or (b) not expressible using existing models. Moreover, it clearly highlights novel instances where interactions *between* processes can create problematic countervailing efforts. Notable discoveries included the following:

It is generally recognised in modelling literature in the field that under a range of circumstances, targeting supporters of FGM may be 'optimal' for generating positive spillovers in the abandonment of the practice. Efferson et al. (2019) recently showed that this result is not always robust. However, analysis in Chapter 8 (via the 'supporters gambit' scenario) goes further. It shows that while targeting supporters can increase the potential for interventions to generate widespread change, it can also precipitate fundamental intervention failure through 'coalition collapse' and 'negative spillovers'.

Similarly, while increases in intervention size are generally represented as increasing intervention effectiveness (UNICEF, 2007)⁷⁶, even if they won't always result in positive spillovers (Efferson et al., 2015), Chapter 8 (via the 'out-of-reach' effect) shows that there

⁷⁶ E.g. to help reach a 'tipping-point'.

can be fundamental downsides to increased intervention size. In some cases, increasing intervention size might greatly reduce efficiency and increase the likelihood of *negative* spillovers⁷⁷.

Both sets of effects can be precipitated by a lack of persuasive strength (i.e. lack of preference change) in the intervention process. They point to the need for policy-makers to carefully consider potential conflicts between different aims. On the one hand, the desire to create widespread change (e.g. through large interventions targeting supporters of FGM), on the other, the need for acceptable levels of efficiency and assurances that *some* change may occur (e.g. through small interventions targeting those already predisposed abandon FGM).

Norm enforcement and differentials of power and authority between social actors have not played a meaningful role in coordination models⁷⁸ of FGM (prior to this thesis), despite being present in the theoretical and empirical literature (see Chapter 5). PFSA in Chapter 8 provides important and novel insights into the ways these features of social interaction might precipitate intervention failure, and interact with other policy-relevant social processes.

Intervention failure might be precipitated not only by a positive association between the approval of FGM and authority, but especially by increases in the dispersal of power across the community, and the shape of the distribution of power - with L-modal distributions potentially placing the bulk of decision-making authority in the hands of small-number of FGM supporters.

In the presence of strong norm enforcement, targeting actors who are willing to abandon FGM could precipitate fundamental intervention failure, because much of the social pressure to continue in the practice arises from norm enforcers who are *not* willing to abandon FGM. This pressure is not relieved by targeting those willing to abandon the practice.

These latter failure scenarios might appear to have an ‘obvious’ solution: that interventions should focus on targeting powerful or FGM norm-enforcing individuals within communities. Yet, previously mentioned dynamics, like the ‘supporters gambit effect’ might undermine such a strategy. Authoritative individuals, and norm enforcers (in these

⁷⁷ I.e. that the level of behavioural change is considerably smaller than the size of the intervention.

⁷⁸ Although authority is briefly discussed as a possible formal feature of coordination models of FGM in Platteau et al. (2017).

scenarios), are also those who support FGM practice. As such, targeting them for an intervention brings with it the possibility of intervention failure through coalition collapse.

Limitations of the Analysis and Future Avenues for Progressive Modelling of FGM

As discussed in detail in Chapter 2, one of the challenges for progressive modelling in the field of FGM, and more widely, is model adjudication. Proposed models of social dynamics, including my own, represent a combination of empirical commitments (e.g. the underlying theory), and subjective 'extra-empirical' commitments (e.g. idealisations, Weisberg, 2012). This makes it hard to compare two 'competing' models, especially when they have different dynamics. These differences may be due to empirical disagreements, diverging extra-empirical design choices, or some combination of two.

In some cases, adjudication is made easier because the new model is a direct response to a previous model. Novak's (2016) model can be understood as a de-idealisation of one of the assumptions of the social convention model. In other cases, a comparison is much more challenging. Models of FGM presented by Novak (2016), Chesnokova and Vaithianathan (2010) and Coyne and Coyne (2014), for instance, all differ from one another in a myriad of ways (c.f. Chapter 2).

Whatever their merits, many of the empirical and extra-empirical commitments involved in the design of my own model can be traced back to specific steps in a model development strategy. This doesn't necessarily mean that the model will be more accurate. Nor does it necessarily make the model design itself easier to compare to others. However, because the model design is the result of applying an explicit modelling strategy, the strategy itself may provide a framework for adjudication. In other words, the explicit logic of the strategy may help other researchers to build on, or productively dispute, the model's design.

In the following sections, I highlight weaknesses in key steps of the development of my model. I also point to the ways in which, following the logic of that methodological 'step', other researchers could productively disagree with or build on, my efforts.

The first step of model development was the identification of a core-theory. This theory would inform the selection of an initial model and act as a conceptual framework to help guide model design. I examined existing empirical evidence and performed my own empirical analysis. I selected social coordination as a core theory.

In doing so, my analysis was subject to at least two kinds of possible limitation. First, there are potential weaknesses in the empirical evidence I used. Second, there are limitations to the theoretical ideas that I initially selected for testing.

My own contribution to the empirical analysis relied on quantitative analysis of survey data. Imperfections in this data (for my purposes) might have led to incorrect conclusions. This includes my use of dichotomous binary variables to measure beliefs, and the measurement of women's 'decisions' in terms of the FGM status of their daughters. My own approach, based on regression and on variable-importance measurement, assumed that the statistical association between beliefs and decisions could be interpreted as supporting a particular theory of decision-making. There is no guarantee that this assumption was correct.

I initially selected three possible 'core' theories for consideration. These were 'social coordination', 'social competition' and 'informational influence'. I selected these because they appear to have distinct implications for modelling social dynamics. Also, they each find some support in the existing literature. However, it might be possible to identify other candidate theories. I interpreted Shell-Duncan and colleagues' research as supporting social coordination in preference to the other two theories. Yet, their research highlighted other aspects of social interaction that are somewhat tangential to social coordination. These included the notion of social capital and the role of social hierarchy. It might be possible to construct a new kind of core theory from their findings, one in which hierarchies of power and social capital eclipse social coordination incentives.

My attempt to evaluate and select an initial core theory in this step could be revisited by others. Other empirical tests of decision-making could be conducted. This could include further secondary analysis of the available survey data. It could also involve further qualitative research in Senegal. If these efforts supported the social coordination account of decision-making, then they would add credibility to the theoretical motivations of the ABM model. If they supported some other account, then this would be a highly productive way of adjudicating the theoretical foundations of the model. It could also provide a clear justification for exploring a different class of models, based on a different 'core' theory.

The second step of the model development strategy was crucial in determining the 'extra-empirical' commitments of the model. Using existing literature on FGM, I selected a range of possible elaborations and de-idealizations of a 'standard' coordination model. These selections determined the model design elements that I was able to consider in the analysis.

They also determined the possible design elements that I did *not* consider. Thus, the idealisations and omissions present in my own model were largely 'set' at this point.

There are many other possible model features that could be explored. One possibility that comes to mind is social hierarchy. Shell-Duncan et al. (2011) and others (e.g. Nam, 2018) have emphasized the role of social hierarchies in FGM practicing communities. My treatment of hierarchy (in terms of influence weighting) was relatively shallow. More expansive explorations of social hierarchy as a part of coordination-dynamics could reveal important new model design possibilities.

Another area that is unexplored in my analysis (and other formal models of FGM) is community deliberation. Empirical accounts of anti-FGM interventions refer to community meetings and other kinds of organized deliberation as important parts of the social change process (e.g. Diop et al., 2008). Yet these are abstracted-away in my model. It might be that explicit representations of community deliberation could have important implications for social dynamics.

In general, any new model feature which is shown to be credible (e.g. is supported by existing research) and important (is demonstrated to impact the dynamics of existing models) should be added to the 'agenda' for models of the dynamics of FGM.

As such, following the logic of this step, the avenue for others to object to the design of the model, or build upon it, is clear. Future researchers can explore other features of coordination dynamics that were not considered in my analysis. If these can be shown to alter the key dynamics of the model, then this would represent important progress for the field. It would suggest that existing models need to be further expanded to take account of yet more features. In repeatedly applying such an analysis, it is possible that the field could progress toward a comprehensive understanding of the 'space' of possible coordination models of FGM. This would be a very significant step in managing the uncertainty present in such models.

Another crucial avenue for model adjudication is presented in Steps 6 and 7. Suppose that an observer is satisfied with the core theory I selected. Suppose that they are also convinced by the set of model features that I selected for analysis. In addition, let's say they also agree with the way I combined these into a general model. There still remains considerable scope for them to disagree with, or build upon, the calibration and validation of the model.

My approach to empirical calibration of the parameters of the ABM model depended on a number of relatively arbitrary (although not implausible) assumptions. Other researchers may be able to provide evidence to support or undermine these assumptions. They also might find alternative approaches to calibrating the same parameters. The results of these alternative approaches may be consistent with my own findings, or they may conflict with them. Researchers may also find other ways to calibrate new parts of the model, including through the use of other data sources. All such efforts have the potential to improve the model, or productively reassess its merit.

Similarly, empirical validation depended on assumptions about the available survey data. These assumptions might be supported or undermined by further analysis. This, in turn, will enhance or undermine the credibility conveyed to the model by the validation test. I employed three kinds of empirical validation. These were correlational analysis (comparing real and predicted rates of FGM in communities), and two empirical 'patterns' which the model partially reproduced. There is considerable potential for the strength of this empirical validation test to be expanded if new and relevant empirical patterns can be identified. If the model also reproduces these new patterns, then this will further strengthen its credibility. If it fails to do so, then this presents a clear challenge for further model improvement.

There is a core to the argument that I am attempting to convey here. I am not claiming that following my modelling strategy will result in a model that is necessarily 'correct'. Nor do I claim that such a model will necessarily be more accurate than one which does not follow it⁷⁹. However, the modelling strategy I've proposed *does* offer a framework for addressing both the empirical and extra-empirical aspects of model design. If other researchers are willing to accept this framework, then it may be possible to locate differences between models in a particular step of the modelling strategy and to adjudicate these differences according to the logic of that step.

All of this depends, of course, on other modellers being convinced by the logic of the strategy itself. The strategy I have proposed emphasizes the need for models to identify 'important features' of their real-world targets (Steps 1-4), and to match those features in their design (Step 5) (see Weisberg, 2012). Yet, a more 'instrumentalist' framework might treat these kinds of efforts as subordinate to the goal of empirical validation. Under an instrumentalist

⁷⁹ Although, at least, following the strategy codifies a process through which I have attempted to make the model more accurate (i.e. targeted empirical calibration), and to test this accuracy (i.e. empirical validation)

framework, the task facing modellers would be framed primarily in terms of the successful reproduction of validation data. As discussed above and in Chapter 7, the popular 'Pattern Orientated Modelling' (POM) framework has elements of this perspective (Grimm et al., 2005; Railsback and Grimm, 2012). The authors do acknowledge the need for the model to be 'structurally realistic' (Railsback and Grimm, 2012). But, they also allow significant aspects of model development to be guided by a process of fitting model structures and parameters to available data patterns. Under POM, model adjudication depends primarily on comparing models' ability to reproduce a set of empirical patterns (see discussion of Heine et al., 2005 in Railsback and Grimm, 2012: 239)

The POM framework is deliberately built on scientific principles of rigorous empirical testing. It has also seen considerable success, especially in the field of ecology. Yet, as I have argued previously, it depends on the assumption that available data is sufficiently rich, and model uncertainty sufficiently small, that aggregate 'patterns' in the available data will be sufficient to discriminate between 'realistic' and 'unrealistic' models. When modelling the social dynamics of FGM, these assumptions may not be credible. The availability of detailed empirical patterns is limited. Moreover, there is substantial uncertainty about model design. As such, there is real potential for a 'bad' model to reproduce existing data on FGM. In a scenario such as this, a strategy like the one I have proposed may be preferable. The strategy is heavily focused on managing uncertainty about model design *before and after* empirical validation.

On the other hand, there is an ongoing project, supported by the Population Council, which is collecting empirical data directly related to the social dynamics of FGM⁸⁰. One can only speculate, but if sufficient empirical data accumulates through projects like this one, then perhaps a POM approach (or something similar) which emphasizes a process of model fitting, would become a viable approach to studying the dynamics of FGM.

⁸⁰ <https://www.popcouncil.org/research/evidence-to-end-fgm-c-research-to-help-girls-and-women-thrive1>

Final Thought: A Model-Centric Perspective on Social Norm Theory as a Theory of Social Dynamics

‘Social Norm Theory’, as a perspective on both decision-making and social dynamics, has been widely applied to problems involving widespread harmful practices. This includes FGM, but also practices such as child-marriage, open-defecation, binge-drinking, corporal punishment, and harassment (Mackie et al. 2015). As such, insights into the status of this theory may have broad implications.

In this thesis, I have defended social norm theory’s claims about decision-making (i.e. that decisions are interdependent and subject to coordination incentives) and emphasised its usefulness as *part of* the construction of a model of social dynamics. In discussing model design uncertainty, I have focused on the challenges facing modellers themselves. My final point is to emphasise the connection between this issue and the challenges facing those drawing on social norm theory more generally.

Modelling in this thesis has extensively demonstrated that there is not a robust connection between social norm theory’s claims about *decision-making* and any *particular* claims about the outcomes of social dynamics. Tipping-points might exist, or might not. Local universality might be expected, or it might not. And so on.

If we want insight into the social dynamics of a particular harmful social practice, we will need a formal model to help us *extrapolate* from our ideas about decision-making. In doing so, we will face all of the kinds of uncertainty discussed in this thesis.

There has been considerable interest in the development literature in ‘diagnosing’ whether a practice is a social norm (Mackie et al., 2015; Bicchieri et al., 2014). Yet, modelling shows us that after identifying a practice as a ‘social norm’, the lion-share of our work to understand its *dynamics* remains unfinished. We will have to contend with the many challenges associated with building and applying a *formal model* of those dynamics.

In this thesis, I have tried to give an account of those challenges and to take a step toward meeting them.

Appendix A: Chapter 2 Appendices

Appendix A1: The Critical Mass in UNICEF (2007)

In addition to the ‘tipping-point’ dynamic, the original social convention model, as described in UNICEF (2007), also includes another feature called the ‘critical mass’. If readers examine Figure 1 in Chapter 2, then they will see that the expected utility for FGM abandonment *and* FGM practice, slopes up as a function of the proportion of others abandoning FGM. This occurs because the model assumes a (small) *miscoordination* incentive for FGM practice, based on cut women’s potential advantages in the marriage market. This means that there is a level of FGM abandonment, *prior to the tipping point*, at which the utility that actors expect for abandoning FGM is greater than the utility that they would expect if *everyone* practiced it. UNICEF (2007) claim that this might be a temporarily stable point in the abandonment process since (acting collectively) actors would do better by reaching this point than if they all practiced FGM. This point is called the ‘critical mass’. UNICEF (2007) claimed that at this point, actors will be incentivised to recruit others to abandon in order to reach the final, stable, tipping-point.

However, there are a number of problems with treating this as a formal implication (as opposed to a kind of ‘bolt-on’ narrative) of the social convention model.

- It violates the basic principle of utility maximisation in the model. Although, at the ‘critical mass’ point, actors would do better by abandoning than if they all practiced FGM, they would still, *at this point* do better individually by practicing FGM.
- The ‘recruitment’ process is not formally specified; it’s just an (optimistic) narrative about how actors might persuade others to abandon FGM *despite* those recruited actors *also* still preferring to practice FGM.

There is also an empirical problem with the logic of the ‘critical mass’. It only exists separately from the tipping-point if we assume that there is a miscoordination incentive favouring FGM practice (e.g. because of marriage competition). If there isn’t, as I show in Chapter 4, then the ‘tipping-point’ and the ‘critical-mass’ point are the *same* point, rendering the latter redundant.

That isn’t to say that the idea of actors recruiting others, or making *conditional* initial commitments to abandon FGM *couldn’t* be formalised in a coherent way. In fact, I introduce formalisations of both these elements to a formal model described in Chapter 5 (see

‘coalition formation’ and ‘organised diffusion’). However, I do so in a way which maintains the assumption that actors will only abandon FGM *if* they actually prefer to do so.

Therefore, as it stands, the critical mass idea is better treated as a narrative about social dynamics which is *associated with* the formal convention model, rather than a demonstrable implication of that model for our understanding of social dynamics.

Appendix A2: Coyne and Coyne’s Identity Economics Model of FGM

Coyne and Coyne (2014) presented a game-theoretic model of FGM which is quite distinctive in its design relative to either the original social convention model (Mackie, 1996) or Chesnokova and Vaithianathan’s (2010) marriage competition model. Coyne and Coyne’s (2014) model focused on the concept of identity - specifically, the costs and benefits, in terms of identity, that social actors may derive from practicing FGM, or from the non-practice of FGM by other actors (which may represent a threat to their own identity).

Coyne and Coyne’s (2014) model consists of two actors engaged in a sequential game. I will refer to these actors as Player A and Player B. Player A represents an uncut women (or her family) who must decide whether to practice FGM or not (be subjected to FGM), and in the analysis Player A is treated as viewing FGM as having an intrinsic cost (i.e. they would prefer not to practice FGM). Player B represents a cut woman (or her family) who must decide whether or not to *punish* Player A *if* Player A decides not to practice FGM. An annotated version of the resulting game is shown in Figure A2.1.

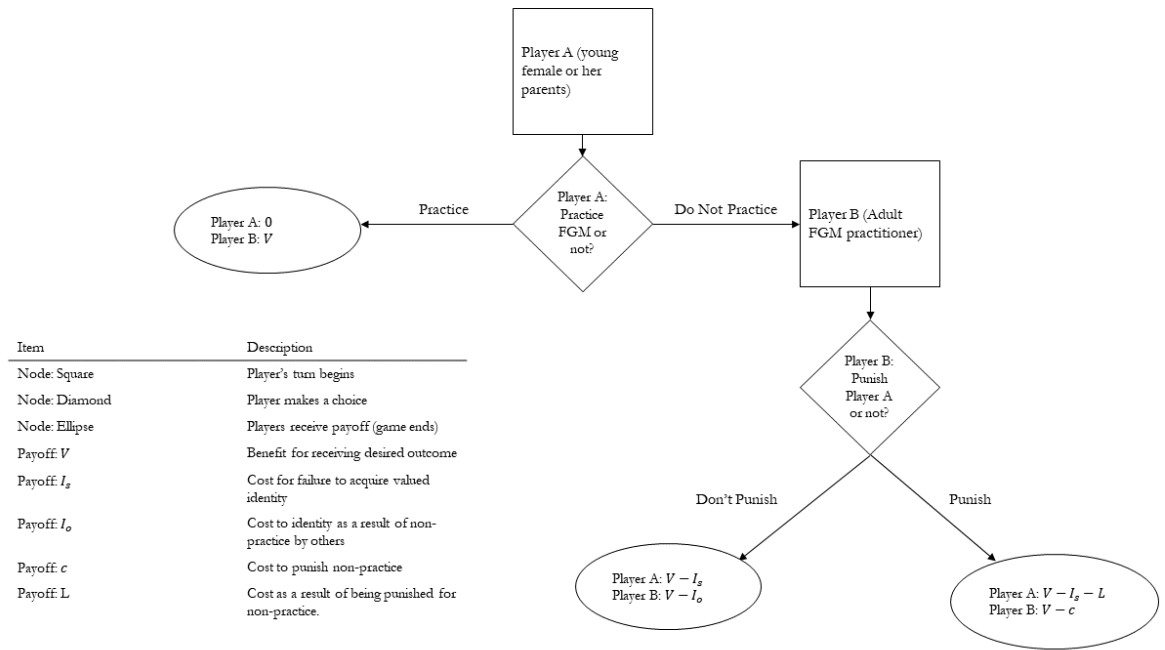


Figure A2.1: An Annotated Version of the FGM Identity Game (Coyne & Coyne, 2014)

In terms of the relevant pay-offs, Player A would prefer not to practice FGM. If they abstain from the practice, they will receive a gross benefit of V . Otherwise, they receive a pay-off of 0. If they choose not to practice FGM, then their net pay-off depends on what Player B decides to do. If Player B decides to punish them for not practising FGM, then they receive a pay-off of $V - I_s - L$. This represents the benefit they get for not practising FGM (assumed to be their preference: V), the passive social cost they face for not being able to take on the identity associated with FGM practice ($-I_s$, since FGM considered a high-status activity in practising communities) and social costs imposed on them by Player B ($-L$). If Player B doesn't punish them, then Player A doesn't pay the cost $-L$.

Player B's net utility also depends on whether they decide to punish Player A. If they do not punish Player A then they will pay a cost of $-I_o$, which represents the cost to their valued identity (as an FGM practitioner) being undermined by Player A's refusal to be cut. If they do punish Player A then, through ostracising them, they are able to avoid the damage to their identity; however, they pay a cost of c (e.g. in terms of the loss of the relationship with Player A, or because punishing others is inherently costly) to do so. Player B always receives a gross payoff of V because they practice FGM and this is their *intrinsically* preferred choice (i.e. they are a true FGM proponent).

As will be clear, these choice (at-least under some parametrizations) are interdependent. Player A's decision to practice FGM may depend on whether they expect to be punished for non-practice by Player B. This, in turn, depends on the relative costs and benefits to Player B for punishing Player A.

Coyne and Coyne (2014) argue that there is evidence from ethnography that identity plays a role in decision-making around FGM. Actors view FGM as part of their identity, view it as a social requirement and may look unfavourably on those who fail to engage in the practice. Coyne and Coyne (2014) also refer to nationally representative survey data that shows that 'tradition' is a much more commonly cited advantage of FGM by practitioners than 'marriage' or 'virginity'.

Coyne and Coyne (2014) are relatively conciliatory in their style of writing, describing their emphasis on identity as *complementary* to the focus on marriage (and marriage related) issues in other models. However, their empirical motivation could be considered an implicit empirical claim that (at least in some contexts) marriage is not an important part of social actors' decision-making. This perspective would seem to represent a substantive empirical disagreement with Chesnokova and Vaithianathan (2010), for example, for whom the notion of marriage competition is indispensable. However, it is far from clear either that (a) this empirical disagreement can adjudicate some of the more technical differences in model design between Coyne and Coyne (2014) and Chesnokova and Vaithianathan (2010), or (b) that the de-emphasis of marriage really constitutes a substantive empirical disagreement with social norm of coordination theory.

Chesnokova and Vaithianathan (2010)'s treatment of marriage as central to decision-making about FGM, versus Coyne and Coyne (2014)'s omission of marriage as a component of decision-making, in favour of identity concerns, may represent a substantive disagreement for which some empirical adjudication is attempted. However, some of the more particular divergences of design between the two models have no such basis. Chesnokova and Vaithianathan (2010) employ an *n-person* game involving actors distinguished by wealth, gender and FGM status, who are randomly and stochastically matched in a sequential game. Coyne and Coyne (2014) employ a two-person game involving actors distinguished by age (and initial FGM status), who engage in a single-shot sequential game. To a large extent, these differences are simply *alternative idealizations* of the relevant social processes. Unfortunately, there is no available analysis of how (and which of) these divergent design features may impact the insights into social dynamics gained from the two models.

In the initial exposition of the convention model of FGM (Mackie, 1996), marriage is a key motivation for assumptions about (coordinated) social pay-offs. However, this assumption has been subsequently relaxed, both by the author (Mackie and LeJeune, 2009) and in extensions of the model (Efferson et al., 2015; Novak, 2016; Platteau et al., 2017). Thus, the contemporary interpretation of the social convention model treats the coordination incentives associated with FGM as deriving from either marriage-related concerns, or general normative social pressures to continue with the practice of FGM (including those related to identity), or both (Mackie and LeJeune, 2009). As such, the empirical research cited by Coyne and Coyne (2014) in favour of an identity model, is largely compatible with the modern incarnations of social norm theory, and associated coordination models. Therefore, no clear or substantive empirical adjudication of the designs of the former and the latter is provided by Coyne and Coyne (2014)'s analysis of marriage as a motivating factor (see Chapter 4 for my critique of their approach to this question).

Instead, once again, many of the differences between the social convention model and the identity model can be accounted for in terms of the preferences of the *modellers* (i.e. for an 'identity economics' framework) and divergent choices about idealisations (e.g. Mackie abstracts away the question of *who* enforces FGM and *why* they do so, Coyne and Coyne (2014) abstract away the notion of an *n-person* population under continuous interaction, in favour of two 'representative' agents in a one-shot repeated game).

As in the comparison of Chesnokova and Vaithianathan (2010) with the social convention model, the comparison of Coyne and Coyne (2014) with the social convention model reveals divergent predictions about social dynamics and divergent recommendation for policy. Here, the most straightforward comparison is between the identity game and the matrix-form convention model, since both are two-player games whose dynamics are defined in terms of their Nash equilibria. Recall that under the matrix form coordination game, there are two equilibria: the inferior equilibrium (where both players practice FGM) and the superior equilibrium (where neither player does). The policy implication arising from the game is that interventions should act as coordination devices to assist players in moving from the inferior to the superior equilibrium (which both players prefer).

The identity game allows has four possible Nash equilibria (two of which involve interdependence):

6. Successful Deterrence Equilibrium: I_o is greater than c , so Player B prefers to punish non-practice. Also, $V < I_s + L$ so Player A is deterred by this threat of punishment and practices FGM:

“This equilibrium describes those situations where FGM persists because of fear of severe responses such as significant losses of family honor, violent harassment, and extreme forms of ostracism.” (p.145)

2. Failed Deterrence Equilibrium: I_o is greater than c , so Player B prefers to punish non-practice. However, $V > I_s + L$ so Player A is *not* deterred by this threat of punishment and refuses to practice FGM:

"This equilibrium indicates that movement away from FGM is possible, but high cost. The shift away from FGM will occur only when Player 2's utility from refraining from participating in FGM is relatively high when compared to the associated cost. (p.145)

3. Unpunished Abandonment Equilibrium: I_o is less than c , so Player B prefers not to punish Player A even if Player A doesn't participate in FGM. Knowing this, Player A prefers to abandon FGM because $V > I_s$.

“This equilibrium represents a situation where the movement away from FGM is most likely.” (p.145)

4. Unconditional Abandonment Equilibrium: $V < I_s$ so Player A continues to practice FGM regardless of what Player B might have done had they abandoned it, in order to avoid the loss of the valued identity.

“Under this scenario, the practice of FGM will continue, since it is deeply intertwined with the identity of community members.” (p.145)

The policy implications that can be derived from the identity game model depend on the way in which equilibria involving cutting (1 and 4) can be transformed into equilibria involving not cutting (2 and 3). Coyne and Coyne (2014) note that key factors in such a transition are the size of the identity costs associated with not practising FGM (I_o and I_s) and the threat of deterrence for non-practitioners (L). According to the model, if FGM can be un-coupled from the valued identities of community members (lowering I_o and I_s) then the net benefit from abandoning FGM increases and the incentive for practising adults to punish non-practitioners declines. Furthermore, if the strength of the deterrence declines, then the incentive to abandon FGM, even in the face of potential ostracization, increases.

Coyne and Coyne, (2014: 147) argue that, in particular, their model motivates interventions which maintain the socio-cultural rituals associated with FGM whilst abandoning the physical aspect of the ceremony (these are sometimes called Alternative Rites of Passage – ARP, cf Droy et al., 2018). They also argue that “marginal” changes in FGM practice are likely to be more effective than “wholesale” changes since the former is likely to limit the associated loss of identity (p.147).

Yet, despite Coyne and Coyne (2014)'s insistence that their model is complementary to existing analyses, in many ways these dynamics and their implications for policy are contrary to those of the social convention model. Under the social convention model, complete "wholesale" abandonment of the practice through coordinated action is expected to be the most effective (and possibly the only) way to ensure stable social change. Instead, under Coyne and Coyne (2014)'s model, *incremental* changes in the identity status of the practice are expected to be the most effective means of reaching a non-practice equilibrium.

In other respects, rather than being contradictory, Coyne and Coyne (2014)'s model is simply incommensurable with the social convention model. The identity game is not an *n-person* game, and it is not clear how it could be extended to become one. As such, it makes no prediction about the stability of different rates of FGM within practicing communities, nor about the potential existence of beneficial tipping-points which may be reached through intervention efforts.

Reconciling these two models would represent a considerable intellectual challenge. Evidence for the association of FGM with valued identities is broadly compatible with a coordination model based on normative social pressures, or with a model of the kind designed by Coyne and Coyne (2014). Some degree of empirical adjudication of the design differences between the two models may be possible in principle. However, it would *still* be necessary to address the different idealizations present in both models. These do not necessarily rest on empirical disagreements per-se, but on divergent (implicit) commitments regarding the importance of different aspects of a shared empirical picture.

Appendix A3: The Diversity of Formal Models of FGM

This appendix provides a high-level summary of the range of formal models of FGM which have been developed. This summary is captured in Table A3.1, which describes notable features of the structure, decision-processes, dynamics and potential derived policy recommendation from these models.

Table A3.1: Formal Models of the Social Dynamics of FGM

<i>Model N</i>	<i>Study/Model</i>	<i>Model Structure</i>	<i>Decision Contingencies</i>	<i>Dynamics</i>	<i>Potential Derived Policy Implications</i>
#1	Mackie (1996) #1	Matrix-form game with two representative actors acting simultaneously	Cost of FGM (homogenous), Miscoordination Costs (Social/Marriage - homogenous)	Two equilibria: Both practice, both abandon	Recommends coordinated action to exit the inferior equilibrium of both practicing FGM
#2	Mackie (1996) #2	The N-person sequential game represented in a Schelling Coordination Diagram	Cost of FGM (homogenous), Miscoordination Costs (Social/Marriage - homogenous)	Exterior stability (0% or 100%), Interior tipping point	Recommends coordinated action to reach tipping-point (preference change as a useful adjunct)
#3	UNICEF (2007) #1	Matrix-form game with two representative actors acting simultaneously	Cost of FGM (homogenous), Unilateral Abandonment Costs (Social/Marriage - homogenous), Unilateral Practice Incentives (Marriage - homogenous)	Two equilibria: Both practice, both abandon	Recommends coordinated action to exit the inferior equilibrium of both practising FGM (preference change as a useful adjunct)
#4	UNICEF (2007) #2	The N-person sequential game represented in a Schelling Coordination Diagram	“ “	Exterior stability (0% or 100%), Interior tipping point	Recommends coordinated action to reach tipping-point (preference change as a useful adjunct)
#5	Chesnokova & Vaithianathan (2010)	Sequential game with a stochastic pairing of actors and mixed-strategy equilibria	Expected return on the FGM investment in the marriage market (in terms of findings rich husband)	Interior stability (below 100%), Exterior stability (at 0%), Interior tipping point	Recommends reducing the rate of FGM below a tipping-point for stable abandonment
#6	Wagner (2011)	Utility function only (implicit dynamics)	Reputation and prestige gained through FGM (conditional on practice in the community), health costs of practicing FGM	Exogenous	Potential limited relevance of health problem information, which is a low cost compared to socio-cultural gains
#7	Ouedraogo & Koissy-Kpein (2014)	Utility function, with parameterized labour-markets and marriage-markets	Marginal gain from investment in FGM (cost of FGM, marriage return from FGM) versus marginal gain from Education (cost of school, labour market return from schooling).	Exogenous	Lower the costs of schooling for girls, improve women's opportunities in the labour market, reduce parental dependence of the financial returns from daughters.
#8	Coyne and Coyne (2014)	Sequential game with two representative agents	Value of preferred outcome (practice FGM/not practice FGM - homogenous), cost to self-identity from non-practice (homogenous), cost to others' identity from non-	Four equilibria (depending on the parameterization of utilities): unconditional practice of FGM, threat-induced practice of FGM, unpunished	Reduce connection of FGM to valued identity (e.g. through alternative rites), increase costs to punish (e.g. through cultural change),

			practice (homogenous), cost to punish (homogenous), cost of punishment (homogenous)	abandonment of FGM, punished abandonment of FGM.	decrease costs of punishment (e.g. through exit options).
#9	Efferson et al. (2015 - Supplement)	Evolutionary (replicator dynamic) game with a stochastic pairing of families	Miscoordination cost for practising and non-practising interactants (homogenous), cost of FGM (homogenous)	Exterior stability (0% and 100%).	NA – critique of Mackie (1996)
#10	Novak et al. (2016)	Threshold model of collective behaviour (see Granovetter, 1978)	Cost of FGM (heterogenous), Social coordination incentives	Multiple internal equilibria, interior tipping-point not guaranteed	It may be inappropriate to focus on collective, action (village-level) interventions under certain threshold distributions; in these cases, a focus on changing individual preferences may be preferred
#11	Ross (2016)	Population equilibrium states estimated by analysis of stochastic utility functions incorporating information about the distribution of characteristics in the population	Costs of practicing FGM (homogenous), rank-related increase in marriage value for cut women (depends on the distribution of value of spouses), social coordination incentives	All rates of FGM are potentially stable.	Coordination focused approaches less likely to be effective if marriage competition based on heterogeneity in the value of male spouses is high
#12	Efferson (2019) # 1	Threshold model of collective behaviour (see Granovetter, 1978)	Costs and benefits as exogenous, actors have a threshold at which they will practice FGM.	Multiple internal equilibria, interior tipping-point not guaranteed.	The effectiveness of coordination efforts will depend on the distribution of preferences, targeting actors most willing to abandon FGM may be less effective, relative to targeting supporters
#13	Efferson (2019) # 2	Threshold model of collective behaviour (see Granovetter, 1978) with <i>networked interaction topology</i>	Costs and benefits as exogenous, actors have a threshold at which they will practice FGM.	Multiple internal equilibria, interior tipping-point not guaranteed.	Homophilous social networks can reduce the potential for positive spillovers from interventions, especially when interventions target supporters of FGM.
#14	Efferson (2019) # 3	A dynamic stochastic model of frequency-dependent social behaviour	Costs and benefits as exogenous, actors practice FGM with a probability which is a function of other choices in an in-group and out-group.	Depending on the parameterization, rates are stable when matched between in-group out-group, or for polarized equilibria, with practice high in one group and low in another.	The increasing tendency for out-group reactivity (distinguishing in-group from out-group) can impair the possibility of positive spillovers through in-group conformity dynamics
#15	Platteau et al. (2017)	Threshold model of collective behaviour (see Granovetter, 1978)	Cost of FGM (heterogenous), Social miscoordination costs	Multiple internal equilibria, interior tipping-point not guaranteed	See Efferson (2019) #1

There has been considerable variation in the design of models employed in formal analyses of the dynamics of FGM. Analysts have used two-player simultaneous games (Models: 1 and

3_ and two-player sequential games (Model: 8). More commonly *n-person* models have been used (Models: 2, 4, 5, 9, 10, 11, 12, 13, 14 and 15).

Mackie (1996) and UNICEF (2007) used the notion of a repeated sequential game (Models: 2 and 4). Chesnokova and Vaithianathan (2010) employed a two-round sequential game with simultaneous decision-making within each round (Model: 5). Coordination models employed in Efferson et al. (2019), Novak (2016) and Platteau et al. (2017) implicitly use a logic of repeated decision-making (Models: 9, 10, 12, 13 and 14).

The pattern of interactions between actors assumed by the models has also varied considerably. In the two-player games, a single pair of actors interact with one-another (Models: 1, 3 and 8). In the *n-person* models, some analysts have assumed that actors interact in pairs as a part of a stochastic matching process (Models: 9 and 5). Others have allowed global interaction (all actors interact with all other actors, models: 2, 4, 10, 11, 12, 14 and 15). At least in one case, interactions have followed a networked structure (actors interact with network neighbours, Model: 13).

Representations of the decision-making processes of social actors have varied in terms of both the assumed sophistication of actor's decision-making processes and the relevant contingencies of those decisions. In many cases actors respond to whether (and how many) actors are expected to practice FGM in the next round of the game (Models: 1, 2, 3, 4, 10, 12, 13, 14 and 15). This is generally assumed to be based on other actors' decisions in the previous round. As such, actors choose the utility-maximizing strategy for the previous round.

Under other models, actors' decision-making has been assumed to have various levels of sophistication. Under Chesnokova and Vaithianathan (2010)'s model, for instance, actors choose a maximising strategy (which is a mixed-strategy) which is based on their knowledge of the rationality of other actors and an understanding of the probabilistic structure of the matching process (Model: 5). Under Coyne and Coyne (2014), actors are implicitly assumed to know one another's pay-off structure, with Player A's decisions depending on their anticipation of the rational response of Player B to their choice (under some circumstances) (Model: 8). Ouedraogo and Koissy-Kpein (2012) assume that parents have sophisticated fore-knowledge of the relative economic returns from educating (labour market returns) or cutting their daughters (marriage market returns) and that they maximise utility with respect to their expected returns (Model: 7). Ross et al. (2016) assume that actor's decisions are responsive to the expected value of choices about FGM that depend on the outcomes of

stochastic processes and the distribution of different features in the population (Model: 11). Under Efferson et al. (2019)'s frequency-dependent social influence model, utility maximization is implicit (e.g. actors simply have an increased probability of practicing FGM as others in their 'in-group' do so, Model: 14). Under Efferson et al.'s (2015 – Supplement) evolutionary model, actors do not evaluate different choices per-se; instead successful strategies increase in relative frequency within the population over repeated rounds (Model: 9).

Alongside these different forms of decision-model are varied assumptions about the costs and benefits associated with decision-making, and the relative uniformity (or not) of these pay-offs. A number of analysts have assumed that actors face social costs for failing to practice FGM when others do so (Models: 1-5, 10, 11, 12, 15). Less commonly, analysts have assumed that these costs are symmetric, i.e. actors *also* face social costs for practising FGM when others have abandoned it (Models: 9, 10, 15). Some models have assumed that actors can face miscoordination *incentives* to practice FGM when others *don't* do so (Models: 4, 5, 11). This can result, for example, from gaining a competitive advantage in the marriage market. In some cases, these benefits are a simple function of other actors' choices (Model: 4). In others, they depend on the distribution of marriage opportunities in the population (Models: 5 and 11), and the consequent likelihood of gaining an advantageous marriage from practising FGM.

Across existing models, FGM has often been associated with some form of intrinsic value. Typically, it has been modelled as a *cost* to health and/or human rights for those individuals who practice it (Models: 1-7, 9, 10, 15). Coyne and Coyne (2014) offer a different perspective; however, treating FGM practice as either an intrinsic good or an intrinsic cost, depending on the subjective view of the actor (Model: 8). In most cases, the *intrinsic* value of FGM is treated as homogeneous across the population (Models: 1-5, 8, 9, 11). However, notably, it has been allowed to vary (either explicitly or implicitly) under some coordination models with heterogeneous thresholds (Models: 10, 12, 13 and 15). As discussed in Chapter 2, this distinction is known to have substantive implications for dynamics.

Unsurprisingly, the variety of model designs is associated with a number of different predicted social dynamics and/or derived policy recommendation. As discussed in Chapter 2, *n-person* social convention models (Models: 2 and 4) with homogeneous preferences exhibit a dynamic of exterior stability (rates of FGM at 0% or 100%), interior *instability*, and a single interior tipping-point (rates above or below this point being driven to universal practice or

abandonment). These dynamics, in turn, imply the value of coordinated action to reach a tipping point, after which social forces will drive the population to complete abandonment.

Chesnokova and Vaithianathan (2010)'s model, by contrast, had interior stability and did not (generally) have stability at universal practice (Model: 5). However, the model *did* exhibit an interior tipping-point, below which the practice was expected to decline through social processes. Efferson et al. (2015 - Supplement)'s evolutionary model also exhibited only exterior stability, although tipping-points were not analyzed (Model: 9). Some models have indicated the potential for stability at any level of FGM practice, with the possibilities for advantageous tipping-points depending on the distribution of characteristics in the population (Models: 10, 11, 12, 13, 14 and 15). In these cases, the recommendation that interventions focus on acting as coordination devices may be called into question, in favour of a greater focus on changing individual preferences. Modellers have also highlighted various other avenues for intervention, including the substitution of social benefits otherwise accrued through FGM (Model: 6), reducing schooling costs and labour-market discrimination for women (Model: 7), and the *incremental* de-coupling of FGM from its association with identity (Model: 8).

Appendix A4: Survey of ‘Validation’ of Formal Models of FGM

A summary of studies involving the validation of formal models of the social dynamics of FGM is presented in Table A4.1. Validation is broadly defined here, as any attempt to deploy empirical research to confirm or dispute the adequacy of a particular model. The first column identifies the model and the study in which the model is explicated. The second column shows the study (or studies) in which some attempt at validation occurred, which is not necessarily the study in which the model is first outlined. The third column summarizes the outcomes and form of this validation. A ‘+’ or ‘–’ symbol is used to indicate whether the findings have broadly been interpreted as supporting or undermining the model (with ‘+/-’ indicating mixed support. I do not cover *all* formal model of the dynamics of FGM; instead, I focus on models for which some form of explicit empirical validation has been attempted.

Table A4.1: Validation of Formal Models of FGM

ORIGINAL MODEL	STUD(IES) PERFORMING VALIDATION	SUMMARY OF VALIDATION
SOCIAL CONVENTION MODEL (MACKIE, 1996; UNICEF, 2007)	Mackie (1996)	<p><u>Validation by analogy with foot binding (+):</u> Mackie argues that there are strong analogies between foot binding in China and Infibulation (Type 3 FGM) in sub-Saharan Africa. He likens the process of rapid abandonment in China to the dynamics of the n-person coordination model, arguing by analogy that infibulation will follow a similar outcome</p> <p><u>Validation by reference to ethnography (+):</u> Mackie argues that a range of ethnographic research highlights social and marriage pressures to practice FGM (Abdalla, 1982), in particular, as a response to marriageability pressures (Boddy, 1982; Gruenbaum, 1982)– validating the focus of the decision model on coordination incentives</p> <p><u>Validation by reference to empirical regularities (+):</u> Mackie argues that FGM is <i>locally universal</i> (within a locality everyone practices FGM or no-one does) and <i>attitude-behaviour discrepant</i> (support for FGM being lower than actual levels of practice). Both of these are consistent with the dynamics of the model, which only has exterior points of stability, and which represents decisions as being driven by social incentives which are independent of attitudes.</p>
SOCIAL CONVENTION MODEL (MACKIE, 1996; UNICEF, 2007)	Mackie (2000)	<p><u>Validation by reference to policy outcomes (+):</u> Mackie likens the Tostan interventions, which resulted in collective declarations of abandonment of FGM by certain members of a set of villages, to the arrangement of a convention shift, of the kind recommended by the convention model, with the alleged success of these declarations treated as highly consistent with of the predictions of the model (p.253, pp.256-257, p.279).</p>
SOCIAL CONVENTION MODEL (MACKIE, 1996; UNICEF, 2007)	Multiple statistical analyses of survey data (Hayford, 2005; Hayford and Trinitapoli, 2011; Kandala et al., 2009; Kandala and Shell-Duncan, 2019; Modrek and Liu, 2013) cited in Mackie (2017)	<p><u>Validation by statistical modelling (+):</u> A variety of statistical studies converge on the finding that, holding constant a range of other correlates of parents decisions to practice FGM, regional and community factors, such as the frequency of the practice within the local social group, are significantly associated with the prevalence of the practice – potentially supporting the focus of the convention model on interdependent decision-making.</p>

SOCIAL CONVENTION MODEL (MACKIE, 1996; UNICEF, 2007)	UNICEF (2013)	<p><u>Validation by aggregate regularities (+):</u> Without relying on sophisticated statistical modelling, a review of a large number of national surveys related to FGM (DHS and MICS) by UNICEF, finds strong geographic and ethnic non-uniformity in the prevalence of the practice – consistent with the ‘relational’ assumptions of the convention model. They also find that references to social acceptance as an advantage of FGM are very common, supporting the representation of social incentives in the decision-model. They also find widespread evidence of attitude-behaviour discrepancies – with prevalence consistently higher than support for the practice.</p>
SOCIAL CONVENTION MODEL (MACKIE, 1996; UNICEF, 2007)	Mackie (2017)	<p><u>Validation by reference to ethnography (+):</u> In defence of the model and surrounding theory, and in addition to citing much of the research noted above, Mackie highlights qualitative research and field reports indicating social and marriage pressures to practice FGM (Almroth et al., 2001; Boddy, 2007; El Dareer, 1982; Gruenbaum, 2001; Hemlund and Shell-Duncan, 2007; Hicks, 1996; Lightfoot-Klein, 1989).</p> <p><u>Validation by statistical regularities (+):</u> Mackie argues that the social convention approach is validated by statistical regularities in survey data, he conducted his own analysis of rates of FGM in local ethnicity-communities from 25 nationally representative surveys in FGM practicing counties (approximately 66,000), and found that 93% had rates above or below 90%/10%, respectively, consistent with the convention model’s predictions about the instability of interior rates of FGM. Mackie also refers to Sudanese Demographic and Health Survey (DHS) data showing that those who oppose the practice say frequently that it continues because of fear of social criticism. Furthermore, a study aggregating responses to surveys across west Africa found that 77% of women give ‘social’ reasons for the practice, especially social acceptance (40%)</p>
SOCIAL CONVENTION MODEL (MACKIE, 1996; UNICEF, 2007)	Efferson et al. (2015)	<p><u>(In)validation by reference aggregate regularities (-):</u> Efferson et al. (2015) undertook local surveys of the rate of FGM (which they defined in terms of the proportion of eligible young girls cut in each of 45 communities in a given year) in a region of Sudan. They didn’t find any evidence of discontinuity in the rate of FGM – concluding that there was no evidence of the exterior stability predicted by the original social convention model. They similarly found little evidence of discontinuity of ‘implicit attitudes’ to FGM within or between communities.</p>
SOCIAL CONVENTION MODEL (MACKIE, 1996; UNICEF, 2007)	Bellemare et al. (2015)	<p><u>(In)validation by statistical modelling (-):</u> Bellemare et al. (2015) conducted a multi-level statistical analysis of the practice of FGM in West Africa, using a number of nationally representative surveys. They found significant variation in attitudes to FGM at the individual and household level, which was large relative to variation at the village level. This has been interpreted as contrary to the idea of social coordination (Platteau et al., 2017: 5).</p>
SOCIAL CONVENTION MODEL (MACKIE, 1996; UNICEF, 2007)	Platteau et al. (2017)	<p><u>(In)validation by policy outcomes (-):</u> Platteau et al. (2017), referring in particular to studies which have (re)visited anti-FGM policy outcomes in Senegal (Camilotti, 2016a, 2016b), argued that there is little evidence of widespread abandonment of FGM in the Kolda region of Senegal, indicating that Tostan public declarations were not successful in producing the kind of wide-spread, tipping-point, abandonment that the social convention model might predict.</p>
SOCIAL CONVENTION MODEL (MACKIE, 1996; UNICEF, 2007)	Shell Duncan et al. (2011)	<p><u>Validation by ethnography (-/+):</u> Shell Duncan et al. (2011) conducted a mixed-methods study in Senegal and the Gambia, with the explicit aim of testing the social convention account. Contrary to the original rationale for the social convention model (Mackie, 1996), they found that it is not generally believed either that FGM is a pre-requisite for marriage or an avenue for an advantageous marriage. However, they did find widespread reference to other kinds of positive social pressures to practice FGM, leading them to support a ‘peer convention’ account of FGM.</p>

PRE-MARITAL INVESTMENT MODEL OF FGM (CHESNOKOVA AND VAITHIANATHAN, 2010)	Chesnokova & Vaithianathan (2010)	<p><u>Validation by ethnography (+):</u> The authors cite ethnographic evidence of the importance of marriage in FGM <i>by proxy</i> by referring to Mackie (1996).</p> <p><u>Validation by aggregate regularities (+):</u> As part of the justification of the design of their model, the authors cite studies of men (Morison et al., 2004) and women (Boyle, 2002) from some FGM practising countries which indicate that men prefer to marry mutilated women.</p> <p>The authors disaggregate the prevalence of FGM by ethnicity, religion and region, and point out that the existence of multiple levels of FGM within these subgroups is consistent with their model's predictions that different stable interior rates of FGM can occur (p.17).</p> <p><u>Validation by statistical modelling (+):</u> The authors conduct regression modelling of the association between FGM and: (a) marriage age and (b) household wealth (post-marriage) in Burkina Faso. They find that cut women marry earlier and are wealthier on average (after controlling for other covariates) – regularities which are consistent with their competitive marriage market account (p.20).</p>
IDENTITY ECONOMICS MODEL OF FGM (COYNE AND COYNE, 2014)	Coyne and Coyne (2014)	<p><u>Validation by ethnography (+):</u> The authors cite ethnographic sources and research reports referring to the importance of FGM in social actors' cultural identity, consistent with the motivations of their model (Althaus, 1997; Gruenbaum, 2001; Williams, 1998).</p> <p><u>Validation by statistical modelling (+):</u> Citing Wagner (2011), the authors argue that there is strong evidence of an association between ethnicity and FGM practice, which they attribute to the role of ethnic identity.</p> <p><u>Validation by aggregate regularities (+):</u> The authors note that 'good tradition' is cited more frequently as an advantage in national survey data than 'marriage prospects' or 'virginity' (Yoder et al., 2004) which they interpret to support an identity model instead of a marriage-focused model. They also argue that their model accounts for a particular regularity in the practice of FGM: which is that older women tend to be the most staunch supporters of FGM – as predicted by the shifting social-identity incentives that their model associates with being cut (Toubia and Sharief, 2003).</p> <p><u>Validation by Policy Outcomes (+):</u> The authors argue that their model predicts that laws banning FGM which do not address the connection of the practice to identity, will be ineffective. They cite evidence that, as predicted, some legal bans have not been effective (Ako and Akweongo, 2009; Rahman and Toubia, 2000) even amongst immigrant diaspora (Kool, 2010).</p>
HETEROGENOUS THRESHOLD MODEL (NOVAK, 2016)	Novak (2016)	<p><u>Validation by statistical modelling (+):</u></p> <p>Novak (2016) tried to estimate the distribution of thresholds in parts of Burkina Faso. Her estimates are consistent with heterogeneity of thresholds, supporting her advocacy for a heterogeneous coordination model in preference to the original social convention model.</p>
PARENTAL INVESTMENT MODEL (OUEDRAOGO AND KOISSY-KPEIN, 2012)	Ouedraogo and Koissy-Kpein (2012)	<p><u>Validation by statistical modelling (+):</u></p> <p>Consistent with their model which treats FGM and Education as competing possible parental investments, Ouedraogo and Koissy-Kpein (2012) use national household survey data from Burkina Faso and show FGM status is negatively associated with the likelihood of school attendance. They also find that the probability of FGM increases with the belief of mothers than men want FGM to continue (consistent with the idea that this suggests an increased 'return' from the FGM investment). Furthermore, they find that families with fewer boys relative to girls are more likely to practice FGM, consistent with their notion that the labour market success of boys can relieve</p>

		families of the economic incentive to practice FGM on daughters.
COSTLY SIGNAL / VIRGINITY ASSURANCE MODEL (ROSS ET AL., 2016)	Ross (2015)	<u>Validation by statistical modelling (+):</u> Using survey data collected in Columbia, Ross (2015) provide evidence of ‘frequency related’ historical, social transmission of the practice of FGM from Afrocolumbian diaspora to indigenous populations. They also show that the practice of FGM is positively associated with geographic isolation from the wider (non-FGM practising) population – consistent with the idea of <i>interpersonal</i> influences on the practice of FGM (i.e. positive social influence)
COSTLY SIGNAL / VIRGINITY ASSURANCE MODEL (ROSS ET AL., 2016)	Ross (2016)	<u>Validation by statistical modelling (+/-):</u> Bayesian phylogenetic modelling of data related to the prevalence of FGM in different African ‘cultural clusters’ finds somewhat limited support for the role of social stratification (i.e. associated with the use of FGM for marriage advantage) in the persistence of FGM, and stronger evidence of the role of frequency-dependent/conformist social pressures (interpretable as coordination pressures).

Appendix B: Chapter 4 Appendices

Appendix B1: Analytical Strategy for Calculating SZOC and SSPC

First, note that SZOC and SSPC are well defined for use with dummy independent variables (Yang et al., 2017).

The squared zero-order correlation between an independent variable (IV) X_1 and dependent variable (DV) Y is defined as:

$$R^2[Y(X_1)]$$

Where R^2 is an operator that returns the coefficient of determination of the model of Y as a function of X_1 .

The coefficient of semi-partial determination (SSPC) (Yang et al., 2017; Zhang, 2017) between (set of IVs) X_1 and DV Y , given a set of other predictors X_2 is defined as⁸¹:

$$R^2[Y(X_1, X_2)] - R^2[Y(X_2)]$$

This is the increase in R^2 (variance explained) when the set of X_1 predictors are added to a model containing only X_2 predictors. It represents the unique additional variance-explained attributable to X_1 (Murray and Conner, 2009; Yang et al., 2017).

The paradigmatic case for calculating these metrics is OLS regression on a continuous dependent variable. In this case, the dependent variable is binary. OLS regression modelling remains meaningful in this context (in which case it is called a linear probability model – LPM). The R^2 statistic remains meaningful, if more controversial (Gronau, 1998), for the LPM.

Typical objections to the use of the LPM as a modelling framework are as follows. First, predictions outside of the unit interval may occur. Second, coefficient estimates may be biased. Third, errors will be heteroscedastic. These objections are of limited concern in this

⁸¹ In a multiple linear regression context this is equivalent to $\frac{SSE(X_2) - SSE(X_1, X_2)}{SST_Y}$, which is the proportional reduction in the total sum of squared errors of the model after the inclusion of X_1

case. When all independent variables are categorical (as here), estimates will be unbiased, and predictions will be within the unit interval. The problem of heteroscedastic errors applies to Null-Hypothesis Significant Testing (NHST). However, NHST is not undertaken here.

An alternative approach is to use logistic regression, for which the question of primary interest is how to define the coefficient of determination (i.e. R^2) outside of an OLS framework. In this case, I am assisted by Zhang's (2017) recent work. Zhang (2017) has defined a generalization of the coefficient of determination for generalized linear models (GLMs - of which logistic regression is a special case). His measure is consistent with the linear regression measure and measures the degree of variation in the dependent variable explained by the model. Moreover, coefficients of partial determination are explicitly defined in terms of this generalized R^2 (cf Zhang, 2017):

$$\frac{R_V^2[Y(X_1, X_2)] - R_V^2[Y(X_2)]}{1 - R_V^2[Y(X_2)]}$$

Consequently, we can define coefficients of *semi*-partial determination for the GLM case as:

$$\frac{R_V^2[Y(X_1, X_2)] - R_V^2[Y(X_2)]}{1}$$

Where Y is modelled through a generalized linear model, and $R_V^2[\cdot]$ measures Zhang (2017)'s coefficient of determination for generalized linear models⁸².

As a basic robustness check, the analysis was undertaken using both an LPM and logistic regression framework. The ranking of variable importance (on the two variable importance metrics) was almost identical.

In calculating the *squared zero-order correlations* for belief variables $i \in \{1, 2, 3 \dots, 10, 11\}$, X_i was defined as belief dummy variable i . In calculating the *squared semi-partial correlation* (i.e. coefficient of semi-partial determination), X_i was defined as a set of terms including belief indicator i and second-order interactions between i and all other variables. $X_{\sim i}$ was defined as the set of all belief variables *other than* i , as well as all second-order interactions between those variables. The coefficient of semi-partial determination for X_i is associated with belief i in the results.

⁸² Implemented in the analysis via Zhang's *rsq* package for the R programming language

Appendix B2: Robustness of Variable Importance Results across Regions of Senegal

Figure B2.1 presents the results of the variable-importance analysis implemented separately for participants in the 2005 Senegal DHS survey from each of the regions in the country. Results are shown for analysis using unweighted (see weighted results below) logistic regression, with variance-based generalised- R^2 as the metric for calculation of squared zero-order correlation and squared semi-partial correlation metrics (see the main chapter). Each plot represents a region in Senegal, with the plot subtitle showing the name of the region and the R^2 of the complete model (all main-effects and second-order interactions) for that region.

The purpose of the analysis was to assess the sensitivity of the conclusions of the main article (that the social coordination hypothesis is the best-supported model of decision-making about FGM for the Senegalese context) to geographic heterogeneity. The analysis indicates some geographic heterogeneity (which is to be expected – see the main chapter); however, results broadly support the main conclusions of the chapter:

1. Out of the 11 regions, social acceptance is shown to make the largest unique contribution to the explanatory power of the model (squared semi-partial correlation) in 7 regions.
2. Social acceptance is among the two most important variance (on either squared zero-order correlation or squared semi-partial correlation) in 9 of the 11 regions.
3. Social acceptance has a higher importance score than belief in improved marriage-prospects on *both* variable importance metrics in 10 out of 11 regions.
4. Belief in improved marriage prospects is not the most important variable in any region, on either metric.

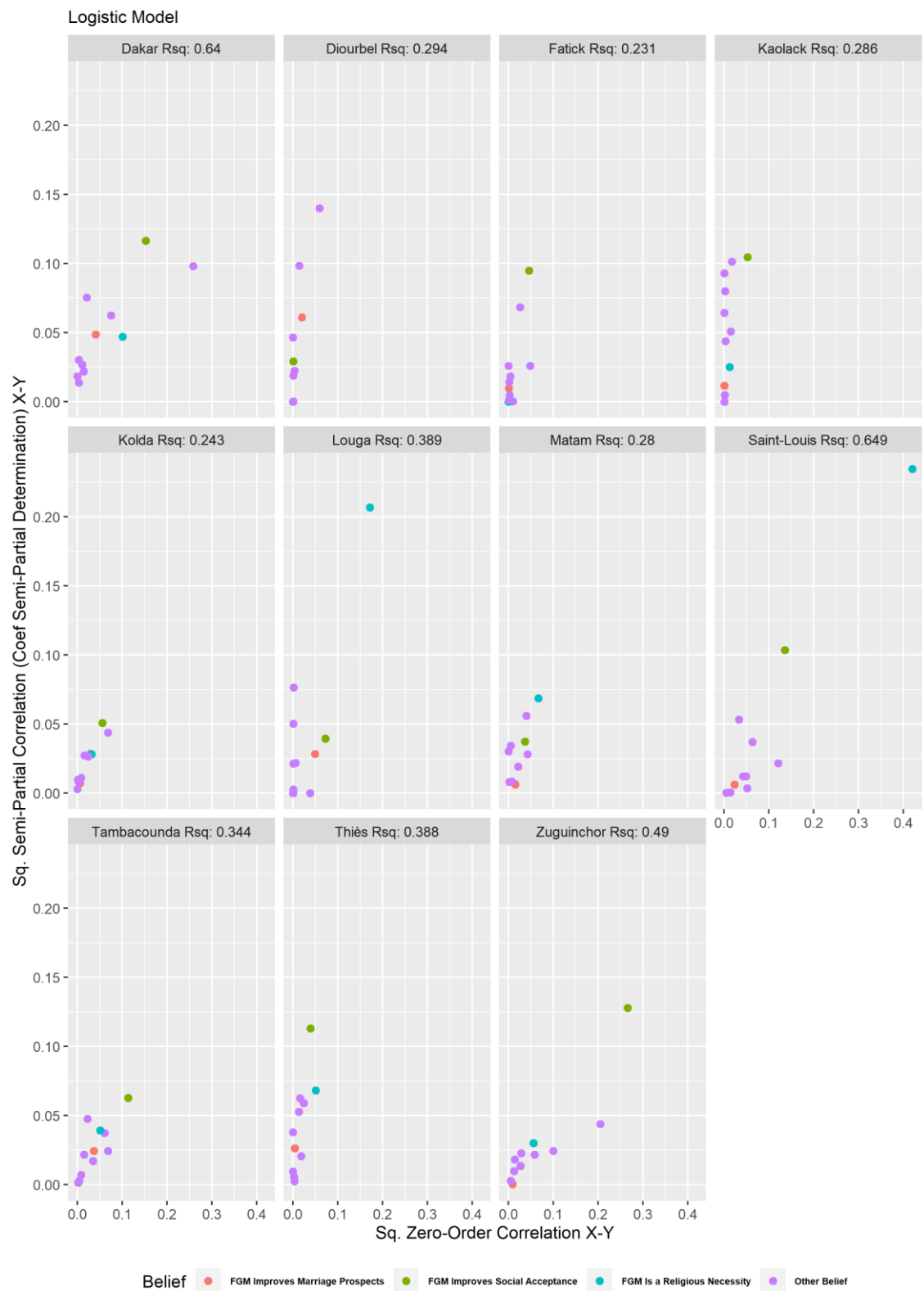


Figure B2.1: Variable Importance Analysis by Region (logistic regression, key three belief variables colour-coded)

Appendix B3: Robustness of Results to Sample Weighting

Figure B3.1 shows the results of variable importance analysis using sampling weights (provided with the Senegal 2005 dataset) in combination with logistic regression, through the facilities of the ‘survey’ and ‘poliscidata’ packages in R. Since Zhang’s (2017) variance-based R^2 has not been implemented for weighted logistic regression, the traditional McFadden’s R^2 metric is used instead (in the calculation of squared zero-order correlations and squared semi-partial correlations).

The purpose of the analysis was to test the robustness of the main conclusions of the chapter to the use of sampling-weights (designed to ensure representativeness with respect to the nation of Senegal). The results are very similar to the un-weighted analysis, showing social acceptance and religion to be the most important key motivators in actor’s decisions about FGM (both metrics) and showing belief in improved marriage prospects to have limited importance (both metrics).

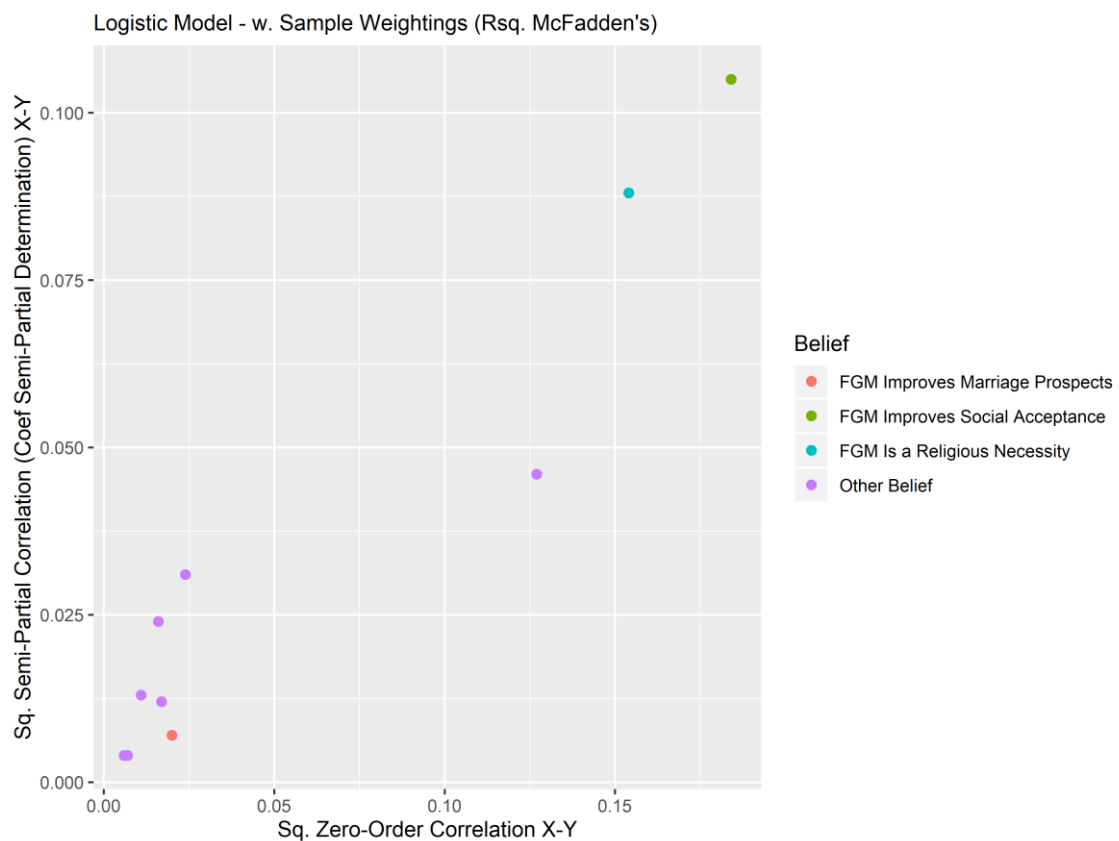


Figure B3.1: Variable Importance Analysis using Survey-Weighted Logistic Regression (key three belief variables colour-coded)

Appendix C: Chapter 5 Appendices

Appendix C1: Sources of Model Elaborations and De-Idealisations in Existing Literature

As noted in the main chapter, in searching for credible de-idealisations and elaborations of the standard coordination model, I considered a range of sources with the potential to highlight important aspects of the social dynamics of FGM which the standard model omits (or inappropriately idealises).

I focused on the Senegalese context. Since the current stage of model development (Chapter 5) is primarily about the representation of micro-level social processes (i.e. action and interaction at the individual level), I focused on qualitative research. I examined a set of studies conducted by Shell-Duncan and colleagues (Hernlund and Shell-Duncan, 2007; Shell-Duncan et al., 2011; Shell-Duncan and Hernlund, 2006), based on fieldwork in Senegal and the Gambia that had a substantial qualitative component. I also examined the ethnographic account provided by Dellenborg (2004), based on her fieldwork among the Jola in Senegal.

I found that these studies emphasised a number of features of decision-making and interaction that could be important to include in a model of FGM as a ‘social norm of coordination’. There was a particular emphasis on power-relations and the different degrees of influence that actors had over the decisions of others. This led me to the consideration of ‘homogeneity of authority’ as a model elaboration. There was also a clear emphasis on localised social interactions, within social networks, families and households.

This led me to a consideration of social and household reference networks, as alternative representations of the ‘global’ reference group assumed in standard coordination models. Furthermore, qualitative accounts strongly suggest that the form of social pressures (and the interactions underlying these) can extend beyond ‘implicit’ coordination with the actions of others to avoid exclusion, to include explicit and deliberate social sanctioning of non-practising actors. For this reason, I considered the potential implications of explicitly representing norm-enforcement activity. Finally, qualitative accounts make clear that some actors are strong supporters of FGM, including for religious reasons, suggesting the need to consider FGM as a ‘subjective benefit’ in decision-making.

I also considered theoretical discussions of the social dynamics of FGM as a social norm of coordination, and associated intervention activities (Cloward, 2015; Mackie, 2000; Mackie and LeJeune, 2009; Mackie et al., 2015). Narrative elements of these discussions go beyond the formal specification of the standard coordination model in a number of respects. These ‘informal’ aspects of the discussion, especially when they are emphasised as important, may point of features of social dynamics which *should* be represented explicitly in the formal model and which might affect its dynamics. These discussions emphasised, among other things, the potential importance of power and authority in coordination dynamics (consistent with the qualitative literature), as well as the idea of actors varying in their susceptibility to social influence (i.e. heterogeneity of autonomy).

Theoretical discussion of interventions to end FGM as a social norm of coordination highlighted important elements of the ‘social norm approach’ which are not included in the standard social coordination model, including coalition formation and organised diffusion. Examination of evaluations of the Tostan program in Senegal, especially those with a qualitative component (e.g. Diop et al., 2008, UNICEF, 2008), provided an opportunity to cross-reference these more ‘theoretical’ ideas about the intervention process with observations real interventions that have been associated with the approach. Although, since the aim of the exercise is to identify credible design *possibilities*, rather than to arrive at a final conclusive representation of the intervention process, I didn’t try to ‘test’ theoretical ideas about the intervention process against empirical accounts. In fact, this might be counterproductive, since a failure of a certain process (e.g. organised diffusion) to occur in a particular situation wouldn’t necessarily imply that it *can’t* occur as part of the intervention. Implementational difficulties might help or obstruct intervention activities independently of the correctness of theoretical accounts of the underlying social dynamics.

It is important that this activity not be confused with an attempt to synthesise qualitative research in Senegal into a ‘new theory’. The aim of the exercise was to identify potentially important aspects of social dynamics which may have been omitted in standard models of FGM as a social norm of coordination. Thus, the aim was not to re-litigate the core theory of decision-making selected in Chapter 4, but to identify important details in *extrapolating* from that theory to a model of the social dynamics of FGM.

I did not, and in practice could not, consider *all* of the potential features of social dynamics highlighted by these accounts. Many of the accounts included detailed narratives with a very wide range of possibilities for formal representation. Instead, I focused on those features

which I judged to have the most potential importance in the context of building a coordination model of FGM.

As with other parts of the modelling strategy in this thesis, there remains considerable scope for others to build on my efforts. Other modellers may wish to identify, formalise and explore the implications of other design possibilities that are suggested in the existing literature but not considered in my analysis. If these design possibilities can be shown to disrupt the key dynamics of existing models, then they should be added to the set of recognised areas of ‘uncertainty’ in the design of such models.

Appendix C2: Calculation of the Rate of ‘Final’ Decision-Making about the FGM Status of Girls within a Local Community

The calculation of the probability that, in a given time-step, a ‘decision-maker’ agent would need to make a ‘final decision’ about the cutting status of one of the girls in their household proceeded as follows. Relevant empirical figures for b and c were taken from an international survey of household characteristics (United Nations Department of Economics and Social Affairs, 2017: 16), assuming that, on average, 50% of children under 15 are girls.

Let a = the number of households in the community, where n is the number of adults

$$a = \frac{n}{6}$$

Let b = the proportion of households with children

$$b = 0.84$$

Let c = the average number of female children (aged under 15) per household

$$c = 2.25$$

Let d = the average number of children in the community

$$d = a \cdot b \cdot c = \frac{n}{6} \cdot 0.84 \cdot 2.25$$

Let l = the number of decisions to be taken each year such that, on average, all girls are cut by age 15 if they are to be cut⁸³

⁸³ Kandala and Shell-Duncan (2019: Table 1) report that less than 1% of girls are cut *after* the age of 15.

$$l = \frac{d}{15}$$

Let \mathfrak{m} = the number of final decision-makers in the community, assuming one per-household

$$\mathfrak{m} = a = \frac{n}{6}$$

Assuming \mathfrak{x} is the number of ‘time-points’ ($\mathfrak{t}_{\mathfrak{k}}$) related to FGM each year, let $\Pr(\text{decision}|\mathfrak{t}_{\mathfrak{k}} = k)_i$ be the probability that decision-maker \mathfrak{i} takes a decision about one-of their girls in a single time-point

$$P(\text{decision}|\mathfrak{t}_{\mathfrak{k}} = k)_i = \frac{l}{x \cdot \mathfrak{m}}$$

Appendix C3: A Simple Spatial Network Algorithm

To create a social network within a population of agents, I implemented the following algorithm based on the work of Hamill and Gilbert (2009):

1. Create a population of n actors distributed randomly in a 2D space (a.k.a ‘network space’) that wraps on its edges
2. Each actor has a fixed ‘social reach’ (r)
3. Actors form network connections with all other actors located within a radius of r
4. For each actor, indexed i , their ‘social reference group’ ($\{\text{social-ref}\}_i$) is all other actors they are connected to

Apart from creating realistic network properties such as ‘clustering’ (Hamill and Gilbert, 2009), this algorithm has the advantage that some of its properties can be established analytically. This, in turn, allows the overall connectivity of the network to be parameterised directly and allows network properties to be independent of the number of agents used, and the size of 2D space chosen.

Theorem (Spatial Network Connectivity)

The average ‘degree’ (number of network connections) of agents under this algorithm can be controlled by the ‘ r ’ parameter, and will be equal to:

$$(n - 1) \cdot \frac{\pi r^2}{A_t}$$

Where A_t is the total area of the 2D space.

Thus, the average degree of the network ($\mu_{Rsocial}$) can be parameterised by defining r as follows:

$$r = \sqrt{\frac{A_t \cdot \mu_{Rsocial}}{\pi \cdot (n - 1)}}, A_t > 0, n > 1$$

Proof (Spatial Network Connectivity)

First, note that actor i connects to actor j *iff* the location of actor j is within r of actor i . Since all positions for j within the 2D space are equally likely, the probability that i is within r of j is the proportion of points in space for which this will occur. This is simply the ratio of the area of the radius r and the area of the wrapped 2D space, which is:

$$Pr_{ij} = \frac{\pi r^2}{A_t}$$

Where A_t is the area of the 2D space and P_{ij} is the probability that actor i will connect to actor j , for $j \neq i$.

Since this probability is independent for all other actors in the agent population, we can expect the number of actors connected to i to be Binomial distributed, with $n - 1$ independent trials, each with probability $\frac{\pi r^2}{A_t}$ of success (connection). Here we are simply interested in the expected value for i , which will be $(n - 1) \cdot \frac{\pi r^2}{A_t}$. We will call this figure (the expected connectivity of each actor): $\mu_{Rsocial}$

We can then parametrize the social circle algorithm by defining r in terms of $\mu_{Rsocial}$:

$$r = \sqrt{\frac{A_t \cdot \mu_{Rsocial}}{\pi \cdot (n - 1)}}, A_t > 0, n > 1$$

Note that, defined in this way, we can use any size of 2D space, and any agent population size, without affecting the connectivity of the network, as long as we update the values of the inputs to the formula as appropriate.

Appendix C4: Creating Household Network Cliques

Household reference networks ‘cliques’ were implemented in the agent-based model using a simple algorithm implemented during the initialisation of the simulation (i.e. before any dynamics occurred). This algorithm is elaborated slightly for the general model (see Chapter 6 and ODD Documentation). The algorithm worked as follows:

1. While there are actors in the population not in a household
 - a. Ask up to x actors at random to form a fully-connected network clique (where x is Poisson distributed, with a mean of 6).
2. For each actor, indexed i , their ‘household reference group’ ($\{\text{household}\}_i$) is all other actors in their network clique

Appendix C5: Implementing an ‘Organised Diffusion’ Process

The organised diffusion process is an elaboration of the formal representation of the intervention process in the simulation. It occurs as part of the intervention, prior to general coordinated interaction within the model.

1. A set of actors is targeted for the intervention.
2. The educational effect of the intervention is applied.
3. Participants decide whether to join the initial coalition (with probability $\frac{H_i}{\bar{M}}$).
4. Initial coalition members reach out to z_4 others (randomly chosen) who are not part of the intervention:
 - a. They ‘transmit’ their knowledge about the harms of FGM to these others. The ‘recruited’ actors increase their perception of the cost of FGM (H_i) by a random amount in-between their original value and the value of the agent trying to recruit them.
 - b. Recruited agents decide whether they want to join the initial coalition (with probability $\frac{H_i}{\bar{M}}$).
 - c. Agents newly joining the initial coalition attempt to recruit z_4 others (return to step 4)
 - d. ... this continues until no more actors join the initial coalition
5. The intervention process proceeds, with a final coalition stabilizing or collapsing and then the standard coordination process occurring across the community.

Appendix C6: Exploratory robustness simulation results

Table C6.1: Exploratory Robustness Analysis of Modifications to the Standard Coordination Model of FGM

Model Variant / Modification	Effect on simulations of persistence	Affect on simulations of intervention
Standard Intervention w. Heterogeneous Response (to the intervention)	NA	<p>Relative to: <i>Standard Coordination Model w. Standard Intervention</i>.</p> <p>Decreased average reduction in FGM when targeting at random ($\mu = 0.49$ vs. $\mu = 0.64$) and positive spillovers disappear on average ($\mu = -0.06$ vs. $\mu = 0.09$). Under random targeting, increased proportion of interventions had negative efficiencies (49% vs. 33%).</p> <p>Relative advantages of targeting willing agents enhanced, with greater average reduction in FGM when targeting willing agents ($\mu_{\text{willing}} = 0.5$ vs. $\mu_{\text{reluctant}} = 0.46$ compared to $\mu_{\text{willing}} = 0.64$ vs. $\mu_{\text{reluctant}} = 0.63$ in the baseline model).</p> <p><i>Processual interpretation:</i> Heterogeneous responsiveness creates increased coalition collapse and smaller initial coalitions, this effect is exaggerated when targeting actors who are reluctant to abandon FGM, and attenuated when targeting actors willing to abandon.</p>
Standard Intervention w. Heterogeneous Response (to the intervention) & Organised Diffusion		<p>Relative to: <i>Standard Coordination Model w. Standard Intervention w. Heterogeneous Response</i>.</p> <p>Greatly increased intervention performance (reduction in FGM) under random targeting ($\mu = 0.73$ vs. $\mu = 0.49$). Greatly increased average positive spillovers under random targeting ($\mu = 0.18$ vs. $\mu = -0.06$). Greatly reduced proportion of interventions with negative spillovers (25% vs. 49%).</p> <p>Reversal of the differential of effectiveness of targeting reluctant agents versus willing agents, with targeting willing agents creating larger decreases in FGM ($\mu_{\text{willing}} = 0.7$ vs. $\mu_{\text{reluctant}} = 0.77$) compared to ($\mu_{\text{willing}} = 0.5$ vs. $\mu_{\text{reluctant}} = 0.46$) in the baseline model.</p> <p><i>Processual interpretation:</i> Organised diffusion reduces the possibility of coalition collapse by recruiting to the initial coalition and spreading the educational effect of the intervention. This removes some of the drawbacks of targeting reluctant actors, whilst capitalising on its advantages (i.e. improved positive spillovers).</p>
Coordination Model w. Social Reference Network ($\mu_{\text{degree}} = 20$)	<p>Relative to: <i>Standard Coordination Model</i> Proportion of interior rates increased to 19% (versus 0%).</p> <p><i>Processual Interpretation:</i> Interior rates of FGM occur because practice and non-practice can stabilise within clusters of the network.</p>	<p>Relative to: <i>Standard Coordination Model w. Standard Intervention</i>.</p> <p>Similar performance of simulated interventions with random targeting. Reduced risk of coalition collapse relative to baseline model (23% vs. 31%) under random targeting. Enhanced average spillover advantage from targeting willing actors versus reluctant actors ($\mu_{\text{opposers}} = 0.11$ vs. $\mu_{\text{supporters}} = 0.05$).</p> <p><i>Processual interpretation:</i> Introduction of network greatly decreases the risk of coalition collapse, since stable coalitions can form in clusters of the network. This effect is enhanced when targeting willing actors (12% of coalitions collapsed) and attenuated when targeting reluctant actors (32% of coalitions collapsed).</p>
Coordination Model w. Reference Network ($\mu_{\text{degree}} = 20$) and Building relationships between coalition members	NA	<p>Relative to: <i>Coordination Model w. Reference Network</i> ($\mu_{\text{degree}} = 20$).</p> <p>Greatly increased average reduction in FGM practice under random targeting ($\mu = 0.86$ vs. $\mu = 0.65$). Far fewer interventions with negative spillovers (4% vs. 34%) and coalition collapse virtually eliminated ($< 0.1\%$ vs. 23%) (under random targeting). Larger coalitions ($\mu = 440$ vs. $\mu = 354$) under random targeting.</p> <p>Reversal of advantages from targeting willing actors, mean reduction in FGM when targeting willing actors was 0.79</p>

(*vs.* 0.89), mean spillover when targeting willing actors was 0.24 (*vs.* 0.34, compared to $\mu_{willing} = 0.11$ *vs.* $\mu_{reluctant} = 0.05$ under comparator model).

Processual interpretation: Building relationships between initial coalition participants increases the mutual support provided by coalition members, virtually eliminates coalition collapse and increases the size of stable coalitions. This leads to increased effectiveness overall. Advantages of targeting opposers (stable coalition formation) are attenuated, advantages of targeting supporters (increased positive spillovers) are enhanced.

Coordination Model w. Decision-Makers as Reference Group	<p>Relative to: <i>Standard Coordination Model</i></p> <p>Proportion of interior rates increased to 19% (versus 0%).</p> <p><i>Processual Interpretation:</i> Interior rates of FGM occur because of temporal delays in decision-making which limited changes in behaviour</p>	<p>Relative to: <i>Standard Coordination Model w. Standard Intervention.</i></p> <p>Intervention performance and other characteristics relatively similar, with slightly worse average reduction in FGM across all three targeting strategies ($\mu_{random} = 0.59$, $\mu_{reluctant} = 0.6$, $\mu_{willing} = 0.61$) compared to standard model ($\mu_{random} = 0.64$, $\mu_{reluctant} = 0.63$, $\mu_{willing} = 0.64$).</p> <p><i>Processual Interpretation:</i> Temporal delays in actors' updates of their decision-making reduced the spread of the intervention, leading to a smaller reduction in FGM in the time-frame of the experiment (30 time-steps).</p>
Coordination Model w. Household as a reference group (80% of social coordination incentives)	<p>Relative to: <i>Standard Coordination Model</i></p> <p>Proportion of interior rates of FGM greatly increased to 38% (<i>vs.</i> 0%)</p> <p><i>Processual Interpretation:</i> Interior rates of FGM occur because of coordination within household network-cliques allowing heterogeneity of decision-making to stabilise.</p>	<p>Relative to: <i>Standard Coordination Model w. Standard Intervention.</i></p> <p>Greatly reduced proportion of interventions with coalition collapse (3% <i>vs.</i> 31%) under random targeting and other targeting biases.</p> <p><i>Processual Interpretation:</i> Mutual support for coalitions can be found among actors within households, this allows initial coalitions to stabilise and prevents coalition collapse.</p>
Coordination Model w. Symmetric Social Pressures	<p>Relative to: <i>Standard Coordination Model</i></p> <p>Distribution of final rates of FGM strongly shifted toward abandonment. Average final rate of FGM was 0.23 (versus 0.5).</p> <p><i>Processual Interpretation:</i> Introduction of symmetric social costs creates an a-symmetry in overall social costs which favours abandonment.</p>	<p>Relative to: <i>Standard Coordination Model w. Standard Intervention.</i></p> <p>Average reduction in the rate of FGM after simulated intervention greatly increased ($\mu = 0.84$ <i>vs.</i> $\mu = 0.64$). Other indicators (spillovers, coalition size and prop. coalitions collapsed) also greatly improved.</p> <p><i>Processual Interpretation:</i> Introduction of symmetric social costs creates an a-symmetry in overall social costs which favours abandonment.</p>
Coordination Model w. FGM as a subjective benefit ($s_2 = 0.4$)	<p>Relative to: <i>Standard Coordination Model</i></p> <p>Distribution of final rates of FGM strongly shifted toward universal practice. Average final rate of FGM was 0.89 (versus 0.5).</p> <p>A small number of interior rates of FGM occurred 0.3% (<i>vs.</i> 0%).</p> <p><i>Processual Interpretation:</i> Introduction of FGM as an intrinsic good for a small number of actors shifts the distribution of costs and benefits in a way which strongly favours practice. Also, under standard decision-functions, actors who attribute positive intrinsic value to FGM practice it unconditionally (creating potential for interior rates).</p>	<p>Relative to: <i>Standard Coordination Model w. Standard Intervention.</i></p> <p>Under random targeting, simulated interventions were substantially worse in terms of reduction in the rate of FGM practice ($\mu = 0.46$ <i>vs.</i> 0.64) and other metrics (spillovers, efficiencies, coalition size and prop. coalitions collapsed).</p> <p><i>Processual Interpretation:</i> Introduction of FGM as an intrinsic good for a small minority of actors shifts the distribution of costs and benefits in a way which strongly favours practice.</p>
Coordination Model w. Heterogeneity of authority (H_i perfect negative rank correlation with authority)	<p>Relative to: <i>Standard Coordination Model</i></p> <p>Distribution of final rates of FGM shifted toward universal practice. Average final rate of FGM was 0.61 (versus 0.5).</p>	<p>Relative to: <i>Standard Coordination Model w. Standard Intervention.</i></p> <p>When targeting at random: Worse average reduction in FGM under random targeting ($\mu = 0.59$ <i>vs.</i> 0.64). Other metrics (spillovers, coalition size and prop. coalitions collapsed) worse.</p>

	<p><i>Processual Interpretation:</i> Heterogeneity of authority, with highest authority among reluctant agents, inflates social costs favouring universal practice.</p>	<p>When targeting reluctant agents: Intervention performance substantially <i>better</i> than in the baseline model. Average reduction in FGM was 0.76 (vs 0.63) all other metrics (spillovers, coalition size and prop. coalitions collapsed), especially average spillovers ($\mu = 0.21$ vs. 0.07) were also improved.</p> <p>Dramatic disadvantage of targeting willing agents created. This covered all metrics including intervention collapse (opposers: 41% of interventions, supporters: 19% of interventions), average reduction in the rate of FGM ($\mu_{willing} = 0.45$ vs $\mu_{reluctant} = 0.76$) and average spillovers ($\mu_{willing} = -0.1$ vs. $\mu_{reluctant} = 0.21$).</p> <p><i>Processual Interpretation:</i> Concentration of authority among those reluctant to abandon created incentives which favour continued practice. However, when these actors were targeted by the intervention, this effectively <i>converted</i> this authority into a driver for abandonment. Conversely, targeting willing agents places those with the least authority in conflict with authoritative supporters of FGM.</p>
Coordination Model w. Heterogeneity of Autonomy	<p>Relative to: <i>Standard Coordination Model</i></p> <p>No meaningful change in simulation outputs.</p> <p><i>Processual Interpretation:</i> See discussion of autonomy below.</p>	<p>Relative to: <i>Standard Coordination Model w. Standard Intervention.</i></p> <p>Little meaningful change in intervention performance overall under random targeting. Similar performance under different targeting conditions.</p> <p><i>Processual Interpretation:</i> See discussion of autonomy below.</p>
Coordination Model w. Norm Enforcement (10% of those reluctant to abandon as enforcers)	<p>Relative to: <i>Standard Coordination Model</i></p> <p>Distribution of final rates of FGM strongly shifted toward universal practice. Average final rate of FGM was 0.71 (versus 0.5).</p> <p><i>Processual Interpretation:</i> Explicit norm enforcement by a small subgroup of supporters concentrates social pressure to practice in those least willing to abandon</p>	<p>Relative to: <i>Standard Coordination Model w. Standard Intervention.</i></p> <p>Extreme reduction in performance of intervention under random targeting. Average reduction in the rate of FGM was only 0.1 (vs. 0.64). Dramatic worsening of all other metrics, with 90% of coalitions collapsing. Same results when targeting those willing to abandon FGM</p> <p>Complete reversal of effect when targeting those reluctant to abandon. All intervention succeeded in promoting completed abandonment of FGM.</p> <p><i>Processual Interpretation:</i> When enforcement of the social norm is concentrated in a minority of those reluctant to abandon, targeting others in the population who have no role in norm enforcement does nothing to eliminate the social incentives that keep the practice in place. By contrast, even the small interventions which target the enforcers (here: those reluctant to abandon FGM) entirely eliminates social costs of abandonment, shifting the population to the 'superior' equilibrium of complete abandonment.</p>
Coordination Model w. 'Uniform' (i.e. central mean, high variance, not strictly uniform) H_i Dist.	<p>Relative to: <i>Standard Coordination Model</i></p> <p>Some increase in interior rates of FGM (1.8% vs. 0%). Reduced Pearson correlation with initial starting rate ($r = 0.52$ vs. $r = 0.84$).</p> <p><i>Processual Interpretation:</i> 'Uniform' dist. approval creates strong sensitivity of coordination dynamics to minor changes in the realised distribution, allowing low rates of FGM to be driven to high rates, and vice-versa. Stable interior points can also occur in the distribution.</p>	<p>Relative to: <i>Standard Coordination Model w. Standard Intervention.</i></p> <p>Under random targeting, slightly worse performance on all metrics. Exaggerated advantage for targeting willing agents on all metrics, including reduction in the rate of FGM ($\mu_{willing} = 0.71$ vs. $\mu_{reluctant} = 0.57$) compared to ($\mu_{willing} = 0.64$ vs. $\mu_{reluctant} = 0.63$ in standard model).</p> <p>However, visual inspection of results shows that spillovers were muted when targeting willing agents. When targeting reluctant agents, stable coalitions were harder to achieve without a larger intervention with high educational strength, but were much larger then they occurred – leading to complete abandonment.</p> <p><i>Processual Interpretation:</i> Uniform distribution of thresholds exaggerated the range of distributions which experienced intervention collapse. Targeting willing agents addresses this issue. However, positive spillovers tended to be highly muted in this case.</p>
Coordination Model w. 'U-modal' H_i Dist.	<p>Relative to: <i>Standard Coordination Model</i></p> <p>The majority of final rates of FGM in the simulation were interior (73% vs. 0%) and the correlation</p>	<p>Relative to: <i>Standard Coordination Model w. Standard Intervention.</i></p> <p>Intervention performance under random targeting is greatly improved for multiple metrics, including reduction in the rate of FGM ($\mu = 0.79$ vs. $\mu = 0.64$), the average positive spillover</p>

	<p>between initial and final rates was greatly reduced ($r = 0.48$ vs. $r = 0.84$).</p> <p><i>Processual Interpretation:</i> The ‘U-modal’ distribution of threshold creates stable interior rates of FGM, with high or low initial rates of practice able to be driven down or up to these points (respectively).</p>	<p>($\mu = 0.24, \mu = 0.09$), and reduced rate of coalition collapse (8% vs. 32%). However, visual inspection suggests some muting of spillovers among larger interventions with high educational strength.</p> <p>Disadvantage of targeting reluctant agents greatly exaggerated, including worse reduction in rate of FGM ($\mu_{reluctant} = 0.52$ vs. $\mu_{willing} = 0.67$) and greatly exaggerated coalition collapse (47% vs. 0%). However, visual inspection shows that when spillovers did occur when targeting reluctant actors, they were much larger, and lead to complete abandonment of FGM.</p> <p><i>Processual Interpretation:</i> The effect of ‘U-shaped’ distributions is to minimise the risk of intervention collapse when targeting at random or targeting willing actors. This accounts for the increased performance overall. However, when targeting reluctant agents the risk is exaggerated, leading to worse performance overall under this condition.</p>
Coordination Model w. ‘L-modal’ H_i Dist.	<p>Relative to: <i>Standard Coordination Model</i></p> <p>Average final rate of FGM greatly increased to 0.91 (vs. 0.5).</p> <p><i>Processual Interpretation:</i> L-shaped distribution of thresholds tends to drive all-but the lowest rates of FGM to universal practice.</p>	<p>Relative to: <i>Standard Coordination Model w. Standard Intervention.</i></p> <p>Under random targeting, performance of the intervention much worse on all metrics, including average reduction in FGM ($\mu = 0.43$ vs. 0.64) and the average spillover ($\mu = -0.11$ vs. $\mu = 0.09$).</p> <p>Strong disadvantage of targeting willing actors created, with worse reduction in FGM ($\mu_{willing} = 0.42$ vs. $\mu_{reluctant} = 0.48$) and positive spillovers virtually eliminated when targeting willing actors (5% of interventions had positive spillovers, vs. 43% targeting reluctant actors). However, intervention collapse was still lower when targeting reluctant actors (36% versus 47%).</p> <p><i>Processual Interpretation:</i> The ‘L-modal’ distribution makes coalition collapse more likely, and suppresses positive spillovers. Better spillovers can be achieved when targeting reluctant actors, although the intervention needed to be large enough (and have large enough educational effect). Targeting willing actors did reduce intervention collapse, but this benefit was counterbalanced by severely reduced positive spillovers under this form of targeting.</p>
Coordination Model w. ‘J-modal’ H_i Dist.	<p>Relative to: <i>Standard Coordination Model</i></p> <p>Average final rate of FGM greatly reduced to 0.09 (vs. 0.5).</p> <p><i>Processual Interpretation:</i> L-shaped distribution of thresholds tends to drive all-but the highest rates of FGM to universal abandonment.</p>	<p>Relative to: <i>Standard Coordination Model w. Standard Intervention.</i></p> <p>Under random targeting, performance of the intervention much better on all metrics, including average reduction in FGM ($\mu = 0.92$ vs. 0.64) and the average spillover ($\mu = 0.37$ vs. $\mu = 0.09$).</p> <p>Strong disadvantage of targeting reluctant actors created, with worse reduction in FGM ($\mu_{willing} = 1$ vs. $\mu_{reluctant} = 0.71$).</p> <p>Positive spillovers did occur when targeting reluctant agents, but less often (61% of reluctant-agent targeting interventions had positive spillovers, vs. 90% targeting willing agents). However, smaller and weaker interventions targeting reluctant agents experienced intervention collapse (29%), whereas this never happened when targeting willing agents (0% of interventions).</p> <p><i>Processual Interpretation:</i> The ‘J-model’ distribution greatly increases the potential for positive spillover. This makes the spillover advantage from targeting reluctant actors redundant, conversely, some of the problems of coalition collapse when targeting reluctant actors remained.</p>

Appendix D: Chapter 6 Appendices

Appendix D1: Generalising Over Sources of Social Influence

I assume that influence from cutting decision-makers and influence from households/social reference groups are broadly distinguishable. The latter is primarily a source of immediate normative social pressure, while the former is about the future state of the marriage-market/social situation of girls if they are not cut. I controlled the relative influence of each with a parameter $s_4 \in [0,1]$. I then further distinguished between social influence from within households, versus the wider social reference group. I controlled the relative influence of each with parameter $s_3 \in [0,1]$. A pseudo-code representation of this would be:

$$[\text{total social influence}]_i = s_4 \cdot [\text{decision-maker influence}]_i + (1 - s_4) \cdot [\text{normative social influence}]_i$$

Where:

$$[\text{normative social influence}]_i = s_3 \cdot [\text{household influence}]_i + (1 - s_3) \cdot [\text{social reference group influence}]_i$$

Given the characterisation of decision-maker influence (above), I defined pro-FGM influence from this source as the proportion of decision-makers who cut their daughters the last time the decision arose (see below), and anti-FGM influence from this source as the proportion who *didn't* cut their daughters last time the decision-arose.

Other sources of social influence (i.e. normative social influence) incorporated heterogeneous weights, representing the authority of individual actors. These other sources of social influence were also divided into explicit social influence (i.e. FGM practice by pro-FGM norm enforcers or FGM abandonment by anti-FGM norm enforcers) and implicit social influence (i.e. FGM practice in general), controlled by parameter $s_5 \in [0,1]$, for example:

$$[\text{social reference group influence}] = s_5 \cdot [\text{explicit 'norm enforcement.'}] + (1 - s_5) \cdot [\text{implicit social influence}]$$

In the case of explicit social influence, norm-enforcers were re-weighted such that the total explicit social influence from all pro/anti-FGM enforcers (whichever group was larger) was equal to the total implicit influence of all actors (see below).

Formal definitions of all of the components of p_{pro_i} (total pro-FGM social pressure facing actor i) were as follows⁸⁴.

Let A_i be equal to the total implicit pro-FGM influence from the social reference group ($\{\text{social-ref}\}_i$) of actor i :

$$A_i = \frac{\sum_{j \in \{\text{social-ref}\}_i} w_j \cdot d_j}{\sum_{j \in \{\text{social-ref}\}_i} w_j}, i \neq j$$

Where $w_i \in [0,1]$ is the authority weight of actor j and d_j is a decision-indicator which is 1 when practicing FGM and 0 otherwise.

Let B_i be equal to the total *explicit* pro-FGM influence from the social reference group of actor i :

$$B_i = \frac{\sum_{j \in \{\text{social-ref}\}_i} w_j \cdot d_j \cdot \beta_{1j} \cdot w_2}{\sum_{j \in \{\text{social-ref}\}_i} w_j}, i \neq j$$

Where β_{1j} is a dummy indicator that is 1 if actor j is a pro-FGM norm enforcer and 0 otherwise. w_2 is a weighting coefficient chosen such that, the total influence ‘weight’ of all pro/anti-FGM enforcers in the community (whichever group is larger), is equal to the total influence ‘weight’ of all actors on average⁸⁵ (see below).

Let C_i be the total *implicit* social influence from the *household* and let D_i be the total *explicit* social influence from the *household*. These are defined in the same way as A_i and B_i (respectively), except that $\{\text{social-ref}\}_i$ is replaced with $\{\text{household}\}_i$ which is the set of other actors in the household of actor i .

Let E_i be the total social influence from the set of decision-makers that actor i is responsive to (see decision-maker reference group below): $\{\text{decision-makers}\}_i$. This is defined as

⁸⁴ Readers can substitute in $(1-d)$ for d , $(1-\theta)$ for θ and β_2 (indicating an anti-FGM norm enforcers) for β_1 for the full definition of p_{anti_i}

⁸⁵ In the event that, due to stochastic effects, the total norm enforcement influence exceeds 1 (e.g. because the actor is connected to an unusual number of enforcers in the network), the influence is capped at 1 by the simulation.

$$E_i = \frac{\sum_{j \in \{\text{decision-makers}\}_i} \theta_j}{\sum_{j \in \{\text{decision-makers}\}_i} 1}, i \neq j$$

where θ_j is a dummy variable indicating that decision-maker agent j cut a girl in their household the last time the decision-arose.

We can then define p_{pro_i} as follows:

$$p_{pro_i} = (s_4 \cdot E_i) + (1 - s_4) \cdot ([1 - s_3] \cdot [s_5 \cdot B_i + (1 - s_5) \cdot A_i] + s_3 \cdot [s_5 \cdot D_i + (1 - s_5) \cdot C_i])$$

As noted above, I also defined a weighting coefficient, such that the total weighted influence of the largest group of enforcers (pro or anti-FGM, whichever was larger) was equivalent, on average, to the total *implicit* influence of all actors. This maintains the conformist properties of the simulation, and ensures, on average, that social influence is not biased in favour of implicit or explicit social influence (instead, this is explicitly controlled by parameter s_5).

The weighting coefficient for norm enforcement (w_2) was defined as:

$$w_2 = \frac{n}{n_{enforcers}}$$

Where n is the number of agents, and $n_{enforcers}$ is the number of pro-FGM or anti-FGM enforcers (whichever is larger), called ‘the largest enforcement group’. Using w_2 as a weighting coefficient ensures that, on average, the total influence of the largest enforcement group is equal to the total influence of all actors (under implicit enforcement).

Under implicit enforcement, the average (i.e. the *expected*, note the $E[\cdot]$ operator) total influence of all actors is:

$$E \left[\sum_{i=1}^n w_i \right] = \sum_{i=1}^n w_i$$

Where w_i is the weight of actor i .

Under explicit enforcement, the average total influence of the largest enforcement group is:

$$E \left[\sum_{i=1}^n w_i \cdot w_2 \cdot x \right]$$

Where x is an independent random dummy variable (0 or 1) that indicates that actor i is in the largest enforcement group. The expectation of x is $\frac{n_{enforcers}}{n}$. Extracting w_2 and x from the summation we find:

$$E \left[\sum_{i=1}^n w_i \cdot w_2 \cdot x \right] = E \left[w_2 \cdot x \cdot \sum_{i=1}^n w_i \right] = w_2 \cdot E[x] \cdot E \left[\sum_{i=1}^n w_i \right]$$

Substituting the definition of w_2 and evaluating $E[x]$, we find:

$$w_2 \cdot E[x] \cdot E \left[\sum_{i=1}^n w_i \right] = \frac{n}{n_{enforcers}} \cdot \frac{n_{enforcers}}{n} \cdot E \left[\sum_{i=1}^n w_i \right]$$

This then reduces to:

$$1 \cdot \sum_{i=1}^n w_i$$

Appendix D2: Breaking the Notion of ‘Thresholds’

There are important implications of the ‘general’ representation of decision-making developed in Chapter 6, for the notion of individual agent ‘thresholds’. Thresholds have been an important idea in recent formal modelling of coordination dynamics in the FGM literature (e.g. Novak, 2016, Efferson et al., 2019). However, they depend on a particular assumption. This assumption is that the social pressure to *practice* FGM, is the relative-complement of the social pressure to *abandon* FGM.

When social pressure is operationalised in terms of the proportion of others in the reference group who practice FGM, this is always true. In terms of the utility function of the general model, pro-FGM activity (p_{pro_i}) will be the proportion of the reference group practicing FGM, and anti-FGM activity (p_{anti_i}) will be the proportion *not* practicing, i.e. $1 - p_{pro_i}$. When this assumption breaks down, so does the notion of a tipping-point. To see why, consider the following hypothetical scenario.

Assume social pressure is exclusively driven by *norm enforcement* within a social reference group (i.e. $s_4 = 0, s_3 = 0$ and $s_5 = 1$). Also, assume that the maximum social costs for *practicing* FGM are equal to the maximum social costs for *abandoning* the practice (i.e. $s_1 = 1$) and actors are homogenous in their influence (i.e. authority) over others.

Now, consider a social actor who, other things being equal, is indifferent between practicing and abandoning FGM ($H_i = 0$). This actor's threshold *would be* the point at which the level of pro-FGM activity was equal to the level of anti-FGM activity.

Under the standard assumption that social pressure to practice FGM depends on the proportion of others practicing FGM, this would only occur when the level of practice in the actor's reference group is 50%. So, the actors would have one threshold: 50%.

However, under norm enforcement, there are *many* circumstances under which the level of pro-FGM activity could be equal to the level of anti-FGM activity. These quantities have a degree of independence. It could be that 10% of the actor's reference group enforces FGM, whilst 10% of this group enforces abandonment, or it could be that 20% of the reference group enforces practice and 20% enforces abandonment, and so on. The key point is that, in this situation, there are multiple 'points of indifference' (i.e. thresholds) for the actor. As such we cannot abstract away the details of decision-making through the idea of a single threshold (even if we wished to).

Appendix D3: Re-Parameterising the Beta-Distribution in Terms of Mean and Variance

A key aim of the general model is to generalize over the marginal distributions of the attributes of actors in the modelled population. The key challenge here is finding a distribution which is bounded (since all variables in the model have bounded distributions) and which is sufficiently flexible to allow a wide variety of qualitatively distinct probability density functions. The beta distribution, also used in Efferson et al. (2019), fulfils the latter function. It allows a wide variety of forms, including unimodal, bimodal, left-skewed and right-skewed (see Chapter 5). However, it presents further challenges in that it is difficult to generalise over the space of beta distributions because its standard parametrisation is unbounded (i.e. parameters can be infinitely large).

The following outlines the beta family of distributions and demonstrates how it can be re-parametrised in terms of its mean and variance. These are bounded values which allow a full exploration of the space of distributions. Furthermore, I show how this distribution can be decomposed into gamma distributions which allows implementation in the Netlogo programming environment used to implement the model

The beta distribution is a family of probability distributions for which all values outside of 0 and 1 have a zero density. It has PDF:

$$f(x) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha, \beta)}$$

Where:

$$B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha + \beta)}$$

and Γ represents the *gamma function*.

The beta-distribution family spans a wide variety of qualitatively distinct shapes, so can be used to approximate a range of distributions of interest for any finite range of values. The beta distribution is determined by two parameters: α and β . It also has some known dispersive properties in relation to these parameters. Specifically, if X is a beta-distributed random variable, then:

$$E[X] = \frac{\alpha}{\alpha + \beta} = E_X$$

Furthermore:

$$V[X] = \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)} = V_X$$

The main practical barrier to using the beta distribution in a simulation context is that the values of α and β are *unbounded*; they can be any positive real number. As such, it is difficult to define a finite ‘space’ of beta-distributions which can be explored in the simulation. The *solution* to this issue is to define the beta distribution *in terms of* E_X and V_X instead. These values have a more meaningful interpretation (centre and spread), and they are bounded (such that the whole space of distributions can be explored systematically).

We know that E_X is bounded in the interval $[0,1]$, we also know from the *Popoviciu inequality of variance* that the maximum variance of a bounded probability distribution is:

$$\frac{1}{4}(M - m)^2$$

Where M is the upper bound and m is the lower bound. Since these values are 1 and 0 for the beta distribution, this simplifies to a maximum variance of $\frac{1}{4}$ for the beta distribution.

To define the beta distribution in terms of its variance and mean, we simply apply methods for simultaneous equations:

$$E_X = \frac{\alpha}{\alpha + \beta}$$

$$V_X = \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)}$$

First, we define α in terms of β and E_X . Note that here we restrict the solution to the case in which E_X is not equal to 1 (although, of course, it can be arbitrarily close to 1). β is always greater than zero.

$$\alpha = -\frac{\beta E_X}{(E_X - 1)}$$

Then we substitute the definition of α into the variance equation, and solve for β :

$$V_X = \frac{-\frac{\beta E_X}{(E_X - 1)}\beta}{(-\frac{\beta E_X}{(E_X - 1)} + \beta)^2(-\frac{\beta E_X}{(E_X - 1)} + \beta + 1)}$$

$$\Rightarrow \beta = \frac{(E_X - 1)(E_X^2 - E_X + V_X)}{V_X}$$

Substituting this definition of β back into the definition of α , we are left with two definitions of these parameters purely in terms of the desired variance and mean of the distribution:

$$\alpha = -\frac{\frac{(E_X - 1)(E_X^2 - E_X + V_X)}{V_X}E_X}{(E_X - 1)} = -\frac{E_X(E_X^2 - E_X + V_X)}{V_X}$$

$$\beta = \frac{(E_X - 1)(V_X + E_X^2 - E_X)}{V_X}$$

We could then redefine the beta-distribution family in terms of these values, giving them a more intuitive specification. However, in the use case of interest here, we are using built-in functionality from Netlogo to create a beta distribution. Specifically, we are interested in defining our beta distribution in terms of the *Gamma* distribution (which is supported directly

in Netlogo). Here we rely on the following relation between the *Gamma* and *Beta* distributions: if X and Y are independent random variables, where $X \sim \Gamma(\alpha, \theta)$ and $Y \sim \Gamma(\beta, \theta)$, and where α and β are the corresponding parameters of the beta distribution, then the random variable:

$$Z = \frac{X}{X + Y} \Rightarrow Z \sim \text{Beta}(\alpha, \beta)$$

This holds irrespective of θ .

Based on the above, we can define the following random variable in terms of the desired mean and variance of the beta distribution and as a function of gamma-distributed random variables:

$$X \sim \Gamma\left(-\frac{E_Z(E_Z^2 - E_Z + V_Z)}{V_Z}, 1\right)$$

$$Y \sim \Gamma\left(\frac{(E_Z - 1)(V_Z + E_Z^2 - E_Z)}{V_Z}, 1\right)$$

$$Z = \frac{X}{X + Y} \sim \text{Beta}(\mu = E_Z, \sigma^2 = V_Z)$$

These formulas are used to implement the beta distribution in Netlogo and generalize over the distributions of continuous heterogeneous characteristics in the general model.

Appendix D4: Generalising over the Joint-Distribution of Continuous Agent Attributes in the ABM

Appendix D4.1: Constructing Variables with Pre-Specified Marginal Distributions and Pre-Specified Correlations

Let us say that we have two random-variables (i.e. random over the agent population): say, H and W . These have marginal distributions $p_H()$ and $p_W()$, which are specified directly. They also have associated cumulative distribution functions $F_H()$ and $F_W()$.

We also have a random variable H_{noise} which has marginal distribution $p_{H_{noise}}()$, and cumulative distribution function $F_{H_{noise}}()$ and whose Pearson correlation with H has been specified directly by defining H_{noise} as $H + \epsilon_H$ where ϵ_H is a random Gaussian noise variable with a standard deviation specified to create the desired correlation between H_{noise} and H (see Appendix D4.2 below).

We want to specify W such that it maintains the marginal distribution $p_W()$ but the joint distribution $p(H, W)$ has a non-linear correlation similar to the Pearson correlation of H and H_{noise} . To achieve this, we define W so that it retains its marginal distribution but has a perfect non-linear correlation with H_{noise} . We define W as follows:

$$W = F_W^{-1}([F_{H_{noise}}(H + \epsilon_H)])$$

Where F^{-1} is the inverse CDF of W .

Appendix D4.2: Arbitrary Correlations using Random Noise with a Pre-Specified Variance

Theorem (Arbitrary correlations using random noise with a pre-specified variance)

Given:

$$y = x + \epsilon$$

Where x and ϵ are random independent variables and ϵ has an expected value of 0. It will be the case that:

$$\rho_{x,y} = \frac{\sigma_x}{\sqrt{\sigma_x^2 + \sigma_\epsilon^2}}$$

Where $\rho_{x,y}$ is the Pearson product-moment correlation between x and y , and σ_x is the standard deviation of x , etc.

Proof (Arbitrary correlations using random noise with a pre-specified variance)

To prove this, we rely on the following previously established theorems regarding the properties of the expected value operator ($E[\cdot]$), the definition of the variance of a random variable in terms of expected value, and the definition of the Pearson correlation.

The person correlation can be defined:

$$\rho_{x,y} = \frac{E[(x - E[x])(y - E[y])]}{\sigma_x \cdot \sigma_y}$$

Where $E[\cdot]$ is the expected value operator, which has the following established properties:

1. E is distributive with respect to addition: $E(x) + E(y) = E(x + y)$
2. If x and y are independent, E is distributive with respect to multiplication: $x \perp y \rightarrow E(xy) = E(x) \cdot E(y)$
3. The expected value operator applied to a non-random variable returns that variable, e.g., $E(2) = 2, E(E(x)) = E(x)$.
4. Constants can be factored out of the expected value operator, such that $E(x \cdot E(x)) = E(x)^2$

Finally, we rely on the following definition of the variance of a random variable:

$$\sigma_x^2 = V[x] = E[x^2] - E[x]^2$$

Proof of the theorem depends on simplification and substitution within the denominator and the numerator in the definition of the Pearson correlation, for the case in which $y = x + \epsilon$.

$$\begin{aligned} & E[(x - E[x])(y - E[y])] \\ &= E[xy - xE[y] - yE[x] + E[x]E[y]] \\ &= E(xy) - E(xE[y]) - E(yE[x]) + E(E[x]E[y]) \\ &= E(xy) - E(x)E(y) - E(x)E(y) + E(x)E(y) \\ &= E(xy) - E(x)E(y) \\ \text{sub. } y = x + \epsilon &\Rightarrow \dots = E[x(x + \epsilon)] - E[x]E[x + \epsilon] \end{aligned}$$

$$\begin{aligned}
&= E[x^2] + E(x \cdot \epsilon) - E[x](E[x] + E[\epsilon]) \\
&= E[x^2] + E(x \cdot \epsilon) - E[x]^2 - E[x]E[\epsilon] \\
x \perp \epsilon &\Rightarrow \dots = E[x^2] + E[x]E[\epsilon] - E[x]^2 - E[x]E[\epsilon] \\
&= E[x^2] - E[x]^2 = \sigma_x^2
\end{aligned}$$

So, in the case where $y = x + \epsilon$, $E(\epsilon) = 0$ and $x \perp \epsilon$ (independence), the covariance of x and y will be equal to the variance of x .

Therefore, we can re-state the Pearson correlation between x and y as:

$$\begin{aligned}
\rho_{x,y} &= \frac{\sigma_x \cdot \sigma_x}{\sigma_x \cdot \sigma_y} \\
\Rightarrow \rho_{x,y} &= \frac{\sigma_x}{\sigma_y}
\end{aligned}$$

We can assume that σ_y is unknown, whereas σ_x and σ_ϵ are known. As such, it is helpful to re-write this as:

$$\Rightarrow \rho_{x,y} = \frac{\sigma_x}{\sigma_y} = \frac{\sigma_x}{\sqrt{\sigma_x^2 + \sigma_\epsilon^2}}$$

This follows from the distributivity of the variance operator $V(\cdot)$ with respect to addition, under the independence of random variables:

$$x \perp \epsilon \Rightarrow V(x + \epsilon) = V(x) + V(\epsilon)$$

This can be proved by substituting the definition of $y = x + \epsilon$ into the definition of the variance of y in terms of $E[\cdot]$ and then simplifying. The distributivity of the variance operator with respect to addition depends on the independence of ϵ and x but doesn't require that $E(\epsilon)$ or $E(x)$ be equal to 0 (as other parts of the proof do).

As such, we can note that:

$$\sigma_y = \sqrt{V(y)} \Rightarrow \sigma_y = \sqrt{V(x) + V(\epsilon)} = \sqrt{\sigma_x^2 + \sigma_\epsilon^2}$$

Application (Arbitrary correlations using random noise with a pre-specified variance)

The intended application of the theorem is the generation of random variables which have an arbitrary degree of correlation with some prior variable. Rearranging the theorem (assuming all values are positive and greater than 0) above shows that this can be achieved using the following formulation:

$$y = x + \epsilon$$

Where:

$$\sigma_{\epsilon} = \frac{\sqrt{\sigma_x^2 - \rho_{x,y} \cdot \sigma_x^2}}{\rho_{x,y}}$$

With the desired $\rho_{x,y}$ chosen arbitrarily. ϵ can have any probability density as long as $E(\epsilon) = 0$.

Appendix D5: Results for the Global Sensitivity Analysis of Intervention Parameters

Table D5.1: Global Sensitivity Analysis Results (Intervention Parameters)

Parameter	Description	Sobol Total Order Sensitivity [95% CIs]	Sobol First Order Sensitivity [95% CIs]
z_1	Size of the intervention	0.193 [0.1,0.287]	0.123 [0.003,0.229]
z_2	Strength of Educational Effect	0.395 [0.235,0.556]	0 [0,0.081]
z_3	Minimum probability that opposers of FGM will join the intervention	0.215 [0.122,0.315]	0.012 [0,0.111]
z_4	Number of actors that initial coalition members try to influence and recruit	0.225 [0.138,0.309]	0 [0,0.061]
z_5	Proportion of actors in the initial coalition who a given coalition-member forms a social connection with	0.148 [0.065,0.232]	0 [0,0.025]
z_7	Probability that a non-norm-enforcer in the initial coalition becomes an anti-FGM enforcer	0.134 [0.046,0.216]	0 [0,0.07]

Appendix E: Chapter 7 Appendices

Appendix E1: Demonstration of the Latent Variable Estimation Approach

As a more concrete illustration of the use of logistic regression to estimate a latent variable, consider the following demonstration.

We begin with a latent variable: y_{latent} , which is determined by 10 independent standard-normal Gaussian random variables and a random error-term, with fixed weighting coefficients (β_{1-10}):

$$y_{latent} = \beta_1 X_1 + \beta_2 X_2, \dots, \beta_9 X_9 + \beta_{10} X_{10} + \epsilon$$

We assume that y_{latent} is not observed, but an indicator, y_{obs} is. The question is whether we can usefully recover the y_{latent} values in a given sample, using logistic regression as an estimation technique. The following simple Monte-Carlo simulation demonstrates the answer. Based on 10000 independent samples of y_{latent} (using arbitrary β coefficients and treating ϵ as a zero-centred random logistic variable with variance 1), and performing a logistic regression of Y_{obs} on X_{1-10} , we obtain estimates of the unobserved coefficients β_{1-10} . Figure 1 plots the linear predictor for each of the samples against the true y_{latent} value for that sample. The procedure recovers Y_{latent} almost perfectly (near perfect linear correlation with the linear predictor) with the divergence of estimates attributable to the random error term.

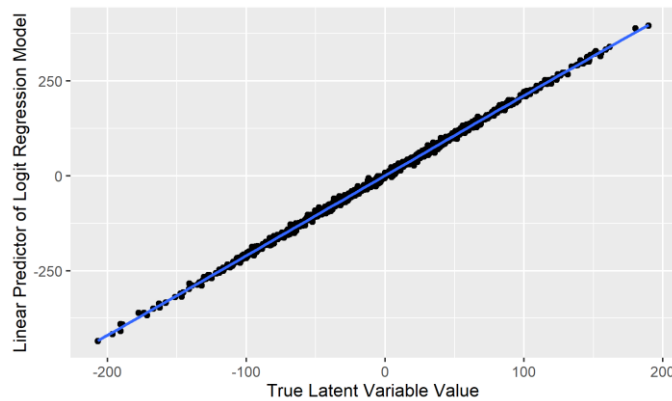


Figure E1.1: Demonstration of Recovering Latent Variables via Logit Regression

Appendix E2: Necessary Assumptions to Estimate H_i for participants in the 2005 Senegal DHS survey

A range of assumptions are required to employ Chapter 7's approach to estimating H_i for participants in the 2005 Senegal DHS survey. Since the regression model is to be estimated on the whole dataset (with averages then disaggregated to the community level), the method implicitly assumes *causal homogeneity* in the relationship between the actors' beliefs and their perception of the intrinsic value of FGM. In other words, if, for example, a belief is heavily weighted by the model either for or against the value of FGM, the method assumes that this weighting is homogenous across the population (or can be usefully approximated as such) and that, when held, that belief is equally important to all actors.

The most striking alternative assumption would be that the causal antecedents of actor's perceptions of the value of FGM are heterogenous at the individual level – i.e. that the beta-coefficients are unique for each individual. It would be difficult to rule out this assumption, although one might suggest that there is a certain upper-bound to this kind of heterogeneity. While, for example, two individuals who view FGM as a religious necessity might differ in the extent to which this 'matters' to them, we would expect them both to view FGM more positively than otherwise identical individuals who did not hold this belief. One other possible source of reassurance is that the model actually used to estimate H_i (see below) is able to explain almost 70% of the variation in actors view that FGM should continue – suggesting that there is, at least, considerable meaningful regularity in the relationship between actors' beliefs and their view of the value of the practice.

Further assumptions relate to the sample characteristics of the survey. Since we are assuming a homogenous causal process, the national representativeness of the survey is of limited concern⁸⁶. However, the method involves estimating the average H_i and variance of H_i at the community level, as an estimate of μ_H and σ_H^2 for that community. As such, representativeness at the community level is a key concern. At the cluster level, the survey involved randomly sampling 21 households within a census enumeration area (typically a large village or city block), with all women aged 15-49 within the household interviewed. We

⁸⁶ The survey is representative of Senegal if weights are applied. This addresses the over-sampling of rural communities.

However, the issue of sampling weights is orthogonal to our concerns in estimating a causal relationship - if a causal relation is homogeneous then (if otherwise valid) results for the available sample will correspond to results for the population (Kohler et al., 2019).

can, therefore, expect cluster level averages to be unbiased with respect to adult women in the cluster⁸⁷. However, in relying on this data as an indicator of the average in the community, we are required to assume that the sampled population (adult women) do not differ substantially from the non-sampled population (adult men) of interest.

Furthermore, as noted below, some women in the survey were excluded because their response on the dependent variable (whether FGM should continue) was uninterpretable, because of missing data on the independent variables (i.e. beliefs about FGM, and education), or because they hadn't heard of FGM (and so were asked no further questions about the practice). We similarly need to assume that the averages estimated from the available sample do not differ substantially from those that would have been estimated from these excluded cases.

Some reassurance with respect to the first assumption comes from the following consideration. The magnitude of bias in estimates will depend on the correlation between inclusion in the sample and the H_i attribute of actors. When there is no such correlation, bias is strictly zero. National aggregate data from Senegal in 2005 shows that the proportion of men and women who said that FGM should *not* continue, was almost identical (75.4% and 74.8% respectively)⁸⁸, suggesting limited correlation between gender and perception of the value of the practice. With respect to the second assumption, of the 1,984 participants (13.58%) in the survey who were excluded due to non-response, 42.3% ($n = 814$) were excluded because they had not heard of FGM, suggesting that they, arguably should not be included in the analysis in any case. Adopting this perspective, the remaining 1170 cases, represent only 8.48% of the surveyed population.

Another key assumption relates to the interpretation of measures in the survey. In the analysis, I assume that actors' perceptions of the intrinsic value of FGM can be modelled as a combination of their beliefs about the practice, their level of education and their wealth (as a potential proxy for their exposure to media and ideas that are critical of FGM). I view this as a relatively safe assumption. Even if such factors don't capture *all* of the influences involved in the formation of actor's perception of the value of FGM, it is difficult to see how any they could fail to play some meaningful role in the formation of such beliefs. Some

⁸⁷ Most households had only one or two respondent women, indicating that intra-class correlation effects at the household level will not be pervasive enough to be a significant concern.

⁸⁸ ICF, 2012. The DHS Program STATcompiler. Funded by USAID. <http://www.statcompiler.com>.

reassurance for this view comes from the high explanatory power of the model (generalised R^2 of almost 70%).

Finally, the analysis requires us to assume that participants' response to the question of whether FGM should continue, or be stopped, can be taken to be a proxy of their evaluative judgement about the net intrinsic benefits or costs of the practice. This assumption is difficult to test, although it seems to be a reasonable interpretation of the question. Nevertheless, it remains a source of uncertainty in the analysis.

Appendix E3: Arriving at the ‘Latent Variable’ Strategy through a Pragmatic Approach

I would like to offer a defence of the latent variable estimation procedure to a sceptical reader and to show that one arrives at a similar procedure even from a position of scepticism. Let us suppose that we reject the idea of H_i as a latent variable that can be estimated from the available data. Nevertheless, we still want to calibrate those aspects of the ABM pertaining to the average preference for FGM, μ_H , and the variation in that preference, σ_H^2 .

If we wanted to heuristically classify communities according to these characteristics, we might start by creating an ‘index’ of actors views of the value of FGM, perhaps defined in terms the number of pro and anti-FGM beliefs that they hold. We might then average this index (and take its variance) at the community level and then heuristically map these values onto the parameters of the simulation (‘high’ support communities, ‘high variance’ communities, and so on). To assist in this task, we might scale the values in the index to match the bounds of the parameters (e.g. scaling the average to between 0 and 1).

This might work quite well. However, we would be left with the conceptual problem that a homogenous weighting of items in our index is somewhat arbitrary and even implausible. It might seem strange to treat all beliefs as equally important in actors’ assessment of the value of the practice. Some may ‘matter’ more than the others.

To address this issue, we might be tempted to find an objective way to ‘weight’ the items in our index – perhaps by findings values for these weights that fulfilled some objective criteria, such as predicting when actors will say they think FGM should continue or stop. At this point, of course, we would arrive at a method like logistic regression, and a procedure analogous to the one adopted here.

The key point is that something similar to the latent-variable estimation procedure adopted here can be justified on a pragmatic basis, without necessarily invoking the concept of latent variables directly.

Appendix E4: Description of Monte-Carlo Experiment

This section reports the use of Monte-Carlo simulation to estimate the expected error associated with estimating μ_H , σ_H^2 (at the community-level) and s_2 (at the population-level) from the 2005 Senegal DHS dataset. The Monte-Carlo simulation is designed to reproduce the conditions under which ‘true’ values of these parameters are generated in the real data, as well as the conditions under which these values are estimated (imperfectly) from the data. For each simulation, since the ‘true’ values of the parameters are known, the error of the (simulated) estimation process(es) can be observed directly. The expected error of an estimation process can, therefore, be equated with the average error of that estimation process under the Monte Carlo simulation.

The estimates of error provided by the Monte Carlo simulation will only be approximate. They depend on plausible but ultimately arbitrary assumptions about the data generating process. They also import the assumptions of the estimation method itself (such as that the determinants of H_i across the population are homogenous). Thus, the results of the simulation can be characterised as providing an *indication* of the expected estimation error, based on the assumption that the simulated data is similar to the real data, and given that the other assumptions necessary for the estimation process are met.

The Monte Carlo simulation was designed to reproduce the following properties of the *population* that formed the DHS sampling frame:

1. A population is divided into 376 community clusters, within which beliefs about FGM are also clustered (e.g. ‘high FGM’ and ‘low FGM’ communities).
2. Each community consists of (on average) 253 adult women who were eligible for participation in the 2005 Senegalese DHS survey⁸⁹.
3. Each member of each community has a set of beliefs about FGM (and other attributes) which add together to determine the total intrinsic value that they attribute to the practice (based on a homogenous set of weighting coefficients).
4. Each member of the community also has an indicator variable which is 1 if the total intrinsic value they attribute to FGM is greater-than-or-equal-to 0, and 0 otherwise.

⁸⁹ Average number of eligible women per-household in the 2005 Senegalese DHS was 2.28, and the average number of households per census enumeration area (site of the cluster) was 111.

The simulation also aimed to reproduce the following properties of the *estimation procedures*, which were used to estimate population characteristics:

1. All estimates are based on a sample of 38 women⁹⁰ from each community cluster
2. Only the determinants of actors' perceived intrinsic value of FGM, and the indicator of whether that perceived value is greater than zero are ever observed in the estimation process (the underlying latent variable is not)
3. Typically, only a *subset* of the determinants of actors' perceived intrinsic value of FGM is observed – as such, estimation must take place in the context of omitted variables.

Based on simulations with these properties, the aim was to estimate the expectation of the average error of the μ_H and σ_H^2 estimators at the community-cluster level (i.e. averaged over 376 clusters per simulation), as well as the average error of the s_2 estimator over repeated simulations (one estimate per simulation).

Simulating the Population (Sampling Frame)

The generation of the simulated population was based on the generation of 376 separate local community populations (i.e. sampling frames for the cluster samples taken in the survey). Each local community population was 'seeded' by one of the communities sampled in the real Senegalese DHS so that the properties of the simulated local populations (individually and in aggregate) would be similar to those of the survey. This 'seeding' process was implemented as follows (it was re-implemented afresh with each simulation).

I aggregated data on beliefs about FGM in each of the 376 communities included in the 2005 Senegalese DHS survey. This aggregation involved taking the mean of each of the 13 dichotomous beliefs recorded in the survey about FGM, as well as the mean and standard-deviation of the number of years of education of respondents in the community. This resulted in a 'seeding' matrix with the following structure (values for illustration)...

Community (i)	Belief 1	Belief 2	Belief 13	Education (M)	Education (SD)
1	0.45	0.7	0.1	2.5	3.57
...	0.1	0.2	0.5	5.4	1.9
376	0.05	0	0.8	4.9	1.12

⁹⁰ The average number of women sampled in the 2005 DHS at the cluster level.

... where each row of the matrix is a community in the real survey, and the columns represent the mean (or standard deviation) of the ‘determinants’ of the perceived intrinsic value of FGM in that community.

Then, for each row in the matrix, a set of 253 individuals was generated, representing all adult women in the sampling frame of the DHS for that community. The named characteristics of the associated 253 individuals corresponded to the names of the columns of the matrix. The individual values of those characteristics were drawn stochastically from distributions whose mean (and standard deviation) matched information in the matrix-row. So, for example, the ‘Belief 1’ attribute of actor j in community i would be a single draw from a Bernoulli distribution (0 or 1) with the expectation of that distribution equal to the mean of ‘Belief 1’ in real community i .

In this manner, the Monte-Carlo simulation was able, for each run of the simulation, to stochastically generate a population with similar properties (in terms of population-level and community-level distributions of characteristics) to those of the real DHS survey.

Determining the Perceived Intrinsic Value of FGM for Each Case in the Simulated Population

Having set the ‘determinants’ of actors beliefs about the intrinsic value of FGM, these were converted to an underlying intrinsic value that the actor attributed to FGM, by weighting each determinant and summing them together. There were 13 dichotomous belief variables for each simulated individual (x_1, \dots, x_{13}) as well as one continuous variable corresponding to the number of years of education of the simulated individual (x_e).

In each simulation, a random set of weights (i.e. Beta-Coefficients) for these variables were drawn and applied to all simulated individuals (across all communities). Although the weights in each simulation were randomly selected (once for each simulation), they were drawn from probability distributions with fixed parameters (across all simulations). In the real data, 9 belief variables corresponded to ‘positive’ beliefs about FGM, and 4 belief variables corresponded to ‘negative’ beliefs about FGM. Years of education is also negatively correlated with the view that FGM should continue. Beta-coefficients for the belief variables corresponding to ‘positive’ beliefs about FGM (β_{1-9}) were drawn from a truncated normal distribution with mean 1, standard deviation 1, and a lower limit of 0. Beta-coefficients for the belief variables corresponding to ‘negative’ beliefs about FGM (β_{10-13}) were drawn from a truncated normal distribution with mean $-\frac{9}{5}$, standard deviation $\frac{9}{5}$ and an upper limit of 0.

Finally, the beta-coefficients for the ‘years-of-education’ variable (β_e) were drawn from a truncated normal distribution of mean $\frac{-9}{50}$, standard deviation $\frac{9}{50}$ and upper limit of 0.

Although this choice for the distribution of beta-coefficients used in repeated runs of the simulation was essentially arbitrarily, it was designed to fulfil a basic heuristic criterion of ‘balance’. On average, a simulated individual who holds all the positive and all the negative beliefs about FGM, and has 10 years of education, would have a neutral view of the practice. This helped to ensure that, like in the real data, all simulations were populated by many actors whose perception of the intrinsic value of FGM was positive and by many actors whose perception was negative.

Noting the above, the latent H_i attribute of each simulated individual i was defined as follows...

$$H_i = \beta_1 x_{1i} + \beta_2 x_{2i}, \dots, \beta_{13} x_{13i} + \beta_e x_{ei}$$

H_i was unobserved as part of the estimation procedures (below). However, each simulated individual was also assigned an indicator variable H_{obs_i} which was positive when H_i is greater-than-or-equal to 0, and negative otherwise.

Implementing the Estimation Procedure

The first step in estimation from the simulated data was to simulate the sampling process of the survey. This was undertaken by taking a simple random sample of 38 cases from each of the 376 local populations in the simulated data. This resulted in a sample-dataset, in each simulation, of 14,288 cases.

The second step was to estimate the underlying H_i score for each participant in the sample-dataset using logistic regression. In order to incorporate omitted variable bias into the simulation H_{obs} was regressed on an equal-probability random selection of $\alpha \in \{9,10,11,12,13,14\}$ of the 14 determinants of H_i , where α is an input parameter in the simulation. All of the sample-dataset was used in regression estimation. The linear predictor for each individual in the sample dataset was taken as the estimate of H_i for that case: \hat{H}_i , and was scaled to the unit interval.

The third step was generating a set of estimates of μ_{H_j} and $\sigma_{H_j}^2$ for each of the clusters ($j \in \{1, \dots, 376\}$) in the sample-dataset. These were estimated simply as the mean of \hat{H}_i and the

variance of \hat{H}_i respectively among sampled individuals from cluster j (after scaling estimates of H_I to the unit interval).

The final step in the estimation, was to estimate S_2 . The estimator for S_2 was

$$\frac{\max [\hat{H}]}{|\min [\hat{H}]|}$$

where both operators act across the set of all \hat{H}_i in the sample-dataset.

Calculating Error Per-Simulation

To calculate the estimation error for an individual simulation, the ‘true’ values at the community-level (mean and variance of perceived intrinsic value at the population level (true ratio of maximum to minimum intrinsic value), were calculated directly from the simulated population (after scaling true values of H_i to the unit interval). The error of each community level estimator was calculated at the square root of the average squared difference (RMSE) between the estimator and the true value at the community level. The error of the estimate for S_2 was calculated as the square root of the squared difference between the estimate and the true value (root squared error).

Overview of the Simulations

In total, the Monte-Carlo experiment was run 6000 times, representing 1000 simulations for each value of alpha (9, 10, 11,12, 13 or 14), where alpha represents the number of determinants (out of 14) available for the regression analysis. During each simulation, the generalised r-squared (Zhang, 2017) of the logistic regression model was also calculated.

Results: Assessment of the Expected Error of the Estimators

In analysing the results of the MC simulations, I sought to exploit the fact that the error of the estimators was likely to be related to the r-squared of the logistic regression model, which, in turn, would depend on the determinants omitted from the regression analysis. As such, I was primarily interested in the expected (I.e. average error) of each of the estimators at different levels of model fit (r-squared). I divided the results for each estimator into 6 equally sized quantiles based on the r-squared of the corresponding model fit and averaged the reported error (Mean of RMSE/Mean of Root Squared Error). This is used as an estimate for the expected error of the estimator with a model r-squared within each interval. The r-squared of the model used in the main analysis was 69% with the expected error of the estimators read from the corresponding rows of the tables below.

Table E4.1: MC Error rates for Mean(H) Estimator (by r-squared of the model)

Model R-Squared	Mean of the RMSE of μ_H estimator (at the community-level)
(0.212,0.57]	0.083092538641096
(0.57,0.664]	0.068677675656155
(0.664,0.749]	0.058608997444881
(0.749,0.851]	0.050358711527079
(0.851,0.97]	0.039737478619149
(0.97,1]	0.023828878416248

Table E4.2: MC Error rates for s_2 Estimator (by r-squared of the model)

Model R-Squared	Mean of Root Squared Error of s_2 estimator (at the population-level)
(0.212,0.57]	0.316172567234501
(0.57,0.664]	0.238575698965909
(0.664,0.749]	0.204525172922742
(0.749,0.851]	0.177491606783698
(0.851,0.97]	0.140200735656714
(0.97,1]	0.056148824815644

Table E4.3: MC Error rates for Var(H) Estimator (by r-squared of the model)

Model R-Squared	Mean of the RMSE of σ_H^2 estimator (at the community-level)
(0.212,0.57]	0.005667753451287
(0.57,0.664]	0.004880537425101
(0.664,0.749]	0.004517681711487
(0.749,0.851]	0.004049794312236
(0.851,0.97]	0.003365527517668
(0.97,1]	0.002659158703809

Appendix F: Chapter 8 Appendices

Appendix F1: Exploratory Analysis of Failure Scenarios

In order to explore the possibility of ‘intervention failure’ in the ABM, repeated simulated experiments were undertaken. These followed the design of the ‘Dynamics of Intervention’ experiment described in Chapter 5. This experiment was repeated 24,000 times. In each run of the simulation, *non-calibrated* parameters had their values selected uniformly at random from the intervals described in Chapter 6 (Table 9) and Appendix D5 (Table D5.1) for the global sensitivity analysis. Calibrated parameters ($\mu_H, \sigma_H^2, s_2, \rho_{Hy}$) had their values drawn from the calibrated parameter set constructed in Chapter 7.

The outcomes of these simulated interventions were analysed along with the parameter values that precipitated them, with the aim of discovering patterns that might indicate distinctive ‘failure scenarios’. The strategy employed in the analysis was to analyse three separate failure outcomes:

1. No Change: No reduction in the rate of FGM
2. Negative Spillover: The number of actors abandoning is lower than the number targeted.
3. Partial Change: Some actors keep practicing FGM

For each of these outcomes, the relationship between the parameters of the simulation and the probability of the outcome occurring, was analysed. This involved creating a ‘report’ for each parameter analysing its relationship to the different outcomes, individually, and in interaction with basic features of the simulated intervention – such as intervention size. It also involved using simple tree-based machine learning tools to build a multi-predictor model of each failure outcome, in terms of combinations of the different parameters. These analyses were instructive in developing my understanding of the dynamics of the model, although primarily in pointing toward issues for further investigation. The reports for each parameter, as well as visual displays of the tree-based machine learning models (using the RPART algorithm with default parameters, Therneau and Atkinson, 2019) are included in the DVD-ROM attached to the thesis.

Appendix F2: Failure Scenario 2 (Out-of-Reach Effects)

Under the original social convention model, increasing intervention size is an unambiguously positive strategy. Larger interventions will be more likely to reach the tipping point. As such, larger interventions will always be more likely to create *some* change, more likely to create positive spillovers *and* more likely to generate complete change. Johansen et al. (2013) made a similar suggestion, associating intervention failure with a failure to involve the whole community in change efforts.

However, exploration of the ABM suggested possible scenarios in which increasing intervention size might be a fundamentally *inefficient strategy*. Under some circumstances, greatly increasing intervention size might produce few additional reductions in the rate of FGM⁹¹. I refer to this as ‘out-of-reach’ effects, for reasons that should become clear shortly.

Put simply, interventions with low *educational strength* may experience severe diminishing returns with increasing intervention size. As such, large (and possibly costly) increases in intervention size could become highly inefficient. To understand why this can occur, consider the following scenario. There is an intervention with an educational strength of 35%. Let’s assume that when targeting actors in the population at random with this intervention, around 1 in 10 will be sufficiently receptive to FGM abandonment, and sufficiently persuaded by the intervention, that they will be prepared to join an initial coalition of those abandoning FGM. That means that if the intervention targets 30% of the local population at random, around 3% of actors will join the intervention. Now consider what will happen if one increases the size of the intervention to 60% of the population. Well, the size of the initial coalition will grow to only 6%. Thus, one requires a very large increase in intervention size to generate only a small increase in the size of the initial coalition.

If the initial recruitment to the intervention is inefficient in this way, the overall intervention will be inefficient whenever the large increase in the size of the intervention is not compensated for by a large reduction in the rate of FGM.

⁹¹ Efferson et al. (2019: 9) briefly mention a similar possibility (based on a simpler representation of the intervention process, see Chapter 5). However, their focus was on circumstances under which increasing intervention size might *not* increase relative spillovers, rather than the possibility of gross inefficiency (i.e. ‘negative’ spillovers) with respect to intervention size. There is also no concept of ‘educational strength’ in their analysis, since preference change is assumed to be extreme if it occurs (see Chapter 5).

In terms of the above example, a trivial case would be if initial coalitions of 3% or 6% were both insufficient to generate any change (in which case, increasing the intervention size from 30% to 60% would have no benefit at all). This can easily occur if *localised deviance effects* are absent or strong *social hump* effects are present (see Failure Scenarios 4 and 5).

A more interesting example would be if an initial coalition of size 3% were sufficient to generate a stable coalition, and this generated some positive spillover into the rest of the population, but increasing the initial coalition size to 6% did little to increase this spillover.

This situation can also occur in the ABM model. It can occur if *only a subset* of actors in the population is ‘susceptible’ to the influence of a small coalition. The influence of a small coalition may ‘spread’ through this subset, but fail to penetrate further into the population. The extent of this spread might not improve greatly for small increases in coalition size (e.g. from 3% to 6%).

A straightforward example of where this might occur is if social interaction is relatively localised (see Failure Scenario 4) and the social network of the population is homophilic (as we would expect empirically, see Chapter 7). Here, the final coalition will tend to be restricted to a ‘susceptible’ subset of the social network. Its influence will spread to other ‘susceptible’ parts, but fail to penetrate parts of the network where actors are more resistant to abandoning FGM (i.e. are ‘out-of-reach’ of the intervention). As such, costly increases in the size of the coalition will not translate into large increases in the effectiveness of the intervention.

To demonstrate these potential ‘out-of-reach’ effects, and their relation to inefficiencies in intervention design, I conducted a simulated experiment with a 2 by 2 design. In all cases, I used interventions with weak educational strength ($z_2 = 10\%$) and random targeting of actors in the population. This means that there were four conditions, based on two variables, each with two levels (two possible values). The first variable was the intervention size (z_1). This was either 35% or 70%. The second variable represented the minimum probability that actors who opposed FGM ($H_i < 0$), after its educational effects, would join the intervention. This was set to either 0% or 100%. When it was 100%, any actor who was persuaded (post educational effects) that FGM is intrinsically negative, would join the initial coalition of abandoners. When it was 0%, the probability that actors would join the initial coalition increased linearly from 0% (when $H_i \approx 0$) to 100% as actors’ perception of the intrinsic

value of FGM approached: $-\hat{M}$, the minimum possible value. The results of these experiments are displayed in Figures 2.1-2.4.

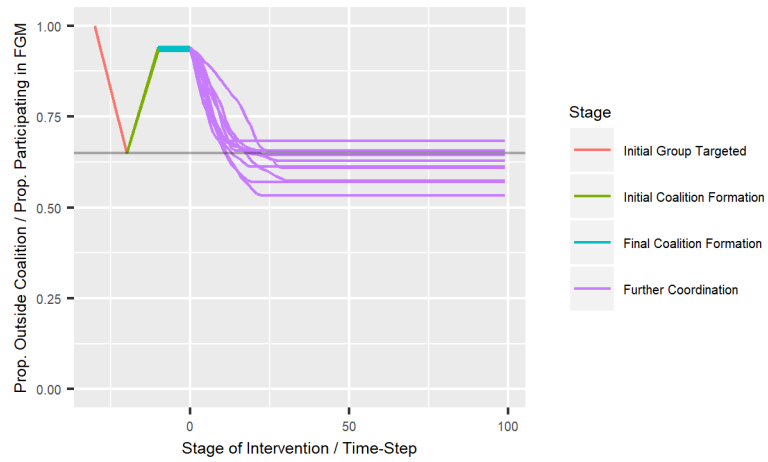


Figure F2.1: Simulated Intervention with Weak Educational Effect ($z_2 = 0.1$), Smaller Intervention Size ($z_1=0.35$) and a Minimum Probability that Opposers ($H_t < 0$) will Join the Initial Coalition of 0%

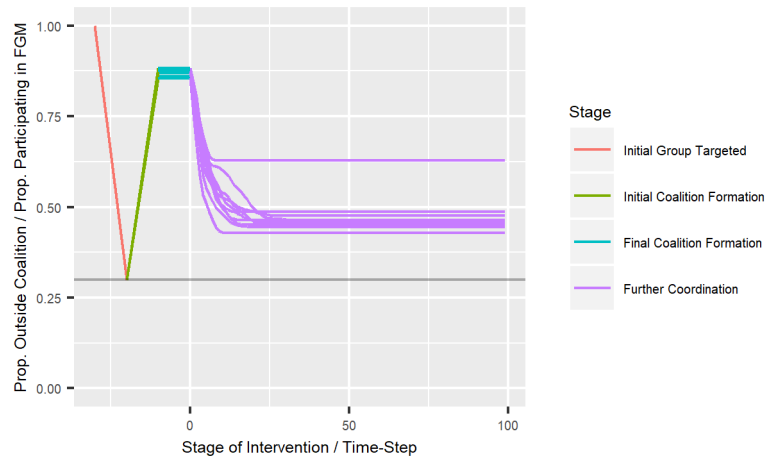


Figure F2.2: Simulated Intervention with Weak Educational Effect ($z_2 = 0.1$), Larger Intervention Size ($z_1=0.70$) and a Minimum Probability that Opposers ($H_t < 0$) will Join the Initial Coalition of 0%

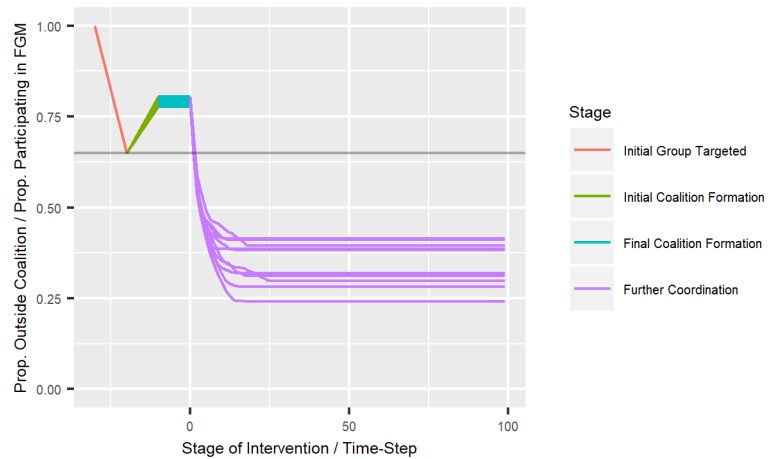


Figure F2.3: Simulated intervention with weak educational effect ($z_2 = 0.1$), smaller intervention size ($z_1 = 0.35$) and a minimum probability that opposers ($H_i < 0$) will join the initial coalition of 100%

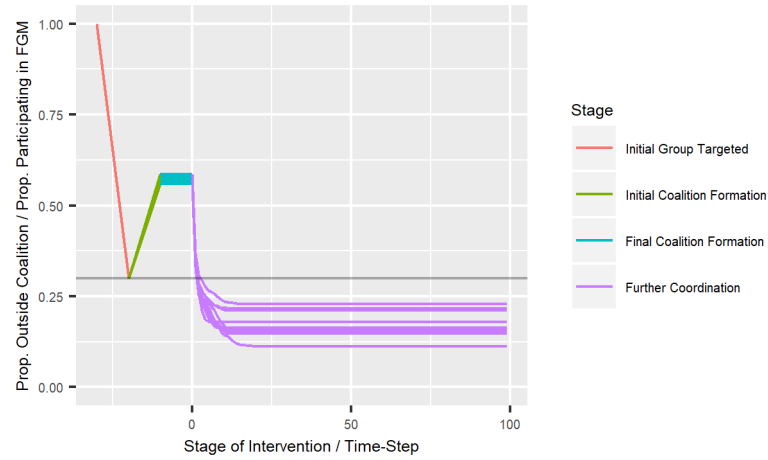


Figure F2.4: Simulated Intervention with Weak Educational Effect ($z_2 = 0.1$), Larger Intervention Size ($z_1 = 0.70$) and a Minimum Probability that Opposers ($H_i < 0$) will Join the Initial Coalition of 100%

Figure F2.1 illustrates the effects of a moderately sized intervention (35%) with weak educational strength (10%), and a strong relationship between the extent of actors' opposition to FGM, and their willingness to join the initial coalition. Here we see that the initial coalition is small, but its effects are able to spread to other parts of the population through coordination dynamics. In many cases, negative efficiencies are avoided, and the reduction in the rate of FGM is greater than the initial size of the intervention.

The intervention displayed in Figure F2.2 has the same parameters, but the intervention size is increased to 70%. Here we see that despite the large increase in the size of the intervention (another 35% of the population), the corresponding increase in the initial coalition size is only a few percent. This change *does* result in an improvement in intervention outcomes, but it is not nearly enough to compensate for the increase in intervention size. In this condition, all of the interventions result in negative efficiencies. Figures F2.3 and F2.4 illustrate the relationship between this effect and the recruitment process of the intervention. When the probability of joining the initial coalition is raised to 1 for all actors who oppose FGM, increasing intervention size results in a more substantive increase in coalition size. This, in turn, partially compensates for the previous inefficiencies.

So, out-of-reach effects could occur when there are constraints on the ability of an intervention to persuade actors to oppose FGM and to recruit persuaded actors to join an initial coalition of abandoners. They could be exaggerated when only a subset of actors is susceptible to the influence of small abandonment coalitions. The issues highlighted for the attention of practitioners are the extent to which they can reliably translate their educational

activities into the creation of an initial coalition who are conditionally committed to abandoning FGM. Also, the extent to which small increases in the size of this coalition will make a meaningful difference to the decisions of others to abandon.

Appendix F3: Failure Scenario 3 (Supporters Gambit Effects)

Supporters gambit effects, as I describe them here, are possible effects which have a similar basis to out-of-reach effects. However, they are distinct in that they relate to the impact of targeting bias and potential trade-offs between different biases. Supporters-gambit effects occur either when an intervention deliberately targets supporters of FGM, but this results in a failure to produce any change, or when interventions deliberately *avoid* targeting supporters of FGM (i.e. they target opposers of FGM), but this results in a failure of the intervention to produce complete abandonment (hence the ‘gambit’).

These contradictory possible outcomes from targeting supporters of FGM result from two countervailing effects. One is the tendency of supporters to help drive positive spillovers *when* enough of them are converted to opposers and recruited to a coalition of abandoners (see Chapter 5). The other is the difficulty of forming a large and influential coalition among those who are (initially) supporters of FGM.

Actors who support FGM will be harder to persuade to view FGM negatively (larger education strength is required). Even when they are persuaded, they may be less likely to join the initial coalition. It will also be harder to form a *stable* coalition from an initial group involving those who oppose FGM because these actors will be the least willing to forgo the initial social pressures to maintain the practice. Finally, if only a small stable coalition of (past) FGM supporters does form, then under network homophily, it may be harder for the influence of this coalition to spread to other parts of the network, because the social contacts of coalition members will be among the least willing to abandon FGM. However, if enough supporters are converted to opposers, and the size of their coalition is large enough, this can lead to large positive spillovers in the rest of the community.

Conversely, these effects are not present when targeting actors who *oppose* FGM. These actors are the most likely to join the initial coalition. They are also the most able to form a stable coalition. Furthermore, under network homophily, coalition members will be preferentially connected to actors who are also more willing to abandon FGM. Thus, there may be more immediate potential for the influence of a small coalition to spread. However, in this case, larger coalitions may have less beneficial effects (since many supporters of FGM

will remain, perhaps sheltered in a different part of the network), resulting in failure of the intervention process to generate widespread change.

So, there is a tension between the possible benefits of recruiting FGM-supporters to abandonment coalitions (which could improve the resulting spillover effect), and the potential difficulties of forming the large stable coalition of these actors needed to generate such effects.

Simulated experiments show that the balance of these risks and benefits may depend crucially on the educational strength of the intervention. As a demonstration, I conducted another 2 by 2 experiment. This involved two parameters, each with two levels. The first parameter was educational strength (z_2). This was set as either a 25% or 75% reduction in actors' perception of the value of FGM. The second parameter was the targeting strategy used in the intervention (z_6). This was set as either preferentially targeting actors most willing to abandon FGM (lowest H_i first) or actors least willing to abandon FGM (highest H_i first). In both cases, the intervention size was 40%. The results are displayed in Figures F3.1 to F3.4.

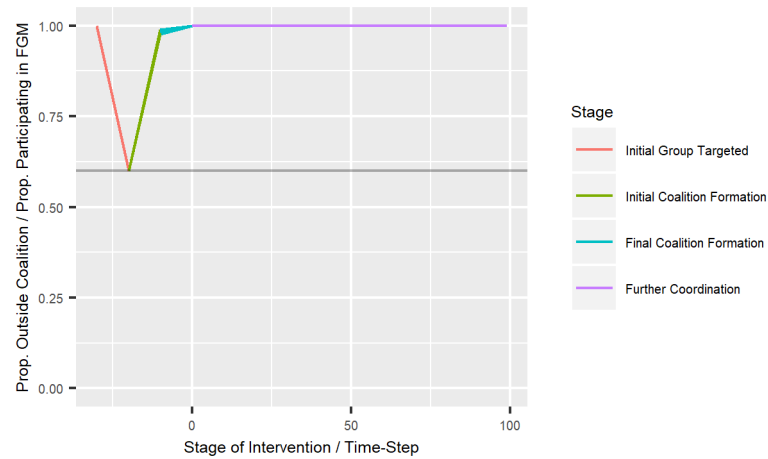


Figure F3.1: Simulated Intervention Targeting Supporters of FGM, with Weak Educational Strength ($z_2 = 0.25$)

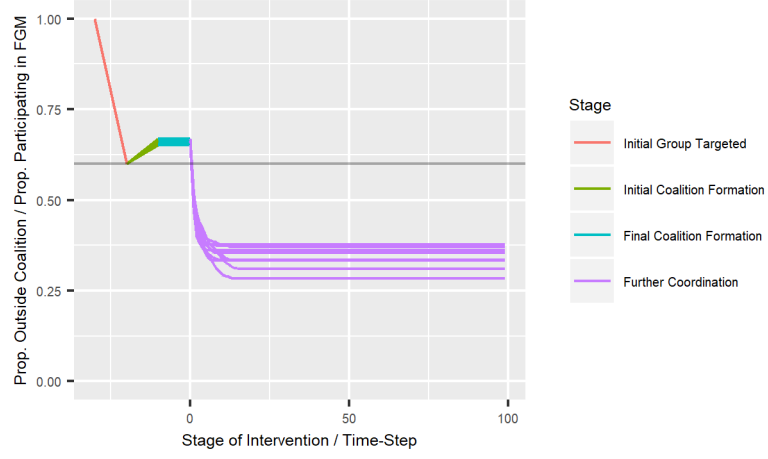


Figure F3.2: Simulated Intervention Targeting Opposers of FGM, with Weak Educational Strength ($z_2 = 0.25$)

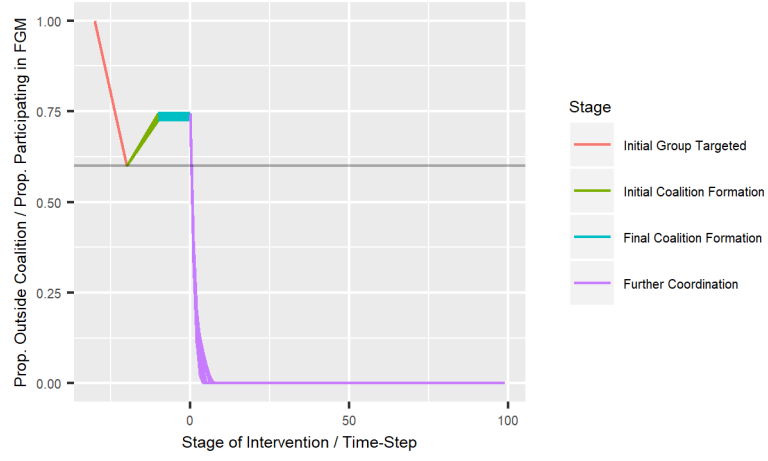


Figure F3.3: Simulated Intervention Targeting Supporters of FGM, with Strong Educational Strength ($z_2 = 0.75$)

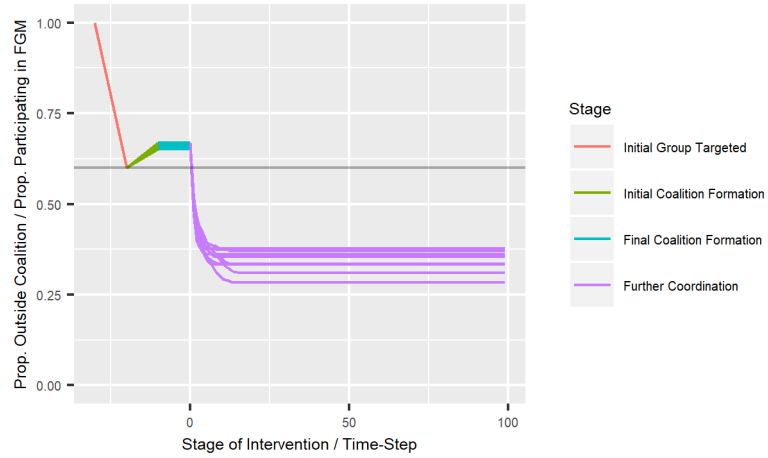


Figure F3.4: Simulated Intervention Targeting Opposers of FGM, with Strong Educational Strength ($z_2 = 0.75$)

These simulations demonstrate that when the educational strength of the intervention was weak, and supporters of FGM were targeted, only a small initial coalition was able to form.

Moreover, this initial coalition was unstable and therefore collapsed. The result was complete intervention failure, with no change. When the same weak intervention was applied to opposers of FGM, a much larger initial coalition was recruited, and it was more likely to remain stable. Furthermore, the influence of this coalition was better able to spread into additional parts of the population.

However, when the educational strength of the intervention was high, the pattern of failure was reversed. Both targeting strategies resulted in a stable coalition of abandoners. Yet, despite forming a larger coalition, the influence of interventions targeting opposers of FGM failed to spill over into the entire population. Its impact was limited to between 50% and 75% abandonment. By contrast, despite forming a smaller coalition, the influence of interventions targeting supporters of FGM ‘spilt over’ into the entire population – leading to complete abandonment.

Appendix F4: Failure Scenario 4 (Localised Deviance Effects)

The *localised deviance* failure scenario relates to the structure of social relationships in the local population and the way that this might affect intervention failure. Almost all formal analyses of FGM have assumed global interaction (or some form of random pairing, which we can reject out of hand), with the exception of Efferson et al., 2019 and early outputs of this project, see Appendix G). This means that all actors in the local population interact with all others. In the context of coordination dynamics, this means that all actors coordinate with all other actors in the local population. One of the key features of global interaction is that (all other things being equal) the coordination incentives faced by actors in the population are the same at a given moment. So, for example, if 80% of the population practices FGM, all actors in the population will experience social incentives associated with 80% of their ‘reference group’ practicing FGM.

On the other hand, when the interaction is localised, this property breaks down. Interaction is localised if some actors in the population only interact with a subset of others. Thus, their ‘reference groups’ only consists of a subset of the local population. Under localisation, while, for example, 80% of the population may practice FGM, the rate of FGM in the reference group of a *particular actor* might be above or below this number. As such, localised interaction can introduce *heterogeneity of coordination incentives* into the population.

Localised deviance, as I define it here, occurs when this heterogeneity of coordination incentives is sufficiently large that some coordinating actors make choices that they would *not* have

made if the interaction were global. In other words, localised deviance occurs when clusters of actors coordinate with one another, rather than with the wider population.

Broadly speaking, we can expect some potential for localised deviance effects to occur whenever there is heterogeneity of coordination incentives in the population. Straightforwardly, this could occur if actors primarily coordinate with key social contacts within their social network. The less interconnected and more clustered this network, the greater the potential for heterogeneity of coordination incentives. However, other factors could create or increase this heterogeneity of incentives. It can be exaggerated by homophily within the network or by coordination within households (see below for more details).

Rather like the supporters-gambit effects (see above) and social-hump effects (see below), localised deviance can induce trade-offs between different kinds of intervention failure. Put simply, when localised deviance is low, this can lead to complete intervention failure. This can occur because the intervention is unable to sufficiently alter the balance of social incentives across the population to allow a stable abandonment coalition to form or spread. Conversely, when localised deviation is high, stable coalitions may be able to form in ‘pockets’ of the social network and spread to other receptive parts of the network. However, they may be unable to spread through the entire network since localised ‘pockets’ which are resistant to abandoning FGM will persist.

These effects are demonstrated through a 2 by 2 simulated experiment using the ABM model and choosing parameter values which (individually and in combination) either enhanced or suppressed localised interaction. The first pair of conditions relates to the localisation of social interaction in the model. This was either high:

- Social reference network connectivity was set at (on average) 10 connections per actor ($\mu_{Rsocial} = 10$)
- Homophily of the network was high ($\rho_{Hy} = 0.991$)
- Actors refer, with equal weight, to their social and household reference networks ($s_3 = 0.5$)

Or localisation was low:

- Social reference network connectivity was set at (on average) 100 connections per actor ($\mu_{Rsocial} = 100$)
- Homophily of the network was low (but within the calibrated range, $\rho_{Hy} = 0.2$)

- Actors refer only to their social reference networks ($s_3 = 0$)

The second pair of conditions was related to the size of the intervention. This was either 10% or 40%. The results of these experiments are displayed in Figures F4.1-F4.4.

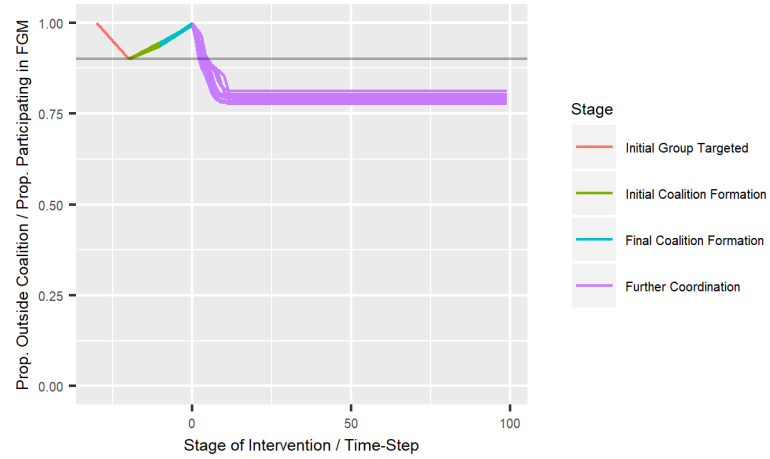


Figure F4.1: Simulated Interventions with High Localisation and Small Intervention Size ($z_1 = 0.1$)

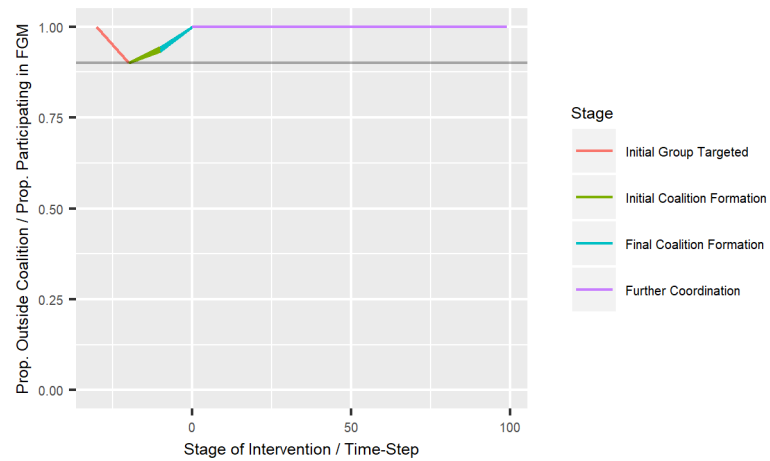


Figure F4.2: Simulated Intervention with Low localisation and Small Intervention Size ($z_1 = 0.1$)

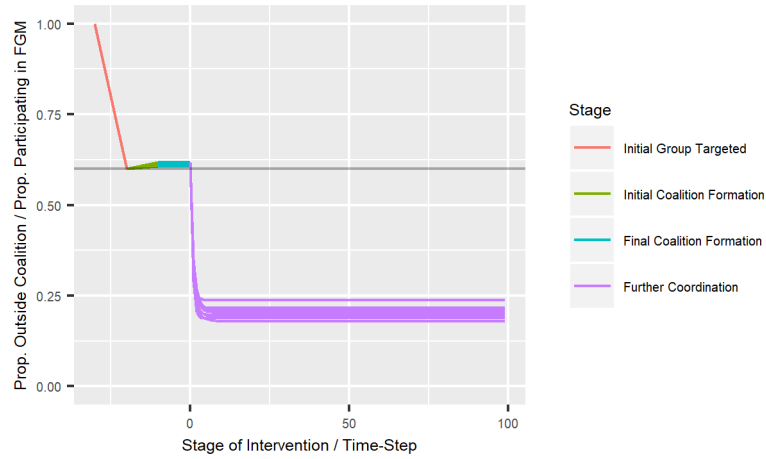


Figure F4.3: High Localisation with Larger Intervention Size ($z_1=0.4$)

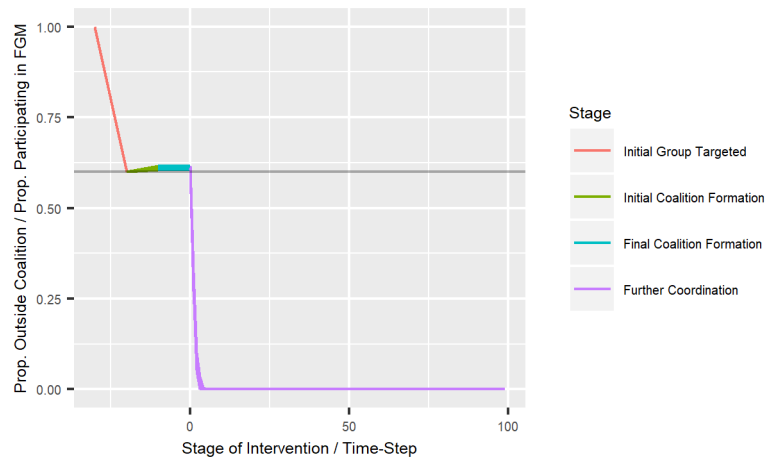


Figure F4.4: Low Localisation with Larger Intervention Size ($z_1=0.4$)

The results of these experiments demonstrate that when localisation is high, and the intervention size is small, a small coalition was able to form and, through coordination dynamics, its influence could spread to other parts of the local population. When localisation was low, and the intervention size was small, coalition formation failed entirely, leading to no change. However, when the intervention was larger, localised deviance became a source of intervention failure. Under high localisation, the effects of the intervention failed to spread through the entire population, leaving some still practicing FGM. Under low localisation, the influence of the intervention was able to spread more successfully, leading to complete abandonment of FGM within the population.

To my knowledge, this is the first formal analysis (in relation to the dynamics of FGM) to recognise the potential contradictory impacts of localised deviance in terms of preventing

complete intervention failure (i.e. preventing coalition collapse) *and* preventing complete intervention success (i.e. preventing complete abandonment)⁹².

Factors Affecting Localised Deviance

A number of factors can affect localised deviance in the ABM model of FGM as a social norm of coordination. The most obvious is the overall connectivity of the social reference network. As connectivity goes down, localised deviance goes up. However, localised deviance can also be increased by homophily of preferences within that network. Finally, localised deviance can be increased if actors refer strongly to their own households as part of decision-making. This occurs because households themselves are disconnected network cliques – allowing for stochastic variation in the social costs facing actors at the start of the simulation.

To illustrate the relations between localised deviance and the above features, we need to isolate the effect of social structure on the variability of social pressures faced by actors. We can begin with a case in which agents are homogeneous and highly interconnected so that they all face the same expected utility for practicing FGM (note that we are choosing arbitrary parameters for illustration purposes)

To achieve this, the simulation is set up with:

- 2000 agents
- Minimal homophily of preferences ($\rho_{Hy} = 0.001$)
- An arbitrarily high level of connectivity of the social reference network ($\mu_{Rsocial} = 500$)
- Implicit pressure from social reference network as the only source of social influence ($s_3 = 0, s_4 = 0, s_5 = 0$)
- Linear relations between social pressures and social costs ($V = 1$)
- Homogenous preferences centred on $H_i = 0$, such that all actors have a threshold of 50% ($s_2 = 1, s_1 = 1, \mu_H = 0.5, \sigma_H^2 = 0.00001$)
- Initially, 50% of agents practicing FGM (at random)

⁹² Efferson et al.'s (2019) focus exclusively on the effects of network structure on positive spillovers. Complete intervention failure cannot occur in their analysis, because they assume extreme preference change.

Setting up the simulation in this way, we find (Figure F4.5) a relatively homogenous distribution of the expected utility of agents, clustered around 0 (as we would expect).

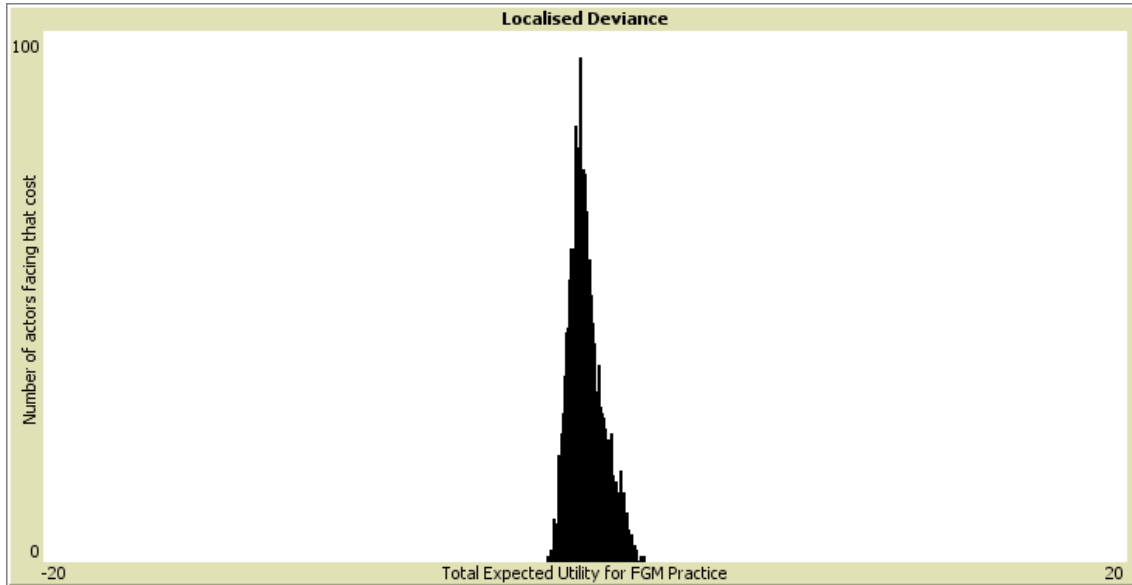


Figure F4.5: Distribution of Expected Utility for FGM Practice at Simulation Setup (homogenous preferences $\sigma_H^2 = 0.0001$ and high network connectivity $\mu_{Rsocial} = 500$)

This is our benchmark for localised deviance. Introducing the factors noted above induces greater variation in these costs, allowing for (potentially) greater heterogeneity of decision-making within the population (the potential policy implications of this are discussed further in the main text). Figure F4.6 shows the distribution of expected utility, after reducing network connectivity to $\mu_{Rsocial}=20$ (rather than 500). Figure F4.7 shows the distribution of expected utility after *also* allowing actors to refer strongly to their household reference group ($s_3 = 0.8$). These show that reducing network connectivity greatly increases variation in the expected utilities of actors (which is driven by variation in social pressures). Adding a strong influence of household reference groups exacerbates this effect further.

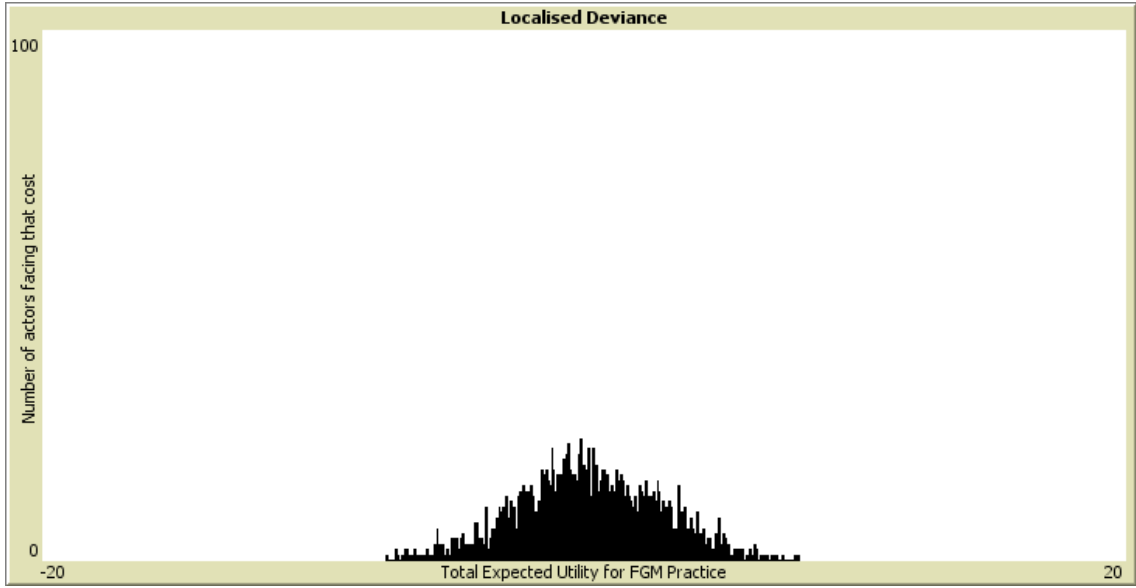


Figure F4.6: Distribution of Expected Utility for FGM Practice at Simulation Setup (homogenous preferences $\sigma_H^2 = 0.0001$ and low network connectivity $\mu_{Rsocial} = 20$)

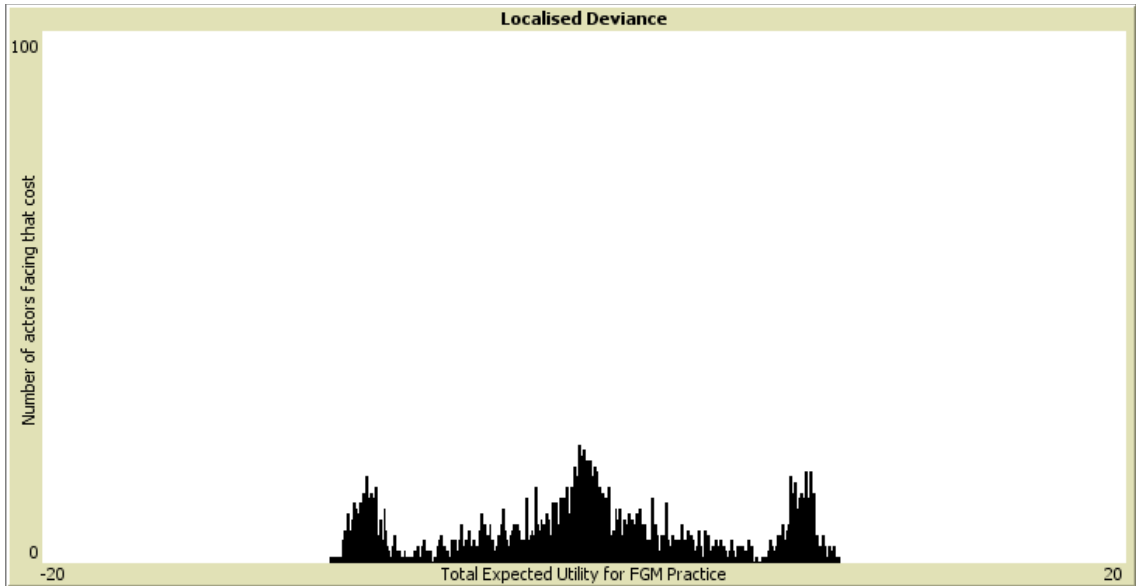


Figure F4.7: Distribution of Expected Utility for FGM Practice at Simulation Setup (homogenous preferences $\sigma_H^2 = 0.0001$, low network connectivity $\mu_{Rsocial} = 20$ and a strong influence of the household reference network $s_3 = 0.8$)

To also illustrate the effects of homophily of preferences on localised deviance, we need to go beyond the *setup* of the simulation. This is because the effect of homophily of preferences is that the social contacts of agents will *react* in different ways to FGM practice by their peers. Under homophily of preference, agents who are reluctant to abandon FGM will also be connected to others who are similarly reluctant, and vice-versa. The effect is that localised deviance ‘forms’ in the network quickly once the simulation is run.

To actually allow homophily of preferences, we also need to allow variation of preferences (H_i), so we start by increasing the variance of this distribution ($\mu_H = 0.5, \sigma_H^2 = 0.031$). Given this distribution of preferences, Figures F4.8 and F4.9 show examples of the expected utility of actors after a 2 steps of the simulation under a network connectivity of 20, with and without homophily of preferences ($\rho_{Hy} = 1, \rho_{Hy} = 0.001$, respectively). Other parameters are the same as in Figure 4.6 (above).

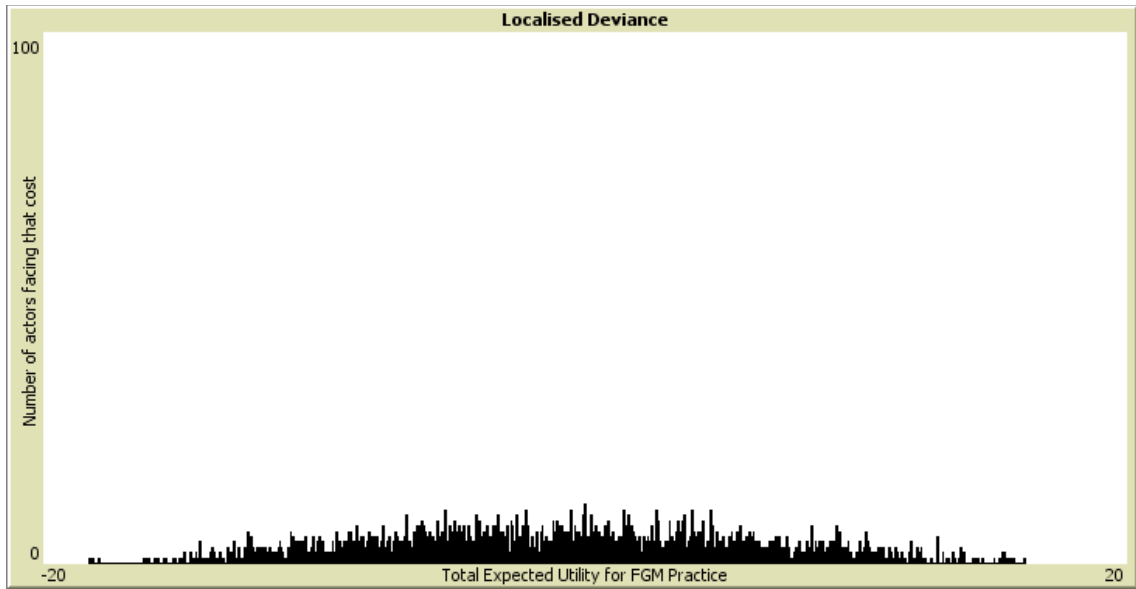


Figure F4.8: Distribution of Expected utility for FGM Practice at Time 2 (heterogeneous preferences, low network connectivity, $\mu_{Rsocial} = 20$ and low homophily of preferences $\rho_{Hy} = 0.001$)

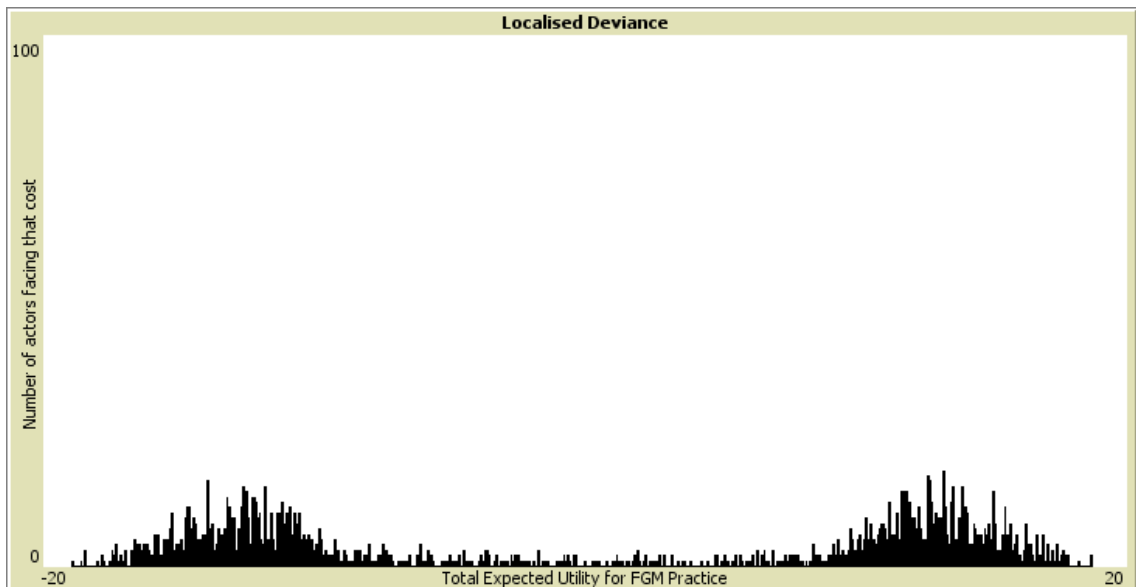


Figure F4.9: Distribution of Expected Utility for FGM Practice at Time 2 (heterogeneous preferences, low network connectivity, $\mu_{Rsocial} = 20$ and high homophily of preferences $\rho_{Hy} = 1$)

Appendix F5: Failure Scenario 5 (Social-Hump Effects)

Parameter estimates suggest that the average approval of FGM within communities where FGM is prevalence (μ_H), is clustered between 0.4 and 0.6 (see Figure 43, Chapter 7), and the maximum approval of FGM is close to the maximum opposition to FGM ($s_2 \approx 1$). Under this circumstance, the distribution of approval of FGM may well be roughly symmetric. On average, actors have relatively neutral view of FGM, whilst some actors have strongly positive, or strongly negative, views of FGM.

Under these circumstances, past analyses of coordination models suggest that strong *variation in* actors' willingness to abandon FGM is frequently bad news for intervention efforts (all other things being equal). As variation in the perceived value of FGM increases, the potential for large positive 'spillovers', in which the population tips into abandoning FGM, are reduced (compare normal, uniform and U-shaped distributions in Chapter 2).

However, I revisit this interpretation here for two reasons. First, the ABM model represents a process of coalition formation, rather than assuming that at least some participants unconditionally abandon FGM. This raises questions about whether low variation in preferences might create a *barrier* to initial change efforts. Second, in Chapter 7, I obtained empirical estimates of the community-level variation in actors' perceived value of FGM. I found that this community-level variation (measured as variance scaled to the unit interval) is typically between 0.01 and 0.07 (see Figure 43, Chapter 7). This is much narrower than the theoretically possible range, which could be arbitrarily close to 0, or as high as 0.25.⁹³

Exploration of the ABM suggests that these naturally occurring differences in the variability of preferences can still have significant implications for intervention failure. They also show that these effects can involve trade-offs between different kinds of intervention failure.

I refer to the effects of changes in the variability of actors' willingness to abandon FGM, as social-hump effects. Put simply; a social-hump may exist when the distribution of approval of FGM is relatively symmetric (see above) and most actors have a similar level of willingness to abandon FGM (variation is low). Under a standard threshold model, this would imply a threshold-distribution which is narrow and unimodal. The effect of the social hump is to make it difficult to form a small stable coalition and/or difficult for the influence of that small coalition to spread to population. However, the spread of influence from larger

⁹³ Based on the Popoviciu's Inequality on variances

coalitions is enhanced and may more easily reach the entire population. Under a standard threshold model, this effect can be understood in terms of a cumulative distribution of thresholds which crosses the diagonal from below (see Chapter 2).

The opposite of the social hump is a situation in which many actors are strongly opposed to FGM, and many others are strongly in favour of it. Under a standard threshold model, this would correspond to a threshold distribution which is more spread-out (like the uniform distribution) or even bimodal (like the U-shaped distribution). In the *absence of* a social hump, it may be easier to form a small coalition, and this coalition may have some further influence on the population. However, the influence of the coalition, even if large, may not spread widely.

These effects are demonstrated through a 2 by 2 experiment in Figures F5.1-5.4. I manipulated the variability of preferences in the model (σ_H^2), setting this at the low end ($\sigma_H^2=0.011$) or high end ($\sigma_H^2=0.071$) of empirically supported values. I also varied the size of intervention between 20% and 35%.

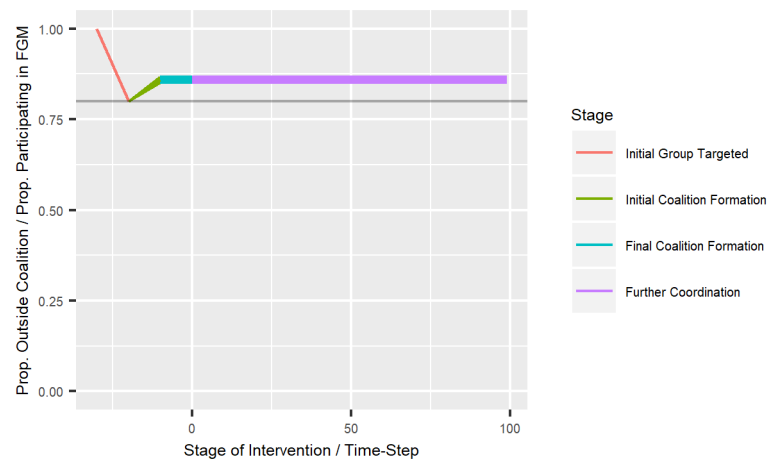


Figure F5.1: Simulated Intervention with High Social Hump ($\sigma_H^2=0.011$) and a Small Intervention Size ($z_1=0.2$)

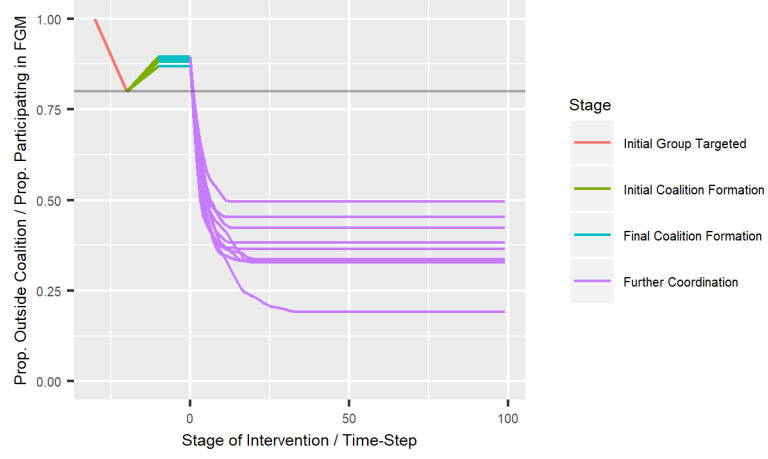


Figure F5.2: Simulated Intervention with Low Social Hump ($\sigma_H^2 = 0.071$) and a Small Intervention Size ($z_1 = 0.2$)

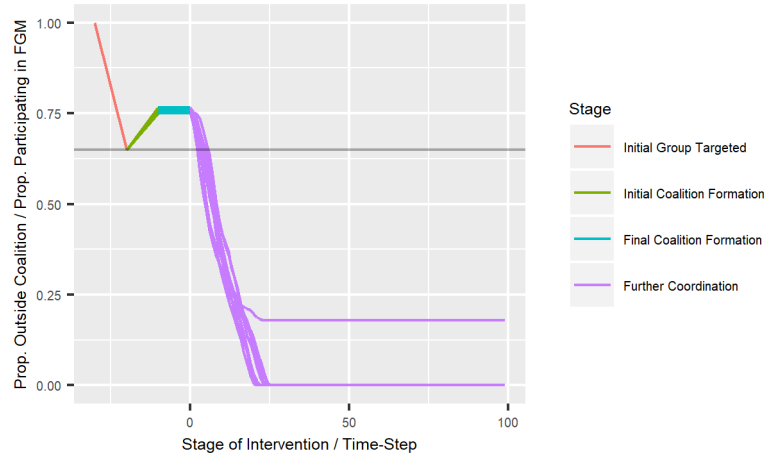


Figure F5.3: Simulated Intervention with High Social Hump ($\sigma_H^2 = 0.011$) and a Larger Intervention Size ($z_1 = 0.35$)

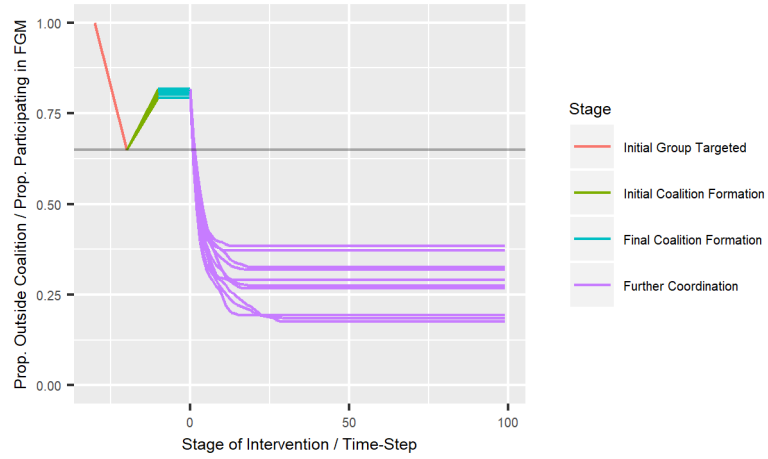


Figure F5.4: Simulated Intervention with Low Social Hump ($\sigma_H^2 = 0.071$) and a Larger Intervention Size ($z_1 = 0.35$)

The results show that when the variability of preferences is at the lower end of empirically supported values, and the intervention is small, this can create a social hump effect (Figure F5.1). The result is that the influence of the coalition is unable to spread to other parts of

the population. Conversely, when the variance is at the upper end of supported values (so the social hump is removed), change is able to spread from the coalition to other parts of the community (F5.2). Thus, when the intervention is small, the social hump precipitates intervention failure. However, when the size of the intervention is increased, the situation is reversed. When a social hump is present, the influence of the coalition can successfully spread through the whole community. When it is absent, the extent of this influence is limited, and some of the community continues to practice FGM.

These observations are in-line with past analyses. However, the analysis here demonstrates the potential importance of differences in variation of preferences within an empirically supported range of parameter settings. It is also important to note that low variability in actors' perceptions of the value of FGM is only *one* possible way that a social hump effect can be created. Variation autonomy has the potential to create or remove social hump effects, independently of changes in the distribution of actors' preferences. This can occur because of the relationship between autonomy and actors' willingness to abandon FGM and because of the relationship between autonomy and the balance of social costs for practicing versus abandoning FGM (see below for details).

The Relationship between Heterogeneity of Autonomy and Threshold Distributions

The relationship between individual autonomy and threshold distributions in the simulation model is complex and not easy to account for in a narrative way. In fact, strictly speaking, agents only have single 'thresholds' if social influence is implicit (i.e. no norm enforcement). This is the scenario I focus on here. However, we can think of thresholds as an indication of the general willingness of the actor to abandon FGM (higher means more willing). Thus the effects of heterogeneous autonomy on thresholds extent to their effects on the distribution of agents' 'willingness to abandon FGM' under norm-enforcement as-well.

When heterogeneity is present in the general model, there is no closed-form representation of the threshold. Also, things are further complicated by the fact that social influence is 'weighted', which means any threshold would be weighted proportion.

Nevertheless, thresholds remain a useful heuristic for understanding the dynamics of the model. In order to demonstrate the various possible effects of heterogeneous authority on the model, I distinguish between a 'simple' and a 'complex' threshold for actors.

I define the ‘simple’ threshold ($p_{i \text{ simple}}^*$) as the value of pro-FGM activity (p_{pro_i}) at which an actor would practice FGM, assuming a linear relation between pro-FGM activity and social costs for abandoning:

$$U(abandon)_i = p_{pro_i}^1 \cdot -\hat{M}$$

$$U(practice)_i = H_i - (s_1 \cdot p_{anti_i}^1 \cdot \hat{M})$$

And assuming that anti-FGM activity is the relative complement of pro-FGM activity ($p_{anti_i} = 1 - p_{pro_i}$), which is true when social influence is implicit. This ‘simple’ threshold has a simple closed-form representation and so is straightforward to calculate for each agent.

I define the ‘complex’ threshold ($p_{i \text{ complex}}^*$) in the same way, *except* I don’t assume a linear relationship between social pressure and social costs, instead, the full utility function is used:

$$U(abandon)_i = p_{pro_i}^{V(-\delta_1 + \alpha_i)} \cdot -\hat{M}$$

$$U(practice)_i = H_i - (s_1 \cdot p_{anti_i}^{V(-\delta_2 + \alpha_i)} \cdot \hat{M})$$

Noting that we assume $p_{anti_i} = 1 - p_{pro_i}$, still.

This complex threshold ($p_{i \text{ complex}}^*$) therefore takes into account potential non-linearities in the relation between social process and social costs. $p_{i \text{ complex}}^*$ is still defined as the value of p_{pro_i} at which actor i would practice FGM. However, there is no closed-form definition available. Instead, however, the value for each agent can be found algorithmically, by arbitrarily ‘testing’ values of p_{pro_i} between 0 and 1 (in ascending increments of $\frac{1}{1000}$) until the expected utility for agent i practicing FGM exceeds the expected utility for abandoning. This will provide an estimate of $p_{i \text{ complex}}^*$ to within $\frac{1}{1000}$ for each agent. An algorithm of this kind is implemented in the ABM in order to explore the issues discussed here.

Since the distribution of $p_{i \text{ simple}}^*$ represents the distribution of thresholds *without* taking into account heterogeneity of autonomy, and the distribution of $p_{i \text{ complex}}^*$ represents the distribution of thresholds *with* heterogeneity of autonomy taken into account, we can compare the distribution of the two metrics to find the ‘effect’ of heterogeneity of autonomy on the distribution of thresholds (note that we will need to set δ_1 and δ_2 to 0 to avoid other effects on the threshold distribution).

Autonomy is an agent attribute in the model (α_i) which affects the relationship between the social pressure felt by agents, and the social costs they experience. The basic intuition is as follows. If agents are more autonomous, then the social costs they experience for practicing or abandoning FGM rise more slowly as the level of social pressure on them increases. Conversely, if agents are less autonomous, the social costs they experience as the level of social pressure increases rise *more quickly*. This is modelled using a non-linear function to connect the level of social pressure experienced by the agent [0,1] to the level of social costs they experience [0,1] (noting that social costs are then scaled according to other parameters in the model).

When there are only social costs for abandoning FGM ($s_1 = 0$ in the simulation) then the effect of autonomy on the distribution of thresholds is relatively straightforward. If autonomy is strongly (positively) correlated with approval of FGM (H_i), such that those who approve of FGM are the most autonomous, then the effect will be ‘quash’ the distribution of threshold (reduce variation). This happens because those who approve of FGM become less susceptible to social pressure to maintain the practice (so their low thresholds go up), whereas those who disapprove of FGM become more susceptible to social pressure to maintain the practice (so their high thresholds go down).

In other situations, such as when autonomy is uncorrelated with approval of FGM, or there are also social costs for *practicing* FGM, the effect is less intuitive.

However, one can classify the effects of heterogeneity-of-autonomy into at least four categories, based on the positive correlation between autonomy and approval of FGM and the whether or not there are symmetric social pressures to *abandon* FGM. A summary of these effects is given in Table F5.1, with illustrations of each scenario in Figures F5.5-F5.12.

Table F5.1: The Complex Relation Between Heterogeneity of Autonomy and Threshold Distributions

<p>Common Assumptions:</p> <p><i>High potential non-linearity of the relation between social pressures and social costs ($V = 4$), with a balance between practice and abandonment ($\delta_1 = 0, \delta_2 = 0$).</i></p> <p><i>Autonomy has a roughly uniform distribution ($\mu_\alpha = 0.5, \sigma_\alpha^2 = 0.071$)</i></p> <p><i>$H_i$ is ‘normally’ distributed with $\mu_H = 0.5$ and $\sigma_H^2 = 0.031$</i></p>	<p><i>No Social Costs for Practicing FGM ($s_2 = 0$) and no actors preferring FGM intrinsically ($s_1 = 0$)</i></p>	<p><i>Symmetric Social Costs for Practicing FGM ($s_2 = 1$) and many actors preferring FGM intrinsically ($s_2 = 1$)</i></p>
--	---	--

Positive correlation with H_i : low ($\rho_{H\alpha} = 0.001$)	Effect Relative to 'Simple' Distribution: Spreading of thresholds, particularly towards low thresholds (more willing to practice FGM)	Effect Relative to 'Simple' Distribution: Moderate spreading of thresholds (relatively symmetric)
Positive correlation with H_i : high ($\rho_{H\alpha} = 0.991$)	Effect Relative to 'Simple' Distribution: 'Quashing' of thresholds around a central point	Effect Relative to 'Simple' Distribution: Spreading of thresholds, particularly towards high thresholds (less willing to practice FGM)

I am not aware of an intuitive explanation for these effects. However, the key point is a relatively simple one. Heterogeneity of autonomy could have profound and non-intuitive impacts on the distribution of thresholds within a community. This might engender the kind of 'social hump' effects described in Chapter 8 or even changes in the overall willingness of the community to abandon FGM. I do not know whether such effects occur in real communities or whether the formal representation is accurate. It's not clear how one *could* know without substantial empirical work. However, it is clear that *heterogeneity of autonomy* is an important and policy-relevant source of uncertainty in the design of coordination models of FGM. It is, no doubt, one of many sources of uncertainty that is only made visible by actually representing and exploring it in a formal model.

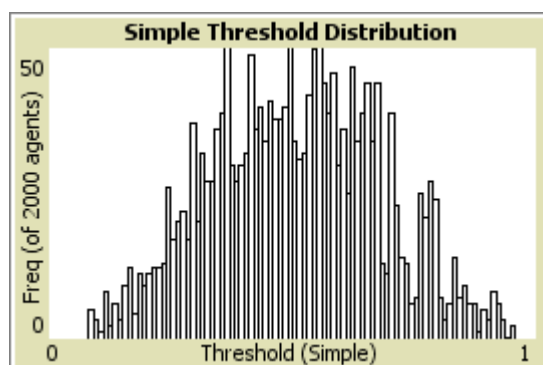


Figure F5.5: Simple Threshold Distribution: Positive correlation with H_i : Low ($\rho_{H\alpha} = 0.001$) and no social costs for Practicing FGM ($s_2=0$) and no actors preferring FGM intrinsically ($s_1=0$)

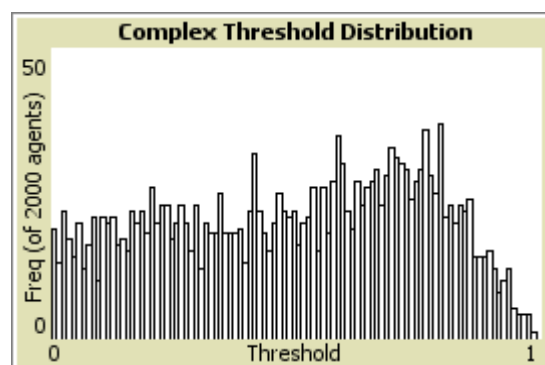


Figure F5.6: Complex Threshold Distribution: Positive Correlation with H_i : Low ($\rho_{H\alpha} = 0.001$) and No Social Costs for Practicing FGM ($s_2=0$) and no actors preferring FGM intrinsically ($s_1=0$)

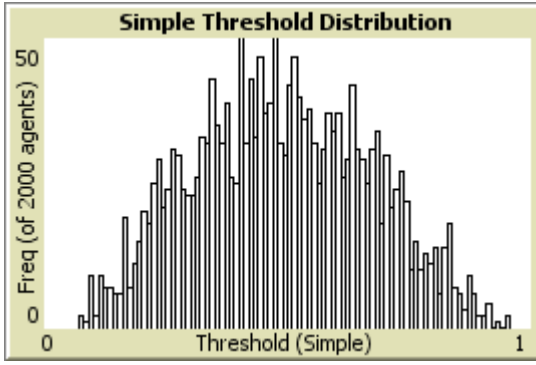


Figure F5.7: Simple Threshold Distribution: Positive correlation with H_i : High ($\rho_{Ha} = 0.991$) and no social costs for Practicing FGM ($s_2=0$) and no actors preferring FGM intrinsically ($s_1=0$)

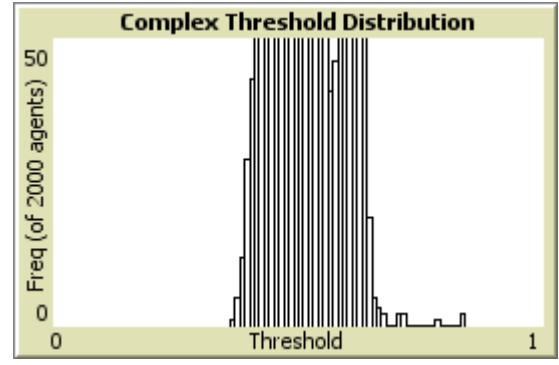


Figure F5.8: Complex Threshold Distribution: Positive correlation with H_i : High ($\rho_{Ha} = 0.991$) and no social costs for Practicing FGM ($s_2=0$) and no actors preferring FGM intrinsically ($s_1=0$)

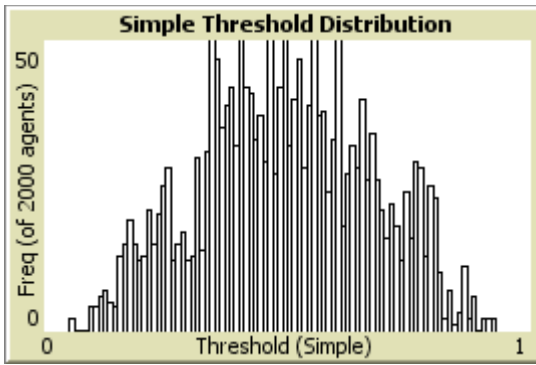


Figure F5.9: Simple Threshold Distribution: Positive correlation with H_i : Low ($\rho_{Ha} = 0.001$) and symmetric social costs for Practicing FGM ($s_2=1$) and many actors preferring FGM intrinsically ($s_1=1$)

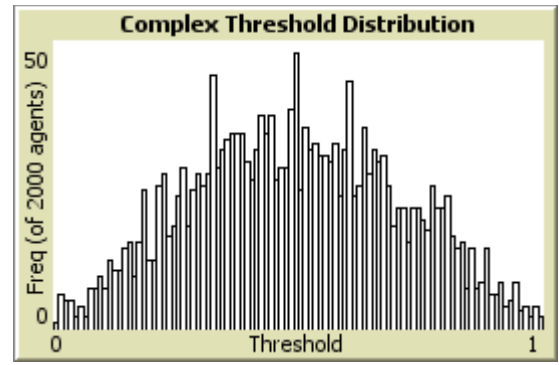


Figure F5.10: Complex Threshold Distribution: Positive correlation with H_i : Low ($\rho_{Ha} = 0.001$) and symmetric social costs for Practicing FGM ($s_2=1$) and many actors preferring FGM intrinsically ($s_1=1$)

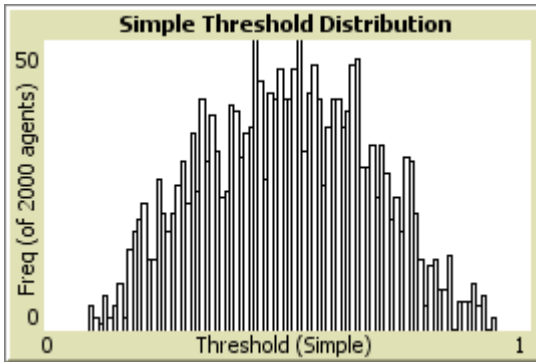


Figure F5.11: Simple Threshold Distribution: Positive correlation with H_i : High ($\rho_{Ha} = 0.991$) and symmetric social costs for Practicing FGM ($s_2=1$) and many actors preferring FGM intrinsically ($s_1=1$)

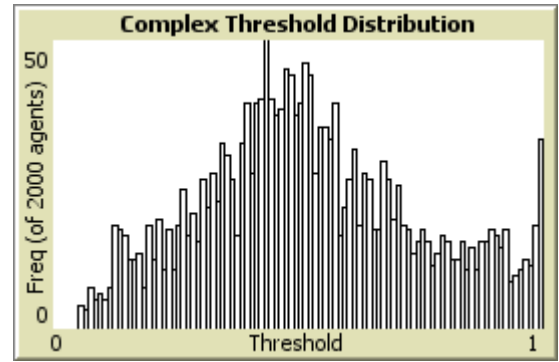


Figure F5.12: Complex Threshold Distribution: Positive correlation with H_i : Low ($\rho_{Ha} = 0.991$) and symmetric social costs for Practicing FGM ($s_2=1$) and many actors preferring FGM intrinsically ($s_1=1$)

Appendix F6: Failure Scenario 6 (Power Imbalance Effects)

In Chapter 5, I demonstrated that if actors vary in their authority within the community, and if this is positively correlated with their approval of FGM, this could influence the outcomes of simulated interventions. Here I extend this insight and note some of the factors which

might influence the magnitude of these effects. In particular, I show that their impact on intervention failure could range from mild to catastrophic, depending on these factors. Furthermore, I show that at least one ‘intuitive’ remedy to this effect: to focus the intervention on actors with the most authority, could be counter-productive.

Power imbalance effects are relatively straightforward to describe. If actors vary in how much authority they have over others, and the actors with the most authority value FGM the most, then this can obstruct the abandonment of FGM within the community. This occurs because the actors who influence others the most are also the least willing to abandon the practice. This, in turn, makes abandonment costly for actors who might wish to stop participating. In the model, authority is implemented through a weighting system, with actors with higher authority ‘weighted’ more heavily in the calculation of coordination incentives than actors with lower authority. The general ABM facilitates variation in three aspects of this ‘weighting’. First, variability in weights is controlled by a parameter σ_w^2 . The variability of weights can vary from narrow (actors all have similar weights) to highly dispersed (some actors have high weights; others have low weights). Second, the average weight in the community is controlled by a parameter μ_w . The average can vary from very low (close to zero) to very high (close to 1). Third, the positive correlation between weights and approval of FGM (H_i) is controlled by a parameter ρ_{Hw} . Approval of FGM can be uncorrelated with authority or highly correlated, such that actors who value FGM the most have the most authority.

Power imbalance effects can occur whenever there is a correlation between authority and approval of FGM. However, they can be strongly *exacerbated* by features of the *distribution* of authority. First, as variance in the authority of actors increases, differences in the influence of FGM supporters and opposers are increased. This exaggerates the influence of FGM supporters. Second, as the average authority of actors *decreases*, social influence is increasingly concentrated in the small number of actors with high authority *who are also supporters of FGM*. The effect is to further imbalance power in the direction of FGM supporters.

These effects are demonstrated through four sets of simulated interventions, depicted in Figures F6.1-F6.4. Each intervention increments the power imbalance effect. In the first set of interventions, power is (approximately) normally distributed with a central average ($\mu_w = 0.5$, $\sigma_w^2 = 0.031$) and is uncorrelated with approval of FGM ($\rho_{Hw} = 0.001$). In the second set of interventions, a strong correlation between power and approval of FGM is added ($\rho_{Hw} = 0.991$). In the third set of interventions, variation in authority is increased, such that

the distribution of authority becomes U-shaped ($\sigma_w^2 = 0.111$). In the fourth set of interventions, the average authority is reduced, so that the distribution of authority becomes L-shaped ($\mu_w = 0.2$). In all interventions, actors were targeted at random.

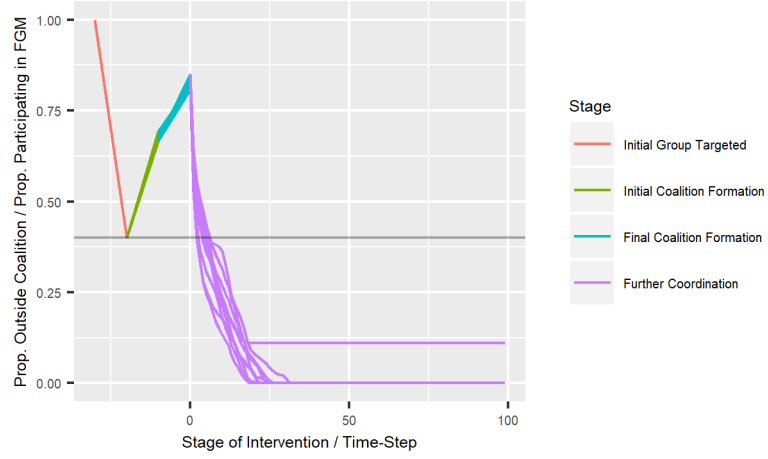


Figure F6.1: Simulated Intervention with 'Normally'-Distributed Authority ($\mu_w = 0.5, \sigma_w^2 = 0.031$), No Correlation with Approval of FGM ($\rho_{Hw} = 0.001$)

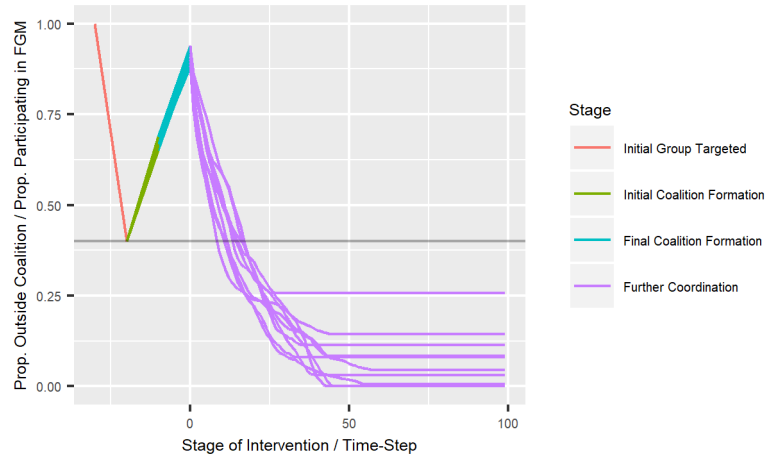


Figure F6.2: Simulated Intervention with 'Normally'-Distributed Authority ($\mu_w = 0.5, \sigma_w^2 = 0.031$), High Correlation with Approval of FGM ($\rho_{Hw} = 0.991$)

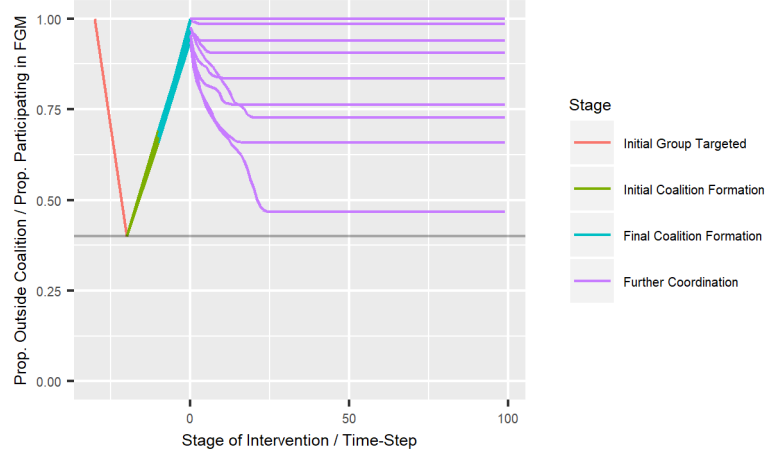


Figure F6.3: Simulated Intervention with High Variance of Authority ($\mu_w = 0.5, \sigma_w^2 = 0.111$) and High Correlation with Approval of FGM ($\rho_{hw} = 0.991$)

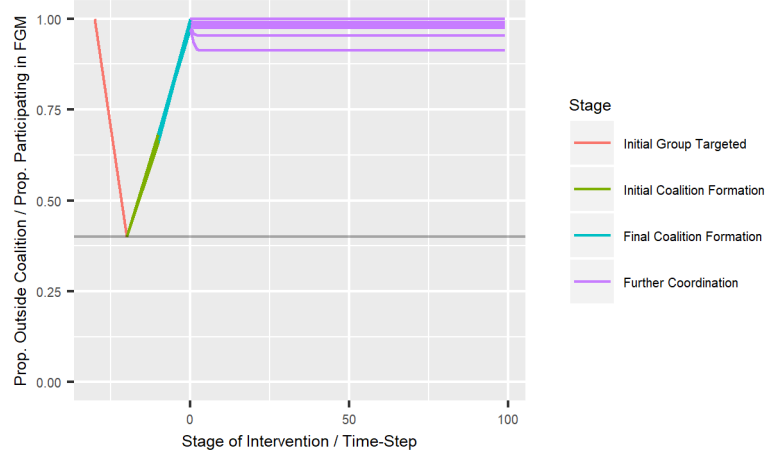


Figure F6.4: L-Shaped Distribution of Authority ($\mu_w = 0.2, \sigma_w^2 = 0.111$), High Correlation with Approval of FGM ($\rho_{hw} = 0.991$)

Figures F6.1-F6.4 clearly illustrate the potential cumulative effects of different aspects of power-imbalance on intervention failure. When authority is uncorrelated with approval of FGM, the intervention almost always results in complete abandonment of FGM. When a positive correlation is introduced, the impact of the intervention is slightly curtailed, with a larger number of interventions failing to produce complete abandonment. When the variation in authority is increased (creating a U-shaped distribution of authority), this curtailment becomes more severe, and introduces negative-spillovers, with the number of actors abandoning FGM fewer than the number targeted. When the average authority in the community is then reduced (creating an L-shaped distribution of authority) the intervention either fails completely or generates only a small reduction in FGM practice.

As noted in Chapter 5, power imbalance effects might well be present in real communities. Shell-Duncan et al. (2011), for example, emphasise the role that authoritative female elders play in perpetuating FGM. One ‘intuitive’ remedy to this situation might be to preferentially target actors who have the most authority in the community. Clearly, if these actors commit to abandoning FGM, then this may improve the influence of the abandonment coalition. However, as illustrated in Figure F6.5, there may be risks associated with this strategy. Figure 85 depicts a simulated intervention which preferentially targets actors with the highest authority in the community. This intervention takes place under conditions present in Figure 82 (normal distribution of authority, high positive correlation between authority and approval of FGM).

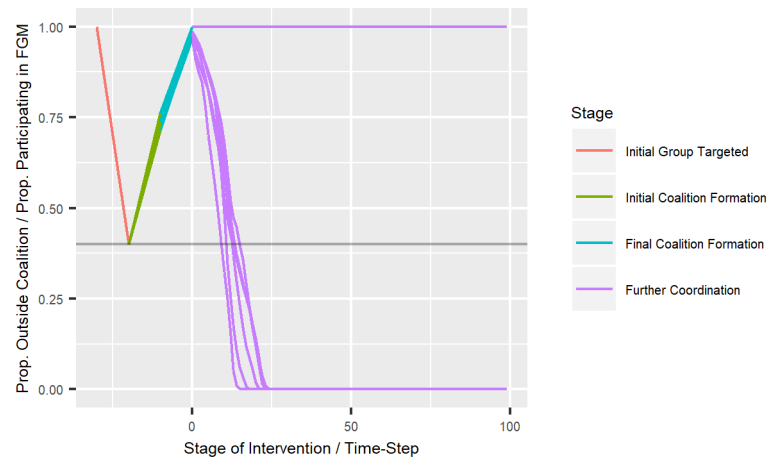


Figure F6.5: Simulated Intervention with ‘Normal’-Distribution of Authority ($\mu_H = 0.5, \sigma_H^2 = 0.031$), High Correlation between Authority and Approval of FGM (ρ_{Ha}), and Preferential Targeting of Actors with the Highest Authority

In Figure F6.5, we can see a kind of bifurcation of the outcome of the intervention. In some cases, the intervention produced complete abandonment. In others, it failed to produce any change. This occurred because the effects of targeting by authority *overlap* with the *supporters-gambit* effects described previously. Since actors with the highest authority are also supporters of FGM, they are hardest to recruit to an initial coalition, and it is more difficult for this coalition to stabilise. We can see that when a stable coalition *does* stabilise, the influence on the remaining population is enormous. However, interventions might run the risk that no coalition will stabilise at all.

Policy-makers will not be surprised by the suggestion that power imbalances in favour of supporters of FGM could make change more difficult to orchestrate. However, the formal analysis provided here demonstrates the substantive and cumulative role that different aspects of power imbalance could have on intervention outcomes. In particular, they show

that social dynamics could be highly sensitive to the *shape* of the distribution of authority in the community. They also highlight that, in the presence of imperfect interventions, attempts to remedy power imbalances by targeting actors with high authority run the risk of other failure scenarios – such as those associated with supporters-gambit effects.

Appendix F7: Failure Scenario 7 (Enforcement Imbalance Effects)

The final failure scenario outlined in this chapter deals with a feature of social dynamics which is acknowledged in theoretical and empirical discussions but which has been absent from formal analyses of FGM. This scenario involves an imbalance between the willingness of FGM supporters to enforce the practice and the willingness of FGM opposers to enforce the abandonment of the practice. The scenario pertains to situations in which the social incentives related to FGM practice stem primarily from explicit enforcement of practice or enforcement of abandonment by specific individuals in the community ($s_5 \approx 1$, see Chapter 6). The model assumes that a certain proportion of actors who support FGM are willing to explicitly enforce that norm (conditional on practicing FGM themselves). It also assumes that a certain proportion of those who oppose FGM are willing to explicitly enforce the abandonment of that norm (conditional on abandoning FGM themselves). Enforcement imbalance effects occur when there are many supporters of FGM willing to enforce the practice, but few opposers willing to enforce abandonment.

In the context of the model, this can be disastrous for intervention efforts. For a number of reasons, it makes an initial abandonment coalition difficult to form. First, the initial formation of a coalition may do little to relieve the social pressures on coalition members. This is because *those who agree to participate initially will tend to also be actors who oppose FGM*. As such, most coalition members will *already* not be enforcing the abandonment of FGM. Second, even when supporters are persuaded to join the initial coalition, only a proportion of these actors will have been enforcers, as such, only some actors will contribute to a reduction in the social pressures to practice FGM. Third, even if many coalition members were previously enforcers of FGM their (conditional) abandonment will only reduce the social pressure to practice FGM, it will not create new social pressures to abandon the practice (as happens under more generic representations of coordination dynamics). For the same reason, if a stable coalition *does* manage to form, this will, at best, reduce the social pressures on others to practice FGM. It will not create new social pressures on others to *abandon* FGM.

The potential effects of norm imbalance are demonstrated through a simulated experiment in Figures F7.1 and F7.2. Figure F7.1 depicts the outcome of simulated interventions with ‘balanced’ norm enforcement, such that around 50% of supports of FGM are willing to enforce the practice, and (after stopping) 50% of FGM opposers are willing to enforce abandonment. Figure F7.2 depicts the same intervention, but in a scenario in which no FGM opposers are willing to enforce abandonment. Both interventions target social actors at random.

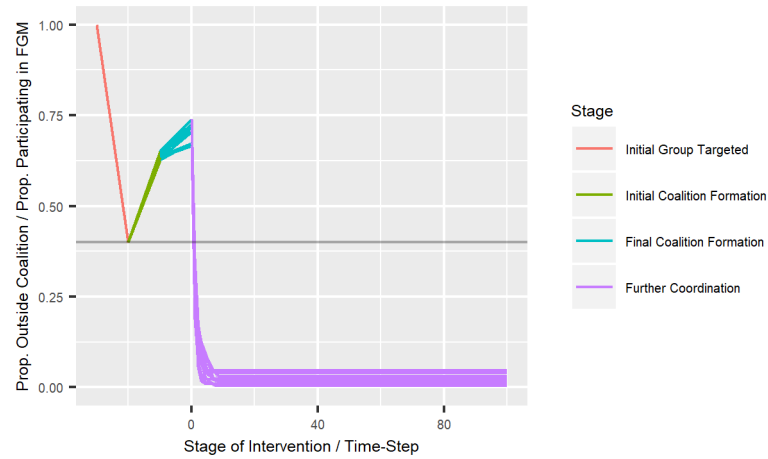


Figure F7.1: Simulated Intervention with Explicit Norm Enforcement and a Balance of Willingness to Enforce

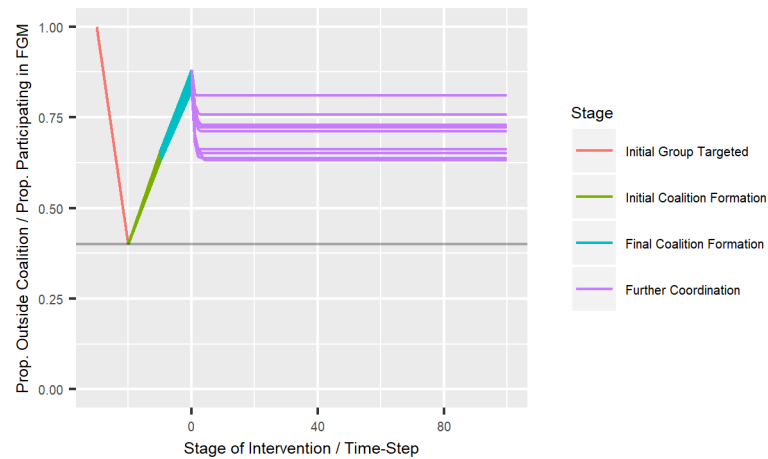


Figure F7.2: Simulated Intervention with Explicit Norm Enforcement, and Opposers Unwilling to Enforce Abandonment

When norm enforcement is balanced, most of the initial coalition remains stable, and the effects of the coalition spread throughout the community, leading to complete abandonment. We see that when norm enforcement is unbalanced, the stable coalition that is able to form is smaller, and its impact on the rest of the community is greatly reduced.

Theoretical discussions of social dynamics, in relation to intervention efforts, have emphasised the need to (a) reduce explicit enforcement of the ‘norm’ of FGM and (b) to create new social incentives against FGM practice. The formal analysis undertaken here provides an opportunity to distinguish between these two strategies and their possible advantages or disadvantages. In particular, analysis of the model suggests that reducing explicit enforcement of FGM practice, without introducing enforcement of the abandonment of FGM, could result in a failure of the intervention to generate complete abandonment of the practice.

This possibility can be demonstrated through simulated interventions, as depicted in Figures F7.3 and F7.4. These simulated interventions are identical to those depicted in Figures F7.1 and F7.2, except in two respects. The simulated interventions depicted in Figure F7.3 explicitly target pro-FGM enforcers for the intervention. The simulated interventions depicted in Figure F7.4 targeted actors at random but introduced an additional ‘effect’ of the intervention. This effect was to encourage a random 50% of initial coalition members to become anti-FGM enforcers (parameter $z_7 = 0.5$).

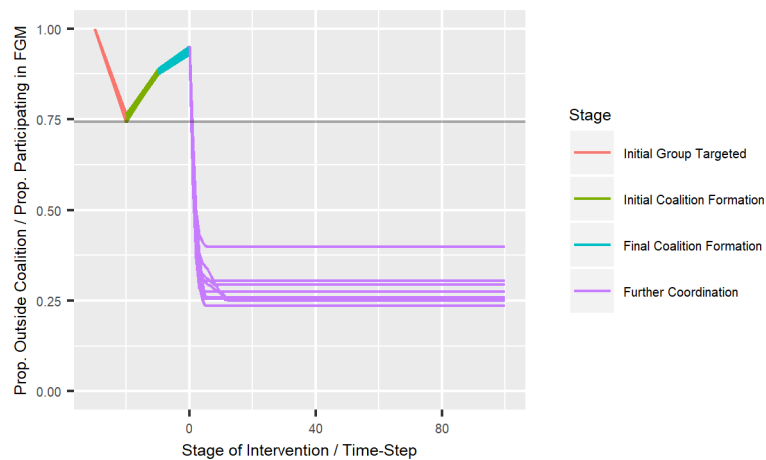


Figure F7.3: Simulated Intervention Targeting Pro-FGM Enforcers and No Creation of Anti-FGM Enforcers

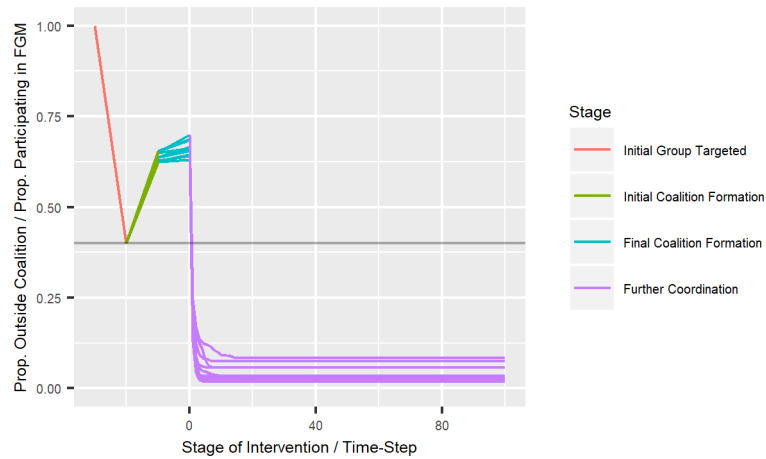


Figure F7.4: Simulated Intervention Targeting Actors at Random, with 50% of Coalition Participants becoming Pro-FGM Enforcers

We can see that if the intervention preferentially targets actors who are enforcers of FGM, the intervention becomes much more effective. This strategy maximises the reduction in social pressure to practice FGM achieved by the intervention. However, some actors continue to practice FGM. This occurs because the intervention has only reduced the social pressure to practice. It has *not* created new social pressures to abandon FGM. As such, the remaining supporters of FGM continue to prefer to practice. By contrast, when the intervention creates new social incentives to engage in the practice (by encouraging actors to enforce abandonment), its influence is able to spread to the entire population, leading to complete abandonment.

Appendix G: Early Outputs of the Thesis Project

There were a number of preliminary outputs of this thesis project, early versions of some of the arguments and methods applied in the thesis appear in these early outputs (although no material has been copied directly into the thesis). These outputs were as follows.

Poster Presentations at Conferences

- Droy L (2016) Female Genital Mutilation (FGM): Explaining the occurrence of a harmful practice with computational simulations of social groups. Poster presented at: *7th ESRC Research Methods Festival (5-7 July 2016)*, *12th Festival of PG Research (University of Leicester, 7 July 2016)* & *9th RC33 International Conference on Social Science Methodology (11th-16th September 2016)*

This poster reported some output from a preliminary (and very simply) agent-based model of FGM as a social norm. It compared the distribution of prevalence of FGM across communities in Senegal, to the distribution of rates of practice across multiple simulation runs. So, it is a preliminary version of the validation exercise reported in Chapter 7.

- Droy L (2017) Can we use simulated experiments to identify important agent-based model (ABM) design choices prior to empirical calibration? A case of modeling a harmful collective practice. Poster presented at: *Interdisciplinary Workshop on Opinion Dynamics and Collective Decision 2017 July 5-7, 2017 @ Jacobs University Bremen, Germany*

This poster reported an exploration of a preliminary (and very simple) agent-based model of FGM as a social norm. It explored different network algorithms and structures and their potential impact on the ‘key dynamics’ of the model, and used random-forest modelling to try to identify the most important design choices among the different aspects of the social network design. It can be considered a prototype of the ‘Exploratory Robustness Analysis’ followed by ‘Global Sensitivity Analysis’ approach employed in Chapters 5 and 6.

Working Papers

- Droy L (2016) Using Agent-Based Modeling to Test Social Coordination Accounts of the Incidence of Female Genital Mutilation against Individualistic Accounts. *Unpublished Manuscript*.

In this working paper, a version of which was presented at the Social Simulation Conference (Rome, 2016), a preliminary (and very simple) coordination model of FGM was used and its outputs were compared to the joint distribution of popularity and prevalence of FGM

observed in Senegal and three other countries. No actual calibration was undertaken, relatively arbitrary parameter values were chosen. Nevertheless, the idea of using the joint distribution of popularity and prevalence of FGM within communities as a ‘pattern’ against which models of FGM could be tested was initially expressed here – so this is a prototype of the validation approach taken in Chapter 7.

Published Articles

- Droy L (2018) Studying Social Processes Underlying the Persistence of Female Genital Mutilation Using Agent-Based Modeling. *SAGE Research Methods Cases*.

This article, which was peer-reviewed, included a simple agent-based model of FGM as a social norm of coordination. It tested this model against two empirical patterns: the distribution of FGM prevalence across communities in Senegal, and the joint distribution of the popularity and prevalence of FGM across communities in Senegal. This model included a preliminary attempt at ‘calibration’, but this wasn’t independent, it involved ‘fitting’ the parameters of the model using an algorithm to minimise the difference between the model output, and the observed data – so this is a kind of prototype of the calibration and validation undertaken in Chapter 7 of the thesis, without independent calibration of parameter values.

Appendix H: Special Materials

Appendix H1: Domain Theory Versus ‘Off-The-Shelf’ Behavioural Frameworks

I’ve argued in this chapter, and previously, that decisions about model design should be driven by methodology. In particular, a methodology with the potential to address key uncertainties in model design in a principled fashion. For example, questions like: which features of a social system are important and need to be included in a model? This kind of consideration isn’t just about the particular design of a given model in a given domain. It is also about searching for a *procedure* for model development that is likely to yield adequate models, to help the field to progress, and so on, over the long run. In fact, I’ve argued that one of the chief advantages of ABM as a modelling framework is that it allows model design to be strongly guided by methodology. In particular, it allows modellers to employ systematic methodological techniques to guide model design without being restricted by the kinds of technical assumption that constrain other modelling frameworks.

Given a candidate model, techniques like robustness analysis and (subsequently) empirical calibration are obvious contenders for a systematic methodological approach to model development, since they can be used to directly investigate (rather than make arbitrary assumptions about) questions like whether a feature of a social system ‘matters’ and (if does matter), how it should be represented in a model. Indeed, the modelling strategy that I propose and follow in this thesis has techniques of this kind at its core.

The question of where to *begin* is perhaps trickier. The answer that I proposed earlier in this chapter is that we should adopt a *procedure* of beginning with a selection of candidate ‘core’ theories (in this case, theories about FGM decision-making), evaluating their empirical support, and using one as a starting point for model development. This is the subject of the next chapter (4). In the FGM literature, the most clearly articulated and widely used core theories of decision-making are Social Norm of Coordination theory (Mackie, 2017), and alternatives focused on marriage competition or informational influence (see Chapter 4). These theories are generally expressed as (boundedly) rational choice accounts of decision-making (UNICEF, 2007) – a format that I retain in my own modelling efforts.

Although, in a given field, it is always possible that the theories favoured by domain experts are ‘wrong’, we would expect that, in general, domain theories represent the most viable hypotheses that are available to the modeller. They are thus *generally* appropriate as a set of

initial ‘priorities’ for the design of a model. Put another way, modellers should need a very strong justification to reject the theories of domain experts in favour of some alternative set of initial design decisions that are divorced from current scientific knowledge of the target phenomenon. Furthermore, by subjecting a selection of initial core hypotheses to empirical adjudication (see Chapter 4), we render our choice to rely on domain theories even more rigorous. We ensure that core theory selection is informed both by domain knowledge and direct empirical evidence.

An entirely analogous argument can be made regarding the procedure for using exploratory robustness analysis to identify important model design variants and (de)idealisation (see Modelling Strategy above). The potential importance of *any* plausible de-idealisation of a model *could* be explored and might be found to ‘matter’ for model dynamics. However, a more fruitful *procedure* is likely to involve prioritising model design features (such as introducing norm enforcement, see Chapter 5) that have been identified as important in the existing domain literature. Other things being equal, we would expect these to be more likely to ‘matter’ for model dynamics (since domain experts already think that these features matter) and to be more likely to be retained by subsequent empirical calibration (since domain experts already think that these features are present in the target system).

This methodological approach (beginning with domain knowledge) strongly conforms to our intuitions about social scientific progress (that domain knowledge accumulates, is employed, tested and subsequently refined). However, such an approach has not always been adopted by modellers, including those using ABM in the social sciences.

An alternative approach among ABM practitioners has been to begin with an ‘off-the-shelf’ behavioural framework (Balke and Gilbert, 2014), rather than a bespoke design based on domain theory. Such ‘behavioural frameworks’ (see examples below) generally involve generic cognitive agent ‘architectures’. These consist of different information processing modules (e.g. ‘beliefs’ and ‘desires’, see below) that (when instantiated as agent code) can furnish a simulation of agents who process information, make decisions, interact and so on. One of the chief advantages of beginning with a generic behavioural framework approach appears to be that it can be imported quickly into the design of a model - requiring only minimal ‘tweaking’ to apply it to a particular target system⁹⁴. As such, behavioural

⁹⁴ Of course, if excessive ‘tweaking’ is required then arguably the modeler is really using domain specific theory anyway.

frameworks have the advantages of being highly convenient (at least in principle). They also have the advantage of helping to create comparability between models in distinct domains.

My main objection to the use of behavioural frameworks, and the primary reason that they are not employed in this thesis, is that the use of a *generic* ‘off the shelf’ agent design is not compatible with the kind of methodological approach that I am espousing. Generic behavioural frameworks, by definition, are not built with a particular target system in mind. Therefore (unlike domain theory) they are not based on an attempt to prioritise the most important features of a particular real-world system. Put another way, they are at cross-purposes with an attempt to address uncertainty about model design systematically through domain knowledge, empirical evidence, or model-based exploration (see Modelling Strategy above). Instead, they constitute a generic assertion of how social systems should be represented. Contrary to my recommendations in this chapter (and throughout the thesis), the use of a generic behavioural framework would obscure (or, if you like, arbitrarily decide) questions of model design uncertainty – rather than resolving them through methodology.

This is not to say that generic behavioural frameworks will never ‘work’ in a particular domain. However, *in general*, we would expect that these frameworks will frequently *neglect* aspects of decision-making that are important for a particular target social system (and associated problems being addressed). Likewise, we would expect that generic frameworks will frequently include features of decision-making that are not relevant or appropriate for a particular system or problem being studied. This expectation follows from a recognition that the most important features of agent decision-making (for a given modelling problem) are likely to vary across real-world domains - yet generic behavioural frameworks, by definition, can only show limited responsiveness to such differences⁹⁵.

The argument so far is at the level of methodology and procedure. This, I believe, is the most important reason why the use of generic behavioural frameworks is inappropriate for this thesis, and for the study of the dynamics of FGM in general. However, the case against the use of behavioural frameworks need not be confined to abstract considerations. It is not only the case that the use of generic architectures of this kind is inadvisable as a *general principle* (given the methodological approach that I’ve advocated). It is also the case that the *particular*

⁹⁵ Again, if they were highly responsive to differences between real-world domains, then behavioural frameworks would become hard to distinguish from domain theory anyway, and their ability to contribute anything meaningful to the model development process would become questionable.

frameworks that might be applied to studying the social dynamics of FGM are a poor fit for the phenomenon and the problems that the field is attempting to address. In the remainder of this chapter, I review examples of behavioural frameworks in the ABM literature that are widely used and (prima-facie) relevant to a phenomenon like FGM. I argue that these display the problem highlighted my general argument – the features of decision-making prioritised in these frameworks do not match the issues of key importance in the field of FGM dynamics.

Behavioural frameworks in the ABM literature are also known as ‘cognitive agent’ frameworks. They are part of a research tradition that has its origins in early AI research and focuses on the representation of agents as information processing systems with distinct modules that interact together to guide knowledge acquisition, decision-making, planning, and action (etc.) (Balke and Gilbert, 2014). As noted above, the design of these cognitive systems is typically stated abstractly in different *cognitive architectures* and then implemented with computer code, as part of agent-based models, or other kinds of simulation.

The archetypal cognitive agent framework used in agent-based modelling is the Belief, Decision, Intentions architecture (BDI) (Balke and Gilbert, 2014). The basic motivation of BDI is to represent agents as having a ‘mental state’ which they use to reason about the world and then choose actions. Agents have beliefs, which can be updated in response to external events. They have desires, which are situations that they would prefer, and intentions which drive actions in service of those desires. Through a deliberation process, and with reference to a ‘library of plans’ (which help the agent to identify actions to reach their goals) agents can act autonomously to select actions that solve problems (i.e. satisfy their desires), within a dynamic environment.

One of the appeals of the BDI architecture is that it can be used to express agents with a certain degree of cognitive sophistication that may be difficult to express using traditional game-theoretic notion (Neumann, 2014: 51). In particular, BDI provides a design for agents who can identify a collection of actions (out of a wide range of options) to help them to achieve a desired outcome (including after taking into account ‘beliefs’ about a changing environment) (Balke and Gilbert, 2014). The depth of capacity of the BDI architecture for modelling complex problem solving is reflected in some of its industrial applications, including its use in control systems for the NASA Discovery Space Shuttle, in the management of complex telecommunications networks (Balke and Gilbert, 2014; Georgeff

and Ingrand, 1990) and in its use in an ABM context to model the challenges faced by individuals when planning their commute to work (Urquhart et al., 2019).

Nevertheless, whilst the capacity of BDI to represent certain kinds of complex problem solving is doubtless useful in some domains, it is not clear that FGM is one of them. The issue which has primarily motivated theories about FGM decision-making has been to understand *the circumstances under which* individual social actors will prefer to abandon FGM, or not (e.g. Mackie, 1996, Efferson et al., 2015). This issue is central to understanding the persistence and abandonment of the practice. By contrast, the issue of *how* individuals abandon FGM (i.e. how they select a set of action plans, given the goal of abandoning FGM) has little relevance for these questions. Stated simply, the question of interest is *whether* actors will choose to abandon FGM, not *how* they will do so. The BDI architecture offers nothing new to the former question, whilst adding a great deal of unnecessary complexity in the handling of the latter.

Nevertheless, whilst being a general framework for decision-making, the BDI architecture was not designed to model collective social behaviours (i.e. social norms) specifically. FGM is widely considered to be a social norm (at least in the generic of being a widely followed social rule⁹⁶) (UNICEF, 2013, Mackie, 2017). As such, a fair assessment of the available behavioural frameworks in the ABM literature needs to consider architectures that have been tailored towards understanding social norms specifically (including derivatives of BDI that include additional cognitive ‘modules’ that address normative behaviour, Neumann, 2014: 60).

EMIL-A, and its derivatives (e.g. EMIL-A-I) are the leading BDI descendants (Andrighetto et al., 2014: 162) focused on social norm behaviour among agents (Campenni et al., 2014). These architectures contribute to the cognitive modelling of social norm behaviour by incorporating specific computational ‘modules’ that give agents the capacity to reason about and respond to information about social norms. Andrighetto et al. (2014) and Campenni et al. (2014) describe the cognitive capabilities that EMIL-A/EMIL-A-I are designed to introduce. The most important of these are the ability of agents to *recognise new norms*, *innovate norms* and *internalise norms* (see also Balke and Gilbert, 2014). These cognitive modules, which are discussed below, permit the exploration of a range of dynamics related to social norms. These include: how new kinds of social norms are created and transmitted across populations

⁹⁶ See Mackie et al. (2015) for an exhaustive list of general definitions of a social norm

(recognition and innovation), and how costly norms can be sustained in the absence of external incentives (internalisation).

Campenni et al. (2014: 96) report an ABM study of cognitive EMIL-A agents endowed with a relatively sophisticated ‘norm recognition’ module. This module endows agents with normative representations (beliefs and goals related to norms and norm following) as well as norm recognition, whereby the behaviours or signals of other agents can lead to the recognition of a new action as a norm, as well as capacities for agents to adopt normative beliefs and comply with normative goals.

To demonstrate the utility of this new framework, Campenni et al. (2014), used a simulated ABM experiment comparing ‘Norm Detectives’ (ND) – cognitive agents endowed with norm recognition, with ‘Social Conformers’ (SC) – simple agents who imitate others. These agents were distributed across a ‘multi-setting’ world, in which agents move between different environments autonomously, thereby exposing them to a changing selection of local agents that might influence them. In each setting, agents had a range of possible actions they could take, some were ‘setting specific’ and one possible action was shared across all settings. Agents within settings interacted. ND agents exchanged messages indicating that some action is a ‘norm’. Based on these interactions, they could recognise and potentially adopt particular actions as normative goals. SC agents, by contrast, would just adopt the action most frequently performed in their local environment. The results of Campenni et al.’s (2014) simulated experiments showed that Norm Detective agents were able to converge effectively on the shared action option (the one common across settings) as a dominant ‘norm’. Whereas, the Simple Conformer agents did not effectively converge on a single action. The authors argued that the ‘Norm Recognition’ module, therefore, provides important insights into how new norms emerge in populations, especially given a changing social environment.

Andrighetto et al. (2014) introduced a ‘Norm Internalisation’ module into EMIL-A agents, which they called EMIL-A-I agents. This added a multi-step cognitive process whereby, after a certain period, highly ‘salient’⁹⁷ social norms would be *internalised* by social actors, meaning that they followed them automatically without further deliberation. Internalisation might also cease if norm salience decreased. The authors simulated mixed populations of EMIL-A and

⁹⁷ Salience as increased by factors like observed compliance and punishment of norm-breakers, and decreased by factors like observing un-punished norm defection.

EMIL-A-I agents playing a Prisoner's Dilemma game (with punishment and sanctioning options between rounds). They found that levels of cooperation increased when the proportion of EMIL-A-I agents in the agent population increased. This increase was due to the tendency for EMIL-A-I agents to internalise a norm of cooperation for an extended period, even in the absence of punishment/sanctions to maintain norm salience.

These kinds of insights, developed through the use of sophisticated cognitive agents with normative cognitive process, are certainly intriguing for the abstract theoretical study of social norms. However, the features of social norms that are addressed do not align well with the problems faced by those studying the social dynamics of FGM specifically.

The origins of FGM as a social norm (i.e. norm recognition/innovation) are, for example, certainly of some intellectual interest (e.g. Mackie and LeJeune, 2009), but they are not a pressing policy issue. FGM is usually a long-established norm in communities where it is practised. Naturally then, the question of primary interest to modellers and policy-makers is *how communities can be moved away* from the established norm of FGM (UNICEF, 2007), *not* how it came to be recognised as a norm in the first instance. Norm recognition and innovation are already implicit in domain theory (FGM is assumed to be an established norm that agents are aware of). Therefore, the highly complex computational reasoning associated with norm recognition or innovation in EMIL appears to be a poor match for the issues of relevance in the domain and field of FGM.

The relevance of norm internalisation to the field of FGM is similarly questionable. There may be individuals who have 'internalised' FGM as a norm, such that they will practice it without reference to the behaviour of others. This can already be accommodated by existing theory and is understood in terms of actors attributing high intrinsic value to FGM (see Chapter 5). However, the introduction of a mechanism whereby agents come to unconditionally follow an established norm *without deliberation*, would only obscure analysis of one of the key questions for the field: How can social actors be encouraged to abandon a long-established norm of FGM (Mackie, 1996)? The notion of norm internalisation, especially when considered to be the end result of a long-term socialisation process, would seem to reject this possibility out of hand – contrary to empirical evidence that individuals *can* be persuaded to oppose FGM (Berg and Denison, 2012; Diop et al., 2008). Introducing norm internalisation would also prevent exploration of a key theoretical contention in the study of the social dynamics of FGM: that individuals may practice FGM despite preferring that the practice ended (Mackie, 2017). This possibility is already clearly expressed in the

rational-choice language of existing domain theory: through competing social and non-social incentives. Introducing norm internalisation, whereby agents adopt long-term norms as an end in themselves, would seem to preclude the possibility of such discrepancies – despite empirical evidence that they occur (UNICEF, 2013).

These issues can be seen as instances of the general problems cited at the beginning of this section. The developers of generic behavioural architectures have neither the domain of FGM nor the key problems addressed by modelling its dynamics, in mind when developing their frameworks. Understandably, therefore, their priorities (what is and is not included in their frameworks) do not align well with the needs of the field. Much of the use of cognitive agents to study social norms in the ABM literature (Neumann, 2008: Table 4) has had a fundamentally *functional* orientation (i.e. concerned with generating patterns of social behaviour that benefit and perpetuate society). Specifically, research has focused on developing cognitive agents with ‘normative’ components (i.e. decision-making processes that allow them to respond to social rules) and analysing the way that these facilitate collective benefits. This is clearly demonstrated in contemporary presentations of the sophisticated EMIL-A-I agent discussed above (Andrighetto et al., 2014), and even in ‘classic’ ABM research using extremely simple cognitive agents (Conte and Castelfranchi, 1995). Yet, this ‘functionalist’ perspective is evidently of limited relevance for the study of FGM – which is an *individually and collectively costly pattern of behaviour* on which social actors hold strongly diverging views and for which the key policy problem is to encourage *abandonment* of the norm (despite social pressures holding the practice place).

Bibliography

28TooMany (2016) *The Medicalisation of FGM*. 28TooMany.

28TooMany (2018a) *SENEGAL: THE LAW AND FGM*. 28TooMany.

28TooMany (2018b) *THE LAW AND FGM AN OVERVIEW OF 28 AFRICAN COUNTRIES*. 28TooMany.

Abdalla RHD (1982) *Sisters in Affliction: Circumcision and Infibulation of Women in Africa*. London: Zed Press.

Abdou M and Gilbert N (2009) Modelling the emergence and dynamics of social and workplace segregation. *Mind and Society* 8: 173–191.

Ako MA and Akweongo P (2009) The limited effectiveness of legislation against female genital mutilation and the role of community beliefs in Upper East Region, Ghana. *Reproductive Health Matters* 17(34): 47–54.

Almroth L, Almroth-Berggren V, Hassanein OM, Hadi EL and Al-Said SS (2001) A community based study on the change of practice of female genital mutilation in a Sudanese village. *International journal of gynaecology and obstetrics: the official organ of the International Federation of Gynaecology and Obstetrics* 74(2): 179–185.

Althaus FA (1997) Female Circumcision: Rite of Passage or Violation of Rights? *International Family Planning Perspectives* 23(3): 130–133.

Amer M, Daim TU and Jetter A (2013) A review of scenario planning. *Futures* 46: 23–40.

Andrighetto G, Villatoro D and Conte R (2014) The Role of Norm Internalizers in Mixed Populations. In: Conte R, Andrighetto G, and Campenni M (eds) *Minding Norms: Mechanisms and Dynamics of Social Order in Agent Societies*. Oxford University Press.

Anzola D (2019) Disagreement in discipline-building processes. *Synthese*: 1-24.

Aodha L and Edmonds B (2017) Some Pitfalls to Beware When Applying Models to Issues of Policy Relevance. In: Edmonds B and Meyer R (eds) *Simulating Social Complexity*. Cham: Springer International Publishing, pp. 801–822.

Axtell R, Epstein J, Dean J, et al. (2006) POPULATION GROWTH AND COLLAPSE IN AMULTIAGENT MODEL OF THE KAYENTA ANASAZI IN LONG HOUSE

- VALLEY. In: Epstein JM (ed.) *Generative Social Science: Studies in Agent-Based Computational Modeling*. STU-Student edition. Princeton University Press.
- Badham J, Chattoe-Brown E, Gilbert N, et al. (2018) Developing agent-based models of complex health behaviour. *Health & place* 54. Elsevier: 170–177.
- Balke T and Gilbert N (2014) How Do Agents Make Decisions? A Survey. *Journal of Artificial Societies and Social Simulation* 17(4): 13
- Ball P (2012) *Why Society Is a Complex Matter: Meeting Twenty-First Century Challenges with a New Kind of Science*. Springer Berlin Heidelberg
- Bellemare MF, Novak L and Steinmetz TL (2015) All in the family: Explaining the persistence of female genital cutting in West Africa. *Journal of Development Economics* 116: 252–265.
- Berg RC and Denison E (2012) Effectiveness of Interventions Designed to Prevent Female Genital Mutilation/Cutting: A Systematic Review. *Studies in Family Planning* 43(2): 135–146.
- Berg RC and Denison E (2013) A Tradition in Transition: Factors Perpetuating and Hindering the Continuance of Female Genital Mutilation/Cutting (FGM/C) Summarized in a Systematic Review. *Health care for women international* 34(10): 837–859.
- Berg RC, Underland V, Odgaard-Jensen J, Fretheim A and Vist GE (2014) Effects of female genital cutting on physical health outcomes: A systematic review and meta-analysis. *BMJ Open* 4(11).
- Bicchieri C (2005) *The Grammar of Society: The Nature and Dynamics of Social Norms*. Cambridge University Press.
- Bicchieri C, Lindemans JW and Jiang T (2014) A structured approach to a diagnostic of collective practices. *Frontiers in psychology* 5: 1418.
- Bicchieri C and Mercier H (2014) Norms and Beliefs: How Change Occurs. In: Xenitidou M and Edmonds B (eds) *The Complexity of Social Norms*. Cham: Springer International Publishing, pp. 37–54.
- Bicchieri C, Muldoon R and Sontuoso A (2018) Social Norms. In: Zalta EN (ed.) *The Stanford Encyclopedia of Philosophy*. Stanford University.

- Boddy JP (1982) Womb as oasis: the symbolic context of Pharaonic circumcision in rural Northern Sudan. *American Ethnologist* 9(4): 682–698.
- Boyle EH (2002) *Female Genital Cutting: Cultural Conflict in the Global Community*. Baltimore: Johns Hopkins University Press.
- Bravo G, Squazzoni F and Boero R (2012) Trust and partner selection in social networks: An experimentally grounded model. *Social Networks* 34: 481–492.
- Brousmiche K-L, Kant J-D, Sabouret N, et al. (2016) From Beliefs to Attitudes: Polias, a Model of Attitude Dynamics Based on Cognitive Modeling and Field Data. *Journal of Artificial Societies and Social Simulation* 19(4): 2.
- Camilotti G (2016) Interventions to Stop Female Genital Cutting and the Evolution of the Custom: Evidence on Age at Cutting in Senegal. *Journal of African Economies* 25(1): 133–158.
- Campenni M, Andrighetto G, Cecconi F, et al. (2014) Hunting for Norms in Unpredictable Societies. In: Conte R, Andrighetto G, and Campenni M (eds) *Minding Norms: Mechanisms and Dynamics of Social Order in Agent Societies*. Oxford University Press.
- Center for Reproductive Rights (n.d.) *Legislation on Female Genital Mutilation in the United States (Briefing Paper)*. Center for Reproductive Rights. Available at: https://www.reproductiverights.org/sites/default/files/documents/pub_bp_fgmlawsusa.pdf.
- Center for Reproductive Rights (2006) *Female Genital Mutilation: A Matter of Human Rights – An Advocate’s Guide to Action*. Center for Reproductive Rights.
- Center for Reproductive Rights (2009) *Female Genital Mutilation (FGM) Legal Prohibitions Worldwide*. Center for Reproductive Rights.
- Chattoe-Brown E (2009) The social transmission of choice: a simulation with applications to hegemonic discourse. *Mind & Society* 8(2): 193–207.
- Chattoe-Brown E (2013) Why Sociology Should Use Agent Based Modelling. *Sociological Research Online* 18(3): 1-11
- Chattoe-Brown E (2014) Using Agent Based Modelling to Integrate Data on Attitude Change. *Sociological Research Online* 19(1): 1–16.

- Chattoe-Brown E (2017a) Why Questions Like “Do Networks Matter?” Matter to Methodology: How Agent-Based Modelling Might Start To Answer Them. *Unpublished Manuscript*.
- Chattoe-Brown E (2017b) *Agent-Based Modelling*. Oxford Bibliographies, Sociology. Oxford University Press.
- Chattoe-Brown E (2019) Un Drôle De Type: The Schelling Model, Calibration, Specification, Validation and Using Relevant Data In: Ahrweiler, Petra (eds) *Proceedings of Social Simulation Conference 2019*. Mainz: Springer.
- Chattoe-Brown E, Colman A, Jenkins D, Krockow EM, Mehtar S, Perera N (2019) The Challenge of Progressive Research in Agent-Based Modelling: A Case Study of Antimicrobial Resistance and Policy Intervention. *Unpublished Manuscript*.
- Chattoe-Brown E (2020a) A BIBLIOGRAPHY OF ABM RESEARCH EXPLICITLY COMPARING REAL AND SIMULATED DATA FOR VALIDATION. *Review of Artificial Societies and Social Simulation*. Available at: <https://rofasss.org/2020/06/12/abm-validation-bib/>
- Chattoe-Brown E (2020b) Why questions like ‘do networks matter?’ matter to methodology: how Agent-Based Modelling makes it possible to answer them. *International Journal of Social Research Methodology*. Routledge: 1–14. DOI: 10/gg9snw.
- Cheikh Seydil Mactar Mbacke (2018) *RE: Mischaracterization and misrepresentation of Tostan and its work*. E-letter 15 January 2018: Science (Magazine). Available at: <https://science.sciencemag.org/content/349/6255/1446/tab-e-letters>.
- Chesnokova T and Vaithianathan R (2010) The economics of female genital cutting. *B.E. Journal of Economic Analysis and Policy* 10(1): 1–26.
- Clark WAV (1991) Residential Preferences and Neighborhood Racial Segregation: A Test of the Schelling Segregation Model. *Demography* 28(1): 1–19.
- Conte R and Castelfranchi C (1995) Understanding the functions of norms in social groups through simulation. In: Gilbert N and Conte R (eds) *Artificial Societies: The Computer Simulation of Social Life*. London: UCL Press.

- Cloward K (2015) Elites, Exit Options, and Social Barriers to Norm Change: The Complex Case of Female Genital Mutilation. *Studies in Comparative International Development* 50(3): 378–407.
- Córdoba C and García-Díaz C (2020) Reflexivity in a Diffusion of Innovations Model. *Journal of Artificial Societies and Social Simulation* 23(3): 9
- Cointet J-P and Roth C (2007) How Realistic Should Knowledge Diffusion Models Be? *Journal of Artificial Societies and Social Simulation* 10(3): 5.
- Coyne CJ and Coyne RL (2014) The identity economics of female genital mutilation. *The Journal of Developing Areas* 48(2): 137–152.
- De Cao E and Lutz C (2018) Sensitive Survey Questions: Measuring Attitudes Regarding Female Genital Cutting Through a List Experiment. *Oxford Bulletin of Economics and Statistics* 80(5): 871–892.
- Dellenborg L (2004) A reflection on the cultural meanings of female circumcision experiences from fieldwork in Casamance, southern Senegal. In: Arnfred S (eds) *Re-Thinking Sexualities in Africa*. Nordiska Afrikainstitutet pp. 79-94.
- Dietze PM, Livingston M, Callinan S, et al. (2014) The big night out: what happens on the most recent heavy drinking occasion among young Victorian risky drinkers? *Drug and alcohol review* 33(4): 346–353
- Diop N, Faye M, Cabral A, Moreau A, Benga H, Cissé F, et al. (2004) *The TOSTAN Program Evaluation of a Community Based Education Program in Senegal*. Dakar, Senegal: Population Council, GTZ and TOSTAN.
- Diop NJ, Moreau A and Benga H (2008) *Evaluation of the Long-term Impact of the TOSTAN Programme on the Abandonment of FGM/C and Early Marriage: Results from a qualitative study in Senegal*. Dakar, Senegal: Population Council.
- Droy L (2018) Studying Social Processes Underlying the Persistence of Female Genital Mutilation Using Agent-Based Modelling. *SAGE Research Methods Cases*.
- Droy L, Hughes L, Lamont M, Nguura P, Parsitau D, Wamue-Ngare G (2018) Alternative Rites of Passage in FGM/C Abandonment Campaigns in Africa: A research opportunity. *LLAS Working Paper Series 1*.

Edmonds B (2015) Using Qualitative Evidence to Inform the Specification of Agent-Based Models. *Journal of Artificial Societies and Social Simulation* 18(1): 18.

Edmonds B (2017) Different Modelling Purposes. In: Edmonds B and Meyer R (eds) *Simulating Social Complexity*. Cham: Springer International Publishing. pp. 39-58.

Edmonds B and Aodha L (2019) Using Agent-Based Modelling to Inform Policy – What Could Possibly Go Wrong? In: Davidsson P and Verhagen H (eds) *Multi-Agent-Based Simulation XIX*. Cham: Springer pp. 1-16.

Edmonds B and Meyer R (2013) *Simulating Social Complexity: A Handbook*. Springer Berlin Heidelberg.

Edmonds B and Moss S (2005) From KISS to KIDS – An ‘Anti-simplistic’ Modelling Approach. In: Davidsson P, Logan B and Takadama K (eds) *Multi-Agent and Multi-Agent-Based Simulation. MABS 2004* Berlin: Springer, pp. 130–144.

Efferson C, Vogt S, Elhadi A, Ahmed H and Fehr E (2015) Female genital cutting is not a social coordination norm. *Science* 349(6255): 1446–1447.

Efferson C, Vogt S, Elhadi A, Ahmed H and Fehr E (2015 - Supplement) Supplementary Materials for ‘Female genital cutting is not a social coordination norm’ (Efferson et al., 2015), Available at: <https://science.sciencemag.org/content/349/6255/1446/>

Efferson C, Vogt S and Fehr E (2019) The promise and the peril of using social influence to reverse harmful traditions. *Nature Human Behaviour*. 4(1): 55-68

El Dareer AE (1983) *Woman, Why Do You Weep: Circumcision and Its Consequences*. London: Zed Press.

Eldin A, Babiker S, Sabahelzain M, Eltayeb M (2018) *FGM/C Decision-Making Process and the Role of Gender Power Relations in Sudan*. New York: Population Council.

Epstein J and Axtell R (1996) *Growing Artificial Societies: Social Science from the Bottom Up*. Washington, DC: Brookings Institution.

Epstein JM (2006) *Generative Social Science: Studies in Agent-Based Computational Modeling*. Princeton University Press.

European Commission (2013) *Towards the Elimination of Female Genital Mutilation*. Brussels: Communication from the Commission to the European Parliament and the Council.

- Evans A, Heppenstall A and Birkin M (2017) Understanding Simulation Results. In: Edmonds B and Meyer R (eds) *Simulating Social Complexity*. Cham: Springer International Publishing. pp. 205–227.
- Flache A, Mäs M, Feliciani T, Chatteo-Brown E, Deffuant G, Huet S and Lorenz J (2017) Models of Social Influence: Towards the Next Frontiers. *Journal of Artificial Societies and Social Simulation* 20(4).
- FORWARD (2016) *Female Genital Mutilation Frequently Asked Questions*. London: FORWARD.
- Freymeyer RH and Johnson BE (2007) An exploration of attitudes toward female genital cutting in Nigeria. *Population Research and Policy Review* 26(1): 69–83.
- Galán JM and Izquierdo LR (2005) Appearances Can Be Deceiving: Lessons Learned Re-Implementing Axelrod's 'Evolutionary Approach to Norms'. *Journal of Artificial Societies and Social Simulation* 8(3).
- Georgeff MP and Ingrand FF (1990) Real-time reasoning: the monitoring and control of spacecraft systems. In: *Sixth Conference on Artificial Intelligence for Applications*, 5 May 1990, pp. 198–204.
- Ghorbani A, Dijkema G and Schrauwen N (2015) Structuring Qualitative Data for Agent-Based Modelling. *Journal of Artificial Societies and Social Simulation* 18(1).
- Gilbert N and Troitzsch K (2005) *Simulation for the Social Scientist*. Maidenhead, UK: Open Univ. Press.
- Glasgow G and Alvarez RM (2009) Discrete Choice Models. In: Box-Steffensmeier JM, Brady HE, and Collier D (eds) *The Oxford Handbook of Political Methodology*. Oxford University Press. pp. 513–529.
- Goldthorpe JH (2015) *Sociology as a Population Science*. Cambridge University Press.
- GTZ (2006) *The Maputo Protocol of the African Union An Instrument for the Rights of Women in Africa*. Deutsche Gesellschaft für Technische Zusammenarbeit (GTZ) GmbH.
- Granovetter M (1978) Threshold Models of Collective Behavior. *American Journal of Sociology* 83(6): 1420–1443.

- Grimm V, Revilla E, Berger U, Jeltsch F, Mooij WM, Railsback SF, Thulke H, Weiner J, Wiegand T, DeAngelis DL (2005) Pattern-Oriented Modeling of Agent-Based Complex Systems: Lessons from Ecology. *Science* 310(5750): 987–991.
- Gronau R (1998) A Useful Interpretation of R-Squared in Binary Choice Models (Or, Have We Dismissed the Good Old R-Squared; Prematurely). *IRS Working Papers*.
- Gruenbaum E (1982) The Movement against Clitoridectomy and Infibulation in Sudan: Public Health Policy and the Women's Movement. *Medical Anthropology Newsletter* 13(2): 4–12.
- Gruenbaum E (2001) *The Female Circumcision Controversy*. University of Pennsylvania Press.
- Haag D and Kaupenjohann M (2001) Parameters, prediction, post-normal science and the precautionary principle—a roadmap for modelling for decision-making. *Ecological Modelling* 144(1): 45–60.
- Hamill L and Gilbert N (2009) Social circles: A simple structure for agent-based social network models. *Journal of Artificial Societies and Social Simulation*, 12(2).
- Hammond RA (2015) Considerations and best practices in agent-based modeling to inform policy. In: Ogawa VA, Geller A and Wallace R (eds) *Assessing the Use of Agent-Based Models for Tobacco Regulation*. US: National Academies Press. Appendix A.
- Hardesty L (2010) Explained: Linear and nonlinear systems. *MIT News Office*. Available at: <https://news.mit.edu/2010/explained-linear-0226>.
- Hayford SR (2005) Conformity and change: community effects on female genital cutting in Kenya. *Journal of Health and Social Behavior* 46(2): 121–140.
- Hayford SR and Trinitapoli J (2011) Religious differences in female genital cutting: a case study from Burkina Faso. *Journal for the scientific study of religion* 50(2): 252–271.
- Hedström P (2005) *Dissecting the Social: On the Principles of Analytical Sociology*. Cambridge University Press.
- Heine BO, Meyer M and Strangfeld O (2005) Stylised facts and the contribution of simulation to the economic analysis of budgeting. *Journal of Artificial Societies and Social Simulation* 8(4).
- Hernlund YK and Shell-Duncan B (2007) Contingency, context, and change: Negotiating female genital cutting in The Gambia and Senegal. *Africa Today* 53(4): 43–57.

- Hicks EK (1996) *Infibulation : Female Mutilation in Islamic Northeastern Africa*. New Brunswick: Transaction Publishers.
- Howard JA and Gibson MA (2017) Frequency-dependent female genital cutting behaviour confers evolutionary fitness benefits. *Nature Ecology & Evolution* 1(3): 1–6.
- Hu H, Lin J and Cui W (2015a) Intervention Strategies and the Diffusion of Collective Behavior. *Journal of Artificial Societies and Social Simulation* 18(3): 16.
- Hu H, Lin J and Cui W (2015b) Local Opinion Heterogeneity and Individual Participation in Collective Behavior: A Reconsideration. *Journal of Artificial Societies and Social Simulation* 18(2).
- Jager W and Edmonds B (2015) Policy Making and Modelling in a Complex World. In: Janssen M, Wimmer MA, and Deljoo A (eds) *Policy Practice and Digital Science*. Cham: Springer. pp. 57–73.
- Johansen REB, Diop NJ, Laverack G and Leye Els (2013) What Works and What Does Not: A Discussion of Popular Approaches for the Abandonment of Female Genital Mutilation. *Obstetrics and Gynecology International* 2013(348248): 1-10.
- Kandala NB, Nwakeze N and Kandala SN (2009) Spatial distribution of female genital mutilation in Nigeria. *The American Journal of Tropical Medicine and Hygiene* 81(5): 784–792.
- Kandala NB and Shell-Duncan B (2019) Trends in female genital mutilation/cutting in Senegal: What can we learn from successive household surveys in sub-Saharan African countries? *International Journal for Equity in Health* 18(25).
- Kohler U, Kreuter F and Stuart EA (2019) Nonprobability Sampling and Causal Analysis. *Annual Review of Statistics and Its Application* 6(1): 149–172.
- Kool R (2010) The Dutch Approach to Female Genital Mutilation in View of the ECHR. *Utrecht Law Review*. 6(1): 51-61.
- Koski A and Heymann J (2017) Thirty-year trends in the prevalence and severity of female genital mutilation: a comparison of 22 countries. *BMJ Global Health* 2(4)
- Korgen KO (2017) *The Cambridge Handbook of Sociology*. Cambridge University Press.
- Krebs F (2017) An Empirically Grounded Model of Green Electricity Adoption in Germany: Calibration, Validation and Insights into Patterns of Diffusion. *Journal of Artificial Societies and Social Simulation* 20(2): 10

- Kuorikoski J, Lehtinen A and Marchionni C (2010) Economic Modelling as Robustness Analysis. *The British Journal for the Philosophy of Science* 61(3): 541–567.
- Leamer EE (1983) Let's Take the Con Out of Econometrics. *The American Economic Review* 73(1): 31–43.
- Leonard L (1996) Female circumcision in Southern Chad: Origins, meaning, and current practice. *Social Science and Medicine* 43(2): 255–263.
- Leonard L (2000) “We Did It for Pleasure Only”: Hearing Alternative Tales of Female Circumcision. *Qualitative Inquiry* 6(2): 212–228.
- Levins R (1966) THE STRATEGY OF MODEL BUILDING IN POPULATION BIOLOGY. *American Scientist* 54(4): 421–431.
- Levins R (1993) A Response to Orzack and Sober: Formal Analysis and the Fluidity of Science. *The Quarterly Review of Biology* 68(4): 547–555.
- Lightfoot-Klein H (1989) *Prisoners of Ritual: An Odyssey Into Female Genital Circumcision in Africa*. New York: Harrington Park Press.
- Lumley T (2004) Analysis of Complex Survey Samples. *Journal of Statistical Software* 9(8): 1–19.
- Mackie G (1996) Ending Footbinding and Infibulation: A Convention Account. *American Sociological Review* 61(6): 999–1017.
- Mackie G (2000) Female Genital Cutting: The Beginning of the End. In: Shell-Duncan B and Hernlund Y (eds) *Female "Circumcision" in Africa: Culture, Controversy, and Change*. Colorado: Lynne Rienner. pp. 253–282.
- Mackie G (2003) Female genital cutting: A harmless practice? *Medical Anthropology Quarterly* 17(2): 135–158.
- Mackie G (2017) Female Genital Cutting Is A Social Norm of Coordination. *Unpublished Report*. Available at:
https://www.dropbox.com/sh/1rqgseblyxu3qon/AAD_7uh1MSMsSUly3-UrO6yOa?dl=0&preview=Long+Mackie+Reply+to+Zurich.pdf
- Mackie G (2018) SOCIAL NORMS OF COORDINATION AND COOPERATION. *Social Philosophy and Policy* 35(1): 77–100.

- Mackie G and LeJeune J (2009) Social Dynamics of Abandonment of Harmful Practices: A new look at the theory. *Innocenti Working Papers* 2009(6): 1–42.
- Mackie G, Moneti F, Shakya H and Denny E (2015) *What are social norms? How are they measured*. UNICEF / University of California, San Diego, Center on Global Justice.
- Macy MW and Willer R (2002) From Factors to Actors: Computational Sociology and Agent-Based Modeling. *Annual Review of Sociology* 28(1): 143–166.
- Magnani L and Bertolotti T (2017) *Springer Handbook of Model-Based Science*. Springer Handbooks. Springer International Publishing.
- Modrek S and Liu JX (2013) Exploration of pathways related to the decline in female circumcision in Egypt. *BMC Public Health* 13(1).
- Moss S and Edmonds B (2005) Sociology and simulation: Statistical and qualitative cross-validation. *American Journal of Sociology* 110(4): 1095–1131
- Morison LA, Dirir A, Elmi S, Warsame J and Dirir S (2004) How experiences and attitudes relating to female circumcision vary according to age on arrival in Britain: a study among young Somalis in London. *Ethnicity & Health* 9(1): 75–100.
- Muelder H and Filatova T (2018) One Theory - Many Formalizations: Testing Different Code Implementations of the Theory of Planned Behaviour in Energy Agent-Based Models. *Journal of Artificial Societies and Social Simulation* 21(4).
- Murray K and Conner MM (2009) Methods to quantify variable importance: implications for the analysis of noisy ecological data. *Ecology* 90(2): 348–355.
- Nair S and Howlett M (2017) Policy myopia as a source of policy failure: adaptation and policy learning under deep uncertainty. *Policy & Politics* 45(1): 103–118.
- Nam YE (2018) The power structure in the perpetuation of female genital cutting in Kenya. *Asian Journal of Women's Studies* 24(1): 128–139.
- Novak L (2016) *Persistent Norms and Tipping Points: Female Genital Cutting in Burkina Faso*. Presented at: American Economic Association ASSA Annual Meeting 2017, January 7th Swissotel Chicago, Monte Rosa
- Neumann M (2008) Homo Socionicus: a Case Study of Simulation Models of Norms. *Journal of Artificial Societies and Social Simulation* 11(4): 6.

- Neumann M (2014) How Are Norms Brought About? A State of the Art of Current Research. In: Conte R, Andrighetto G, and Campenni M (eds) *Minding Norms: Mechanisms and Dynamics of Social Order in Agent Societies*. Oxford University Press.
- Obermeyer CM (1999) Female genital surgeries: the known, the unknown, and the unknowable. *Medical Anthropology Quarterly* 13(1): 79–106.
- OHCHR, UNAIDS, UNDP, UNECA, UNESCO, UNFPA, UNHCR, UNICEF, UNIFEM and WHO (2008) *Eliminating Female Genital Mutilation An Interagency Statement*. World Health Organisation.
- Orzack SH and Sober E (1993) A Critical Assessment of Levins's The Strategy of Model Building in Population Biology (1966). *The Quarterly Review of Biology* 68(4): 533–546.
- Ouedraogo S and Koissy-Kpein SA (2012) *An economic analysis of female genital mutilation: How the marriage market affects the household decision of excision*. Presented at: CSAE Conference 2014 Economic Development in Africa, March 24th, Catherine's College, University of Oxford, UK
- Pashaei T, Ponnet K, Moeeni M, Khaazae-Pool M, Majlessi F (2016) Daughters at Risk of Female Genital Mutilation: Examining the Determinants of Mothers' Intentions to Allow Their Daughters to Undergo Female Genital Mutilation. *PLoS ONE* 11(3).
- Platteau JP, Camilotti G and Auriol E (2017) *Eradicating women-hurting customs: What role for social engineering?* World Institute for Development Economic Research
- Poile C and Safayeni F (2016) Using Computational Modeling for Building Theory: A Double Edged Sword. *Journal of Artificial Societies and Social Simulation* 19(3)
- Polhill G and Salt D (2017) The Importance of Ontological Structure: Why Validation by 'Fit-to-Data' Is Insufficient. In: Edmonds B and Meyer R (eds) *Simulating Social Complexity*. Cham: Springer International Publishing, pp. 141–172.
- Polhill (2010). ODD Updated. *Journal of Artificial Societies and Social Simulation*, 13(4).
- Powell K (2017) A nagging persistence. *Nature Human Behaviour* 1: 0026.
- Poyker M (2016) Regime Stability and the Persistence of Traditional Practices. SSRN.
- Rahman A and Toubia N (2000) *Female Genital Mutilation: A Guide to Laws and Policies Worldwide*. Zed Books in association with Center for Reproductive Law and Policy and Research, Action and Information Network for the Bodily Integrity of Women.

- Railsback SF and Grimm V (2012) *Agent-Based and Individual-Based Modeling: A Practical Introduction*. Princeton University Press.
- Richardson K (2002) ‘Methodological Implications of Complex Systems Approaches to Sociality’: Some Further Remarks. *Journal of Artificial Societies and Social Simulation* 5(2): 1–6.
- Ross C, Campiño PJ and Winterhalder B (2015) Frequency-Dependent Social Transmission and the Interethnic Transfer of Female Genital Modification in the African Diaspora and Indigenous Populations of Colombia. *Human Nature* 26(4).
- Ross C, Strimling P, Ericksen KP, Lindenfors P and Mulder MB (2016) The Origins and Maintenance of Female Genital Modification across Africa. *Human Nature* 27(2): 173–200.
- Royal College of Nursing (2019) *Female Genital Mutilation: An RCN Resource for Nursing and Midwifery Practice*. Royal College of Nursing.
- Salganik MJ, Dodds PS and Watts DJ (2006) Experimental Study of Inequality and Unpredictability in an Artificial Cultural Market. *Science* 311(5762): 854.
- Salecker J, Sciaini M, Meyer KM, Wiegand K (2019) The nlrx r package: A next-generation framework for reproducible NetLogo model analyses. *Methods in Ecology and Evolution* 00: 2041–210X.
- Saltelli A, Annoni P, Azzini I, Campolongo F, Ratto M, Tarantola S (2010) Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. *Computer Physics Communications* 181(2): 259–270.
- Saltelli A and Funtowicz S (2014) When All Models Are Wrong. *Issues in Science and Technology* xxx(2).
- Saltelli A, Pereira A, van der Sluijs J and Funtowicz S (2013) What do I make of your Latinorum? Sensitivity auditing of mathematical modelling. *International Journal of Foresight and Innovation Policy* 9(2–4): 213–234.
- Saltelli A, Ratto M, Andres T, Campolongo F, Cariboni J, Gatelli D, Saisana M and Tarantola S (2008) *Global Sensitivity Analysis: The Primer*. New Jersey: John Wiley & Sons.
- Saltelli A, Tarantola S and Campolongo F (2000) Sensitivity Analysis as an Ingredient of Modeling. *Statistical Science* 15(4): 377–395.
- Schindler J (2013) About the Uncertainties in Model Design and Their Effects: An Illustration with a Land-Use Model. *Journal of Artificial Societies and Social Simulation* 16(4).

- Schoemaker PJ (1995) Scenario Planning: A Tool for Strategic Thinking. *Sloan Management Review* 36(2): 25–40.
- Scott N, Livingston M, Hart A, et al. (2016) SimDrink: An agent-based NetLogo model of young, heavy drinkers for conducting alcohol policy experiments. *Journal of Artificial Societies and Social Simulation* 19(1): 10
- Shell-Duncan B (2008) From health to human rights: Female genital cutting and the politics of intervention. *American Anthropologist* 110(2): 225–236.
- Shell-Duncan B and Hernlund Y (2006) Are there "stages of change" in the practice of female genital cutting?: Qualitative research findings from Senegal and The Gambia. *African Journal of Reproductive Health* 10(2): 57–71.
- Shell-Duncan B, Wander K, Hernlund Y & Moreau A (2011) Dynamics of change in the practice of female genital cutting in Senegambia: Testing predictions of social convention theory. *Social Science and Medicine*, 73(8): 1275–1283.
- Shoemaker PJ, Tankard JW and Lasorsa DL (2003) *How to Build Social Science Theories*. London: Sage Publications
- Squazzoni F (2012) *Agent-Based Computational Sociology*. John Wiley & Sons.
- Squazzoni F, Jager W and Edmonds B (2013) Social Simulation in the Social Sciences: A Brief Overview. *Social Science Computer Review* 32(3): 279–294.
- Stirling A (2010) Keep it complex. *Nature* 468(7327): 1029–1031.
- Swedberg R (2001) Sociology and Game Theory: Contemporary and Historical Perspectives. *Theory and Society* 30(3): 301–335.
- ten Broeke G, van Voorn G and Ligtenberg A (2016) Which Sensitivity Analysis Method Should I Use for My Agent-Based Model? *Journal of Artificial Societies and Social Simulation* 19(1).
- Therneau, T and Atkinson B (2019) *An Introduction to Recursive Partitioning Using the RPART Routines*. Mayo Foundation. Available at <https://cran.r-project.org/web/packages/rpart/vignettes/longintro.pdf>
- Thiele JC, Kurth W and Grimm V (2014) Facilitating Parameter Estimation and Sensitivity Analysis of Agent-Based Models: A Cookbook Using NetLogo and ‘R’. *Journal of Artificial Societies and Social Simulation* 17(3).

- Toubia NF and Sharief EH (2003) Female genital mutilation: have we made progress? *International Journal of Gynecology & Obstetrics* 82(3): 251–261.
- Troitzsch KG (1997) Social Science Simulation — Origins, Prospects, Purposes. In: *Simulating Social Phenomena* (eds R Conte, R Hegselmann, and P Terna), Berlin, Heidelberg, 1997, pp. 41–54. Springer Berlin Heidelberg.
- UNFPA (2015) *Demographic Perspectives on Female Genital Mutilation*. Technical Division. New York: United Nations Population Fund.
- UNFPA and UNICEF (2019) *Accelerating Change: UNFPA-UNICEF Joint Programme on the Elimination of Female Genital Mutilation Annual Report 2018*. New York: UNFPA-UNICEF.
- UNICEF (2007) *Coordinated Strategy to Abandon Female Genital Mutilation / Cutting in One Generation*. New York: UNICEF.
- UNICEF (2008) Long-Term Evaluation Of The Tostan Programme In Senegal: Kolda, Thiès And Fatick Regions. *UNICEF (Working Paper)*.
- UNICEF (2010) *The Dynamics of Social Change*. Florence, Italy: UNICEF Innocenti Research Centre.
- UNICEF (2013) *Female Genital Mutilation/Cutting: A Statistical Overview and Exploration of the Dynamics of Change*. New York: UNICEF.
- UNICEF (2016) *Female Genital Mutilation/Cutting: A Global Concern*. New York: UNICEF.
- UNICEF-UNFPA (2014) *UNFPA-UNICEF Joint Programme on Female Genital Mutilation/Cutting: Accelerating Change - Summary Report of Phase 1 (2008-2013)*. UNFPA.
- United Nations Department of Economics and Social Affairs (2017) *Household Size and Composition Around the World 2017*. UNITED NATIONS.
- Urquhart N, Powers S, Wall Z, et al. (2019) Simulating the Actions of Commuters Using a Multi-Agent System. *Journal of Artificial Societies and Social Simulation* 22(2): 10.
- Vogt S, Zaid NAM, Ahmed HEF, Fehr E, Efferson C (2016) Changing cultural attitudes towards female genital cutting. *Nature* 538(7626): 506.
- Volkery A and Ribeiro T (2009) Scenario planning in public policy: Understanding use, impacts and the role of institutional context factors. *Technological Forecasting and Social Change* 76(9): 1198–1207.

- Wack P (1985) Scenarios: Uncharted Waters Ahead. *Harvard Business Review*.
- Wagner N (2011) *Why female genital cutting persists*. Unpublished Report.
- Watts DJ (2002) A simple model of global cascades on random networks. *Proceedings of the National Academy of Sciences* 99(9): 5766–5771.
- Watts DJ and Dodds PS (2007) Influentials, Networks, and Public Opinion Formation. *Journal of Consumer Research* 34(4): 441–458.
- Wahlberg A (2017) *Continuity or Change? Improved Understanding of Attitudes Towards Female Genital Cutting after Migration from Somalia to Sweden*. UPPSALA UNIVERSITET.
- Webb EJ, Campbell DT, Schwartz RD and Sechrest L (1999) *Unobtrusive Measures*. London: SAGE
- Weisberg M (2006) Robustness Analysis. *Philosophy of Science* 73(5): 730–742.
- Weisberg M (2012) *Simulation and Similarity: Using Models to Understand the World*. Oxford University Press.
- WHO (2018) *WHO Guidelines on the Management of Health Complications from Female Genital Mutilation*. World Health Organization.
- Wilensky U (1997) *NetLogo Traffic Basic Model*. Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL. Available at: <http://ccl.northwestern.edu/netlogo/models/TrafficBasic>.
- Wilensky U (1998) *NetLogo Flocking Model*. Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL. Available at: <http://ccl.northwestern.edu/netlogo/models/Flocking>.
- Wilensky U and Rand W (2015) *An Introduction to Agent-Based Modeling: Modeling Natural, Social, and Engineered Complex Systems with NetLogo*. The MIT Press.
- Williams LA (1998) Eradicating female circumcision: Human rights and cultural values. *Health Care Analysis* 6(1): 33–35.
- Williams-Breault BD (2018) Eradicating Female Genital Mutilation/Cutting: Human Rights-Based Approaches of Legislation, Education, and Community Empowerment. *Health and human rights* 20(2). Harvard University Press: 223–233.

Yang H, Wong WH, Bradley K, Toland M (2017) Partial and Semi-Partial Correlations for Categorical Variables in Educational Research: Addressing Two Common Misconceptions. *General Linear Model Journal* 43(1): 1–15.

Yoder PS, Abderrahim N and Zhuzhuni A (2004) *Female genital cutting in the demographic and health surveys: a critical and comparative analysis*. Calverton, Maryland, USA: ORC Macro.

Zhang D (2017) A Coefficient of Determination for Generalized Linear Models. *The American Statistician* 71(4): 310–316.