UNIVERSITY OF LEICESTER

Behavioural Biases and Contract Theory

Matteo Foschi

Department of Economics
University of Leicester

A thesis submitted for the degree of
Doctor of Philosophy at the University of Leicester.

March 2016

A Maria e Corrado

Behavioural Biases and Contract Theory

by

Matteo Foschi

Abstract

This thesis studies how concepts of behavioural biases and bounded rationality affect classical results in contract theory and industrial organization.

Chapter 2 studies the concept of naïveté (Strotz, 1956) in a principal agent model. Agents are assumed to be unaware of their true type, and form biased (naïve) beliefs about it. The latter depend on the actual type of the agent. Results show how the information about agents' true nature, that can be elicited from their beliefs, plays a crucial role in the principal's optimal contracting strategy. In particular, the principal faces a trade-off between exploiting the agent with the most naïve beliefs and designing efficient contracts for the most widespread type of agent, according to her posteriors.

Chapter 3 and Chapter 4 analyse models where agents suffer from temptation and self-control problems (à la Gul and Pesendorfer, 2001). Chapter 3 presents a new justification for loyalty schemes in the retailing industry. In the literature, loyalty schemes have been mostly studied as competition devices (Caminal and Claici, 2007) or as ways to increase consumers' lifetime value (Caminal, 2012). This work focuses on how a seller can use loyalty schemes to acquire information about consumers' preferences and gain the ability to perform individual pricing. Finally, Chapter 4 presents a two-period mechanism design problem with no commitment. It shows how the presence of consumers that suffer from self-control problems can explain the existence of entry bonuses paid by the seller to the consumer, regardless of whether the latter makes the purchase or not.

Acknowledgements

I remember as if it was yesterday the first day I went on Google maps to see what Leicester looked like. Here I am now, almost four years later, with my thesis in my hands (well, keyboard) not knowing where to begin in thanking all the people in my life that made this possible. Therefore, I will start from someone who will probably never read this thesis.

One of the biggest "thank you" goes to my Italian high school teacher of history and philosophy, Prof. Gianluigi Cavaliere. Gianluigi (to whom I refer by first name for the first time in my life) believed in me. Always. No matter my average high school grades, my passion for laziness, my inability to grasp, at the time, the importance of it all, he *always* believed in me. So here I am Prof! This is not the History book you thought I would write one day, but I am sure you would be proud anyway.

Thank you to Alberto Iozzi, whose example I decided to follow. Your advice and support (especially during hard times) were invaluable, I owe you a lot. Thank you also to Gianni De Fraja, for the support and encouragement to apply to Leicester.

Here I have been supervised by two of the best supervisors a PhD student could hope for. The hours spent at the whiteboard in the offices of Chris Wallace and Subir Bose were the most challenging and exciting experiences of my PhD life. I will be forever grateful to you both. To Chris an additional thank you for having taught me how to *actually* write in the English language (although I know I am still far from "good enough" for him).

Also at Leicester, I have to thank all members of the Department for their support. I am particularly grateful to Vincenzo Denicolò, Gaia Garino, Jesse Matheson, Matt Pollison, David Rojo-Arjona and André Stenzel.

I cannot forget to thank the Royal Economic Society to which I owe the funding for my last year of PhD and the financial support that allowed me to attend an amazing workshop in Brazil where I had the chance to meet John Nash, a dream come true.

Thank you to the 5-a-side football team: Nikos ("DEFENSE!"), Ale ("È tuo, è tuo!), Taha, Curro, Sam, Arek, Seb, Andre, Dan, Akis and Johan for making Wednesdays much more bearable.

Lots of love and thank you goes to the amazing friends I found in Leicester. Nikos, remember our promise about 2023! Sam and Mo, you made me feel at home wherever we were, I love you guys. Alex and Maria (who I may soon have the honor to call co-authors too), people are afraid of change, your "change" was one of the best things that could happen to Leicester, and to me. Narges, Vaggelis, Pasqua, Sneha, Taha, Teddy, Daniel, Johan, Ioli, Anna, Nathan, Maria thank you guys for all the days and the nights out together. Last, but far from least, thank you to Sonia, Alessandro and Giacomo, for simply being part of my life, as I hope you will be in the future.

I could not have done this without Sam Hill, you're the best, thank you so much for all your help! Thank you to my students, who made me proud and crazy at the same time. Thank you to the guy that once thought “Ehi, you know what? I’ll open the best Italian Pizzeria this side of the English Channel! And I’ll do it right in front of Matteo’s doorsteps!”. If I could, I would like to thank Linkin Park, whose lyrics I worshipped when I was a teenager, and tell them that: “in the end, it *does* matter”.

All right, I think I have not missed anyone...now to the serious stuff. Caterina, what would I do without you? As we always say “I wouldn’t!”. I am counting the seconds that separate me from calling you my wife, ci siamo quasi! Ti amo. Thank you to my new family, Enzo, Antonella, Roberta e Lorenzo who always make me feel like I belong.

Thank you to my family, that showered me with love since...well I cannot even say since when, because they probably did even before I existed. Mamma, Martina, Ettore thank you for loving me, you made this possible.

Papá...I wish you were here, you were the supervisor of my life, I wish I did not have had to graduate that soon.

Declaration

All Chapters of this thesis are single authored.
Chapter 2 is scheduled to be presented under the title "Contracting with Type- Dependent Naïveté" at the following conferences:

- March 2016: Royal Economic Society Annual Conference , Sussex, UK.
- March 2016: European University Institute, Italy.
- March 2016: META seminar series, York , UK.

Chapter 3 has appeared under the title "Price Discrimination in the Retailing Industry: Inducing Consumers to Reject the Loyalty Card" in the *Social Science Research Network* online. Various working paper versions were presented at the following conferences:

- April 2015: Royal Economic Society Symposium for Junior Researchers, University of Manchester, UK.
- February 2015: Warwick PhD Conference (poster presentation), University of Warwick, UK.

Chapter 3 was presented under the title "Asymmetric Information, Commitment and Self-Control" at the following conference:

- May 2014: SING10, AGH University of Science and Technology, Krakow, Poland.
- August 2014: International Workshops of the Game Theory Society (poster presentation), FEA-USP, Universidade de São Paolo, São Paolo, Brazil.

Contents

List of Figures

Chapter 1

Introduction

Classical economics assumes that individuals always behave as if perfectly rational. *Homo economicus* is capable of understanding everything around him and to estimate uncertain events with precision. The literature on bounded rationality and behavioural biases studies the partial relaxation of this hypothesis. This thesis builds on previous contributions in this area, with particular focus on temptation and self-control (Gul and Pesendorfer, 2001), and naïveté (Strotz, 1956). The following Chapters present applications of these concepts in contexts of contract theory and industrial organisation, and show how they can strongly affect the results of classical economics.

Chapter 2 studies the case of individuals who are unaware of their true nature, i.e. they are naïve. Bridging the literature on contracting with naïve agents (among others Eliaz and Spiegler, 2006, 2008) and on sequential screening (Courty and Li, 2000), it introduces the assumption that naïveté of individuals depends on the same personal characteristics that they are unable to estimate. Individuals have competence-dependent levels of confidence, which can be used by a counter-party to elicit information about their true nature.

Chapter 3 and 4 focus on temptation and self-control, and on the way they affect the pricing (and non-pricing) strategies of sellers. The most interesting aspect of temptation models is how individuals with high valuations of a good, i.e. the "high" types in contract theory models, also suffer from the strongest self-control problems. These decrease their willingness to pay and, therefore, their *value* as customers in the eyes of the seller. Chapter 3 shows how this may induce a retailer to offer loyalty schemes to consumers. These schemes work as commitment devices. By rejecting them, high type consumers are sure to face a very high price for tempting goods once in front of the purchase decision. This price is set so high that they will not be tempted to buy the good to begin with. Chapter 4, instead, shows how given the self-control problems of consumers, in equilibrium, "high types" may play the role of "low types" and vice versa. Hence, the former may end up being excluded from the market, while the latter may enjoy a positive information rent.

Alongside the above, Chapter 3 and 4 also contribute to the industrial organisation literature. Chapter 3 studies the pricing behaviour of a seller that offers consumers personalised discounts (via loyalty schemes) in exchange for a certain degree of observability over their preferences. This work provides, to the best of the author's knowledge, one of the first theoretical frameworks for individual pricing with consumer tracking.[1] In contrast to classical results on first degree price discrimination — the closest relative of individual pricing — it shows how, under certain conditions, individual pricing may be welfare enhancing.

Chapter 4 studies online markets where temptation plays a crucial role, like casinos or betting web-sites. It presents a mechanism design problem with no commitment in order to explain seemingly puzzling common practices among retailers, such as offering entry bonuses with no strings attached to all consumers, regardless of whether they buy or not.

Finally, Chapter 5 briefly concludes the thesis and suggests directions for further research.

[1]For references on individual pricing, see Shiller and Waldfogel (2011); Mikians, Gyarmati, Erramilli, and Laoutaris (2012, 2013); Hannak, Soeller, Lazer, Mislove, and Wilson (2014); Waldfogel (2015); Shiller (2015)

Chapter 2

Contracting with Type-Dependent Naïveté

Chapter Abstract

I analyse the optimal contracting behaviour of an employer who faces workers with different, incorrect beliefs about their own productivity. While the literature has focused mostly on the exploitative (when the principal knows agents' types, Eliaz and Spiegler, 2006) and speculative (when the principal has priors on agents' types, Eliaz and Spiegler, 2008) aspects of contracts, I introduce the assumption that workers' naïveté depends on their actual productivity level. The employer uses this information to form posteriors on agents' productivity and design more efficient contracts. In particular, I highlight the employer's trade-off between exploiting strongly naïve workers and designing efficient contracts for the most widespread type of worker, according to her posteriors.

2.1 Introduction

I study a two period principal-agent model where agents are hired in period 1 to carry out a task in period 2. Before facing the task they are assigned to, however, agents have limited information about their true type and they are assumed to form biased (wrong) beliefs about it in period 1 — i.e., they are naïve. The main contribution of the paper is to study the optimal contracting behavior of the principal when workers' beliefs depend on their true type — that is, when naïveté is type-dependent. In equilibrium, in period 1, the principal screens among workers with different beliefs, to take advantage of this extra information, and form posteriors on workers' productivity. This allows her to design contracts that are more efficient than when the agents' beliefs are independent from their types, and can exploit agents to a greater degree.

When facing a new task, individuals form expectations about their own ability to carry it out, and about the amount of effort required. Typically, individuals are assumed always to hold unbiased beliefs about their abilities. Often, however, this is not the case. From the workman estimating the time to build a wall to the athlete who forms expectations about the amount of effort to achieve a specific goal, the final result is not always the one expected. In economics we often assume that such "errors" result from specific realisations of random variables either side of the unbiased expectation.

However, estimations can be distorted by one's wrong perception of the situation, or by one's firm beliefs that turn out to be inconsistent with reality. Economics deals with these kinds of situation with the concept of *naïveté* (Strotz, 1956), that is, the inability of an individual to form unbiased expectations about an unknown event. In other words, the systematic over- or underestimation of the realisation of a random variable.

In this paper, I investigate situations where the level of naïveté of an individual depends on his own innate ability. In particular, I study workers who have systematically wrong (naïve) beliefs about their own productivity, which is the realisation of a random variable. Differently from the existing literature — described in section 2.2 — I make the novel assumption that workers' naïveté depends on their own ability (their *type*). On the other hand, the employer, who is perfectly unbiased, designs contracts to hire the workers and can exploit their naïveté.[1] Besides a surplus extraction motive, in order to maximise profits, she is interested in using this information in order to design more efficient contracts.

While it is perfectly reasonable to assume that ability and beliefs are independent — Eliaz and Spiegler (2006, 2008) — it is often the case that this assumption is violated. To illustrate this idea, consider the following example.

Suppose a population of high school students is about to enrol in university, and each student has to choose the right course for him or her. Suppose further that the population can be divided into good students and bad students. All students are unaware of their true ability to succeed at university level until they actually face the lectures, tutorials and coursework. Hence, they form expectations about it and choose their course accordingly. Once they start their courses, they understand their true ability and choose the level of effort to exert before facing the exams. While it may be perfectly reasonable to assume that the student's expectations are independent from their true ability, here students' expectation's bias derives from their innate capabilities. Hence, for example, one can think of the case where good students are naturally more self-confident and self-aware, while bad students are shy and insecure. The former will therefore pick a much more challenging course and succeed, while the latter will pick a less demanding course in order to achieve success. At the other end of the spectrum is the case where ability makes a student aware of the difficulties and complexities of university, generating a pessimistic feeling about his ability to succeed. Hence, a good student would pick a relatively less challenging course and perform strongly beyond his expectations. A bad student, on the other hand, may misunderstand or underestimate the challenges of university and sign up for a relatively difficult course, failing and eventually dropping out of school.

The main message of the paper lies in the importance of the employers' posteriors. General results of the literature on diversely naïve agents (Eliaz and Spiegler, 2006, 2008) emphasise the ability of a principal to take advantage of agents' biased beliefs by achieving the "efficient" outcome at a lower cost, relatively to when agents have unbiased beliefs or are fully informed. This efficiency, however, is achieved

[1]The assumption about the employer having unbiased belief can be thought has her having more experience, or knowing better the suitability of the workers population for the specific job she is hiring for. Ultimately, dropping this assumption simply changes the interpretation of the model, but not its results.

in the states of nature that the principal deems more probable than the agent, according to her priors. While in my paper this result still stands for a portion of the parameter space, it fails to hold more generally, because of the principal's updating of her priors.

The employer designs contracts that, first, screen among differently naïve agents in period 1 and, second, screen among different types of agents in period 2. The key contribution of the paper is owed to the updating of the principal's prior from period 1 to period 2. Given the screening of period 1, the employer updates her beliefs according to the correlation between workers' beliefs and abilities. This originates a trade-off for her: to design efficient contracts either for the most naïve types, or for the ones she deems most probable given her posteriors. The main result of the paper is to show how the efficiency of optimal contracts changes given this new trade-off. In particular, I show how the principal may find it optimal to design efficient contracts for the type she deems most probable according to her updated beliefs, *regardless* of the type's naïveté.

The paper is organised as follows. In section 2.2 I present the related literature. In section 2.3 I explain the model and the assumptions. In section 2.4 I study the case of perfect correlation between naïveté and agents' types. I then relax this assumption and study the general case in section 2.5. I conclude the paper in section 5. All the proofs of Lemmas, Results and Propositions are relegated to the Appendix.

2.2 Related Literature

Extensive experimental evidence motivates the main assumption behind this work. A first set of papers (among others: Svenson, 1981; Chi, Glaser, and Rees, 1982; Dunning and Kruger, 1999; Dunning, Ehrlinger, Johnson, and Kruger, 2003; Banner, Dunning, Ehrlinger, Johnson, and Kruger, 2008) show that the skills needed to evaluate competence in a specific domain are exactly the same required to engender this competence. Hence, individuals without such skills would find it relatively hard to estimate their own competence correctly. Building on these findings, a second set of papers (Dittrich, Güth, and Marciejovsky, 2005; Banks, Lawson, and Logvin, 2007; Moore and Healy, 2008; Ferraro, 2010) present further experimental evidence on the positive correlation between competence and self-awareness. Finally, a third set of papers (Lichtenstein, Fischoff, and Phillips, 1982; Loewenstein, O'Donoghue, and Rabin, 2003; Conlin, O'Donoghue, and Vogelsang, 2007) focuses on the concept of "over-confidence" and projection-bias providing strong experimental evidence on the bias of individual's expectations.[2]

The contributions listed above highlight two main facts: (i) individuals are not perfectly capable of estimating their own skills and (ii) often their estimation of their capabilities depend on the same skills they are trying to evaluate. To date, the economics literature has dealt with these facts only separately.

Sequential screening of consumers who do not know their true valuation of a good was studied by Courty and Li (2000). They propose a model where agents hold unbiased beliefs about their type, but the precision of their estimation depends on their type itself. In line with the results of this paper, Courty and Li (2000) find that optimal contract design depends on the "informativeness" of initial knowledge of agents rather than on the principal's priors. Differently from the present work, however, they assume non-naïve agents — i.e. agents of unbiased expectations — leaving no space for exploitation. On the contrary, the optimal mechanism features "refund contracts", that grant agents the option to claim a refund after they learn their true willingness to pay. In recent years, the model of Courty and Li (2000) has been extended and applied. Among others, Kovác and Krähmer (2013) study sequential delegation, Deb and Said (2015) study the case of a principal with limited commitment power, Evans and Reiche (2015) relax the commitment assumption completely, Grubb (2009) applies the model to the cellular phone service market.

Self-awareness and naïveté were first introduced by Strotz (1956) and applied in contract theory later on. O'Donoghue and Rabin (2001), Asheim (2008) and Heidhues and Köszegi (2010) (among others) study the interaction between naïveté and self-control, modeled as present-biased preferences. Amador, Werning, and Angeletos (2006) analyse the trade-off faced by a multi-self agent who is aware of his time-inconsistency problems, but is not aware of his true preferences until later periods. Gilpatric (2008) studies the problem of moral hazard in the presence of

[2] Less related, Bagues and Perez-Villadoniga (2012) show that these findings extend to the estimation of other people's skills as well. They use evidence from a field experiment to show how recruiters prefer to hire applicants with capabilities to their own. One of the proposed explanations is that evaluators' accuracy is higher when evaluating those dimensions in which their knowledge is greater.

naïvé agents with time-inconsistent preferences.[3]

The papers that most relate to this one are Eliaz and Spiegler (2006) and Eliaz and Spiegler (2008). In both papers, time-inconsistent agents differ in their level of naïveté, with some of them being perfectly self-aware. In Eliaz and Spiegler (2006), the employer has full information about consumers' preferences. The optimal menu provides a commitment device for self-aware agents, who would like to play according to their present preferences, as opposed to their future preferences. Relatively naïve agents, instead, are exploited because of their inability to correctly estimate their actual type. In Eliaz and Spiegler (2008), the authors extend the model to one where the employer has priors over consumers' preferences-change. Hence, two screening processes take place, exactly as in this paper. The first screening separates differently self-aware agents; the second separates with respect to their preferences. In both papers, however, agents' beliefs and types are assumed to be independent.

My model builds on these contributions to study situations where types (or preferences) affect agents' beliefs. My results bridge the findings of screening models with diversely naïve agents (as in Eliaz and Spiegler, 2006, 2008) with the ones of sequential screening (as in Courty and Li, 2000), providing a new perspective on the connections between these two literatures.

2.3 The Model

An employer (the principal, she) seeks to hire a worker (the agent, he) from a population. Workers are hired in period 1 and asked to complete an individual task in period 2. The outcome of the task depends on the level of effort $e \in [0,1]$ a worker exerts then. I assume that the level of effort exerted by the worker is perfectly observable.[4]

To hire workers, the employer, in period 1, offers a set of contracts $w(e) : [0,1] \rightarrow \mathbb{R}$ that each worker can either accept or reject. When a worker accepts a contract in period 1, and exerts effort e in exchange for wage $w(e)$ in period 2, the employer enjoys profits $\Pi = y(e) - w(e)$, where $y(e)$ is increasing and concave in e.

[3]Further, but less related, is Von Thadden and Zhao (2012) that study a classical principal agent model where agents do not know their action space until a later stage.

[4]Extending the model to a moral hazard framework where e is partially, or not at all, observable, is left for future research.

When a worker accepts the contract, he enjoys utility $U_j = w(e) - \theta_j e$, where θ_j is the cost of effort and represents a worker's *productivity type*. Finally, if a worker rejects a contract, both he and the employer obtain zero utility/profits.

The population of workers is composed of a portion λ of *productive* types, who have $\theta_j = \theta_P$, and a portion $(1-\lambda)$ of *unproductive* types, who have $\theta_j = \theta_U > \theta_P$.

The first main assumption of the paper is that in period 1 neither the employer nor the workers are aware of a worker's productivity type. While the employer forms unbiased expectation, however, workers have biased heterogeneous beliefs about themselves, that is, they are *naïve*. Given this, the employer's expectation about a worker's utility is given by $E(\theta) = \lambda U_P + (1-\lambda)U_U$. A worker's belief about his own utility, instead, depends on his *belief type*. A worker can be *optimistic* or *pessimistic* about his true productivity. In the first case, the agent believes himself to be a productive type with probability $\phi > \lambda$, that is $\Pr\{\theta_j = \theta_P\} = \phi$. In the second case, he believes himself to be a productive type with probability $\delta < \lambda$, that is $\Pr\{\theta_j = \theta_P\} = \delta$. An i-belief type expects his productivity to be $E_i = i\theta_P + (1-i)\theta_U$, $i = \{\phi, \delta\}$. Notice that an agent is considered optimistic (pessimistic) with respect to the average of the population and not with respect to his actual productivity.

The second main assumption of the paper, and the one that constitutes the main departure from the literature, states that a worker's beliefs and productivity are not independent. Here, I assume that the distribution of belief types is conditional on a worker's true productivity. In particular, there is a proportion p_P (p_U) of pessimistic types among productive (unproductive) workers. Hence, the employer has priors: $p_P = \Pr\{\delta|\theta = \theta_P\}$ and $p_U = \Pr\{\delta|\theta = \theta_U\}$. This allows her to update her priors on a worker's productivity when she knows his belief type.

Workers update their prior only when they face the task. In period 2, they learn their true productivity *before* choosing the level of effort to exert.

Given the assumptions above, the employer faces two different connected screening problems. In period 1 she wants to separate workers according to their belief type. This allows her to update her priors in period 2 and separate workers on the basis of their productivity type. Notice that the employer and the agents have always different beliefs throughout the game. In period 1, the employer forms unbiased expectations, while workers rely on their naïve beliefs. In period 2, the employer updates her priors given the separation of period 1, while workers learn their true productivity and behave as fully informed agents. This implies that the maximisation problem the employer solves is subject to period 1 constraints, that depend on workers' belief type, and period 2 constraints that depend on workers' true productivity type.

Before stating the problem formally, I define $(w_i^j, e_i^j) \equiv (w_i(e_i^j), e_i^j)$ as the wage and effort level that a worker of i-belief type and j-productivity type chooses in period 2. Notice that workers' utility depends only on the level of effort they choose (or believe they will choose) in period 2, and that once they sign a contract they are constrained to carry out the task — i.e. there is no individual rationality constraint in period 2. Therefore, I can restrict my attention, without loss of generality, to four effort levels, and the corresponding wages set by the employer: $e_\delta^U, e_\delta^P, e_\phi^U, e_\phi^P$.

Given this, the employer solves:

$$\max_{\{w_i^j\}_{i=\delta,\phi,j=P,U}} \quad E(\Pi) \tag{2.1}$$

$$\begin{aligned}
\text{s.t.} \quad & E_\delta(U_j(w_\delta(e))) \geq 0, && (IR_\delta)\\
& E_\phi(U_j(w_\phi(e))) \geq 0, && (IR_\phi)\\
& E_\delta(U_j(w_\delta(e))) \geq E_\delta(U_j(w_\phi(e))), && (IC_\delta)\\
& E_\phi(U_j(w_\phi(e))) \geq E_\phi(U_j(w_\delta(e))), && (IC_\phi)\\
& U_P(w_\delta^P, e_\delta^P) \geq U_P(w_\delta^U, e_\delta^U), && (IC_{P,\delta})\\
& U_U(w_\delta^U, e_\delta^U) \geq U_U(w_\delta^P, e_\delta^P), && (IC_{U,\delta})\\
& U_P(w_\phi^P, e_\phi^P) \geq U_P(w_\phi^U, e_\phi^U), && (IC_{P,\phi})\\
& U_U(w_\phi^U, e_\phi^U) \geq U_U(w_\phi^P, e_\phi^P). && (IC_{U,\phi})
\end{aligned}$$

She maximizes her expected profits with respect to two different contracts: $w_\delta = \{(w_\delta^P, e_\delta^P), (w_\delta^U, e_\delta^U)\}$ and $w_\phi = \{(w_\phi^P, e_\phi^P), (w_\phi^U, e_\phi^U)\}$. These contracts induce separation among belief types in period 1 and among productivity types in period 2. In order to achieve this, the contracts have to satisfy eight different constraints.

The first two are period 1 individual rationality constraints that ensure that each belief type is willing to accept the contract designed for him as opposed to his outside option. The second two are period 1 incentive compatibility constraints that induce separation among belief types. Notice that since these four constraints relate to period 1, they are expressed in expected utility terms, and the expectations are weighted by workers' beliefs.

The last two pairs of constraints are "contract specific" period 2 incentive compatibility constraints. They ensure that belief type i, once he has self-selected in period 1 and learned his true productivity in period 2, chooses the wage/effort pair designed for him. Hence, they are expressed in the actual utility the worker obtains.

Notice that the principal does not have to satisfy any period 2 individual rationality constraint since it is assumed that workers cannot "drop out" of the contract once it has been signed in period 1.

In the next sections, I solve the problem for the optimal set of contracts offered by the employer. I do this under different assumptions about the level of information obtained by knowing a worker's belief type. I start with the case of perfect correlation between the two type dimensions, i.e. when beliefs are perfectly informative about workers' productivity, so that separation in period 1 perfectly reveals the agent's productivity.

Before doing that, however, in order to better express the results, let me define the concepts of *exploitation* and *efficiency*, in line with the existing literature's terminology.

Definition 2.1 (Exploitation)**.** *A worker of i-belief type and j-productivity type is exploited if*

$$w_i^j - \theta_j e_i^j < 0.$$

That is, he is exploited if he accepts a contract $w_i(e)$ that a fully informed agent of his same productivity type would not accept.

The concept of exploitation was first introduced in Eliaz and Spiegler (2006). It generally applies to a situation where a principal takes advantage of an agent's naïveté in order to extract surplus from him beyond the limits of the IR. In the context of this paper, a worker may be exploited not because he does not know his true ability (although they do not), but rather because he has systematically wrong beliefs about it.

Secondly, I define efficient levels of effort as the values of e that equate marginal product to workers' productivity.[5]

[5] These are the effort levels exerted when agents are uninformed about their true type, but are not naïve (Harris and Raviv, 1979; Laffont and Martimort, 2002).

Definition 2.2 (Efficient Effort)**.** *A worker of i-belief type and j-productivity type exerts* efficient *effort if*

$$e_i^j : y'(e_i^j) = \theta_j.$$

That is, if at e_i^j the marginal product of effort equals the worker's productivity.

Given this, a contract $w_i(e)$ may induce either productive or unproductive workers (or both) to exert efficient levels of effort. Hence, the definition of efficiency at the top and at the bottom:

Definition 2.3 (Top vs. Bottom Efficiency)**.** *A contract $w_i(e)$ features* efficiency at the top *if it induces productive workers to exert the efficient level of effort, i.e. e_i^P : $y'(e_i^P) = \theta_P$. It features* efficiency at the bottom *if it induces unproductive workers to exert the efficient level of effort, i.e. e_i^U : $y'(e_i^U) = \theta_U$.*

2.4 Perfectly Informative Beliefs

In this section, I study the case where knowing a worker's belief perfectly reveals his productivity type. This scenario requires that p_P and p_U belong to $\{0,1\}$ and $p_P + p_U = 1$. That is, either all productive workers are optimistic, and all unproductive workers are pessimistic, or vice-versa. If all workers, regardless of their productivity, were pessimistic (or optimistic) then no information could be learned by knowing their beliefs.

The first result shows that separation in period 1 is not affected by the extent, or direction, of belief-productivity correlation. This is because, at this stage, the employer has only a prior on workers' belief types and cannot exploit the correlation between beliefs and productivity. The next Lemma identifies the binding constraints of period 1.

Lemma 2.1 (Period 1 Screening)**.** *Regardless of the correlation between naïveté and productivity, the period 1 constraints are such that:*

(i) (IR_δ) *binds while* (IR_ϕ) *is slack.*

(ii) (IC_ϕ) *binds while* (IC_δ) *is slack.*

Lemma 2.1 presents findings similar to a classical screening model. In this case, the optimistic type plays the role of the "high type" while the pessimistic type the role of the "low type". To see this, notice that what determines period 1's type ranking — high vs. low — is not worker's actual productivity, but rather their subjective expectations about it. Therefore, optimistic (pessimistic) workers play the role of the high (low) type in the population. As in classical screening problems, optimistic workers' IR is slack as is the IC of pessimistic types.

Notice, also, that the employer's only purpose of inducing separation in period 1 is to be able to form posteriors on workers' productivity, since she gains no direct profits from this separation.

Given the above, and substituting for the expected profits and utilities, the problem that the employer solves is reduced to:

$$\begin{aligned} \max_{\{w_j^i\}_{j=\delta,\phi,i=P,L}} \quad & \lambda\Big[p_P(y(e_\delta^P)-w_\delta^P)+(1-p_P)(y(e_\phi^P)-w_\phi^P)\Big]+ \\ & +(1-\lambda)\Big[p_U(y(e_\delta^U)-w_\delta^U)+(1-p_U)(y(e_\phi^U)-w_\phi^U)\Big] \end{aligned} \tag{2.2}$$

$$\text{s.t.} \quad \delta(w_\delta^P-\theta_P e_\delta^P)+(1-\delta)(w_\delta^U-\theta_U e_\delta^U)=0 \tag{IR_δ}$$

$$\phi(w_\phi^P-\theta_P e_\phi^P)+(1-\phi)(w_\phi^U-\theta_U e_\phi^U)=\phi(w_\delta^P-\theta_P e_\delta^P)+(1-\phi)(w_\delta^U-\theta_U e_\delta^U) \tag{IC_ϕ}$$

$$w_\delta^P-w_\delta^U \geq \theta_P(e_\delta^P-e_\delta^U) \tag{$IC_{P,\delta}$}$$

$$w_\delta^P-w_\delta^U \leq \theta_U(e_\delta^P-e_\delta^U) \tag{$IC_{U,\delta}$}$$

$$w_\phi^P-w_\phi^U \geq \theta_P(e_\phi^P-e_\phi^U) \tag{$IC_{P,\phi}$}$$

$$w_\phi^P-w_\phi^U \leq \theta_U(e_\phi^P-e_\phi^U). \tag{$IC_{U,\phi}$}$$

As expected, period 2 separation depends on the correlation between beliefs and productivity. The reason for this is that the period 2 ICs are contract specific. Hence, their relevance depends on the posterior the employer forms on a specific belief type's productivity.

I now present the results under two different scenarios of perfect correlation. In the first scenario, I study *positive correlation* between beliefs and productivity, that is, the case of optimistic-productive and pessimistic-unproductive workers. In the second scenario, I study the opposite case of *negative correlation*. These scenarios have both specific features, relevant only in the special case of this section, as well as more general features revisited in section 2.5.

2.4.1 Perfect Positive Correlation.

In this section I study the case of $p_P = 0$ and $p_U = 1$. In this case the only types present in the labour force population are productive optimistic, (P, ϕ), and unproductive pessimistic, (U, δ). While the employer understands this and behaves accordingly, workers still believe to be part of a population with four different types of workers. That is, they fail to understand that their beliefs are a perfect indicator of their actual productivity. This creates an opportunity for the employer to take advantage of workers' naïveté and exploit them.

Take the contract for pessimistic workers, for example. The employer sets the contracts in period 1, when workers are unaware of their true productivity. While she knows, however, that each pessimistic worker is unproductive, the latter believes himself to be productive with a positive probability. This creates two "channels" of exploitation for the employer to use.

First, when facing a contract that extracts full surplus from unproductive types, a pessimistic unproductive worker expects to obtain positive utility from it. To see this, consider any contract for pessimistic types that does not separate between productivity types and offers $w_\delta(e_\delta^U) = \theta_U e_\delta^U$. In period 1, an unproductive pessimistic worker evaluates this contract with $E_\delta(U_i(w_\delta(e))) = \theta_U e_\delta^U - E_\delta(\theta) e_\delta^U > 0$. Given this, the employer can decrease the wage even further, increasing profits, until the (IR_δ) binds. A this point, the worker in period 1 expects to obtain zero utility, while in period 2 he faces the truth and obtains negative utility. He is exploited through the first channel.

The second channel of exploitation takes place through an "imaginary offer" (Eliaz and Spiegler, 2008).

Definition 2.4 (Imaginary Offer)**.** *An* imaginary offer *is a pair* (w_i^j, e_i^j) *never chosen by any worker in equilibrium, but that workers believe they will choose with positive probability because of their naïveté.*

The imaginary offer satisfies incentive compatibility and it is used by the employer to increase the expected utility of a worker from contract $w_i(e)$, while not increasing the actual utility he obtains. In this section, the imaginary offer in the contract for the pessimistic worker is set to yield a positive surplus to a productive type. If it is added to contract $w_\delta(e)$, in fact, the pessimistic worker assigns a positive probability to the event of choosing it (and of being a productive type). This increases his expected utility from contract $w_\delta(e)$ and allows the employer to decrease the utility given by the "actual offer" even further, increasing exploitation.

To avoid any confusion, notice that imaginary offers do not affect profits directly but only through workers' naïveté. In other words, the employer knows that these offers are never chosen, hence, they are assigned no positive probability in the expectations of profits. Since workers, however, believe they may choose these offers in period 2 with some positive probability, the equilibrium values of the actual offers depend on the imaginary offers. Hence, their effect on profits is indirect.

Given the two channels, I define two possible levels of exploitation.

Definition 2.5 (Mild vs. Strong Exploitation)**.** *Exploitation is* mild *if the employer does not take advantage of the second channel — i.e. she does not design an imaginary offer. It is* strong *if she does so.*

In the case of positive perfect correlation between type dimensions, it is straightforward to understand that the period 2 binding constraints are $(IC_{P,\phi})$ and $(IC_{U,\delta})$. This is because they are intended for the only types that actually exist in the population. Nevertheless, in order to take advantage of the two channels of exploitation, $(IC_{U,\phi})$ and $(IC_{P,\delta})$ should still hold.

Hence, when belief and productivity are perfectly positively correlated, the employer solves:

$$\max_{\{w_i^j\}_{i=\delta,\phi,j=P,L}} \lambda(y(e_\phi^P) - w_\phi^P) + (1-\lambda)(y(e_\delta^U) - w_\delta^U) \tag{2.3}$$

$$\text{s.t.} \quad \delta(w_\delta^P - \theta_P e_\delta^P) + (1-\delta)(w_\delta^U - \theta_U e_\delta^U) = 0 \tag{IR_δ}$$

$$\phi(w_\phi^P - \theta_P e_\phi^P) + (1-\phi)(w_\phi^U - \theta_U e_\phi^U) = \phi(w_\delta^P - \theta_P e_\delta^P) + (1-\phi)(w_\delta^U - \theta_U e_\delta^U) \tag{IC_ϕ}$$

$$w_\delta^P - w_\delta^U = \theta_U(e_\delta^P - e_\delta^U) \tag{$IC_{U,\delta}$}$$

$$w_\phi^P - w_\phi^U = \theta_P(e_\phi^P - e_\phi^U) \tag{$IC_{P,\phi}$}$$

$$w_\phi^P - w_\phi^U \leq \theta_U(e_\phi^P - e_\phi^U) \tag{$IC_{U,\phi}$}$$

$$w_\delta^P - w_\delta^U \geq \theta_P(e_\delta^P - e_\delta^U). \tag{$IC_{P,\delta}$}$$

First of all, notice that period 2 incentive compatibility for contract $w_i(e)$ is possible as long as $(e_i^P - e_i^U) \geq 0$. Once again, recall that this incentive compatibility is only "imaginary" since there are no workers with different productivity but same belief type.

Solving the binding constraints for $w_\delta^P, w_\delta^U, w_\phi^P, w_\phi^U$,

$$\begin{aligned}
w_\delta^U &= E_\delta(\theta)e_\delta^P + \theta_U(e_\delta^U - e_\delta^P),\\
w_\delta^P &= E_\delta(\theta)e_\delta^P,\\
w_\phi^U &= (E_\delta(\theta) - E_\phi(\theta))e_\delta^P + E_\phi(\theta)e_\phi^U,\\
w_\phi^P &= (E_\delta(\theta) - E_\phi(\theta))e_\delta^P + E_\phi(\theta)e_\phi^U - \theta_P(e_\phi^U - e_\phi^P),
\end{aligned}$$

and substituting the relevant solutions in the maximisation, I obtain:

$$\begin{aligned}
\max_{\{e_j^i\}_{j=\delta,\phi,i=P,L}} \quad & \lambda\Big(y(e_\phi^P) - (E_\delta(\theta) - E_\phi(\theta))e_\delta^P - E_\phi(\theta)e_\phi^U - \theta_P(e_\phi^P - e_\phi^U)\Big) + \\
& +(1-\lambda)\big(y(e_\delta^U) - E_\delta(\theta)e_\delta^P - \theta_U(e_\delta^U - e_\delta^P)\big). \qquad (2.4)
\end{aligned}$$

From this new problem, the actually chosen levels of effort e_ϕ^P and e_δ^U are easily calculated to be $y'(e_\phi^P) = \theta_P$ and $y'(e_\delta^U) = \theta_U$. Hence, each worker hired exerts the efficient level of effort for his productivity type. This result is common with the case of negatively correlated beliefs and productivity, and it is generalised in Proposition 2.2 in the next section.

The values of the imaginary offers can also be derived from the maximisation problem. Starting from the effort level for unproductive optimistic workers, it is easy to see that the effect of e_ϕ^U on profits is negative. Hence, in equilibrium $e_\phi^U = 0$. The intuition behind this is that the (w_ϕ^P, e_ϕ^P) offer is already inducing an efficient level of effort. Since the agent believes himself to be an unproductive type with some positive probability, the imaginary action has to require a *low* level of effort. In this way, the worker feels "safe" that in case she turns out to be unproductive, she can always enjoy a small surplus without exerting too much effort; in fact, no effort at all: $e_\phi^U = 0$.[6]

The intuition behind the optimal value for e_δ^P, and its derivation, instead, are not as straightforward. On the one hand, a lower e_δ^P for a given w_δ^P increases $E_\delta(U_i(w_\delta(e)))$.

[6]To see that $U_\phi^U > 0$ notice that $e_\delta^P > 0$ as described below. Hence, $w_\phi^U > 0$ even if $e_\phi^U = 0$

This relaxes the IR of the pessimistic unproductive worker, and allows the employer to decrease even further the wage paid to this type, increasing profits. On the other hand, this also increases $E_\phi(U_i(w_\delta(e)))$, violating the (IC) of the optimistic productive worker. This forces the employer to increase $E_\phi(U_i(w_\phi(e)))$ by the same amount, decreasing her profits. Which of these two opposite effects prevails, depends on the effect of e_δ^P on (2.4). If it is positive, e_δ^P is set to the highest possible value, 1. If the effect is negative, e_δ^P is set to the lowest possible value. Finally, however, notice that for incentive compatibility to be possible — i.e. for $(IC_{U,\delta})$ and $IC_{P,\delta})$ to hold, e_δ^P cannot go below the value of e_δ^U. Hence, whether the effect of e_δ^P on profits is positive or negative does not determine the "direction" of the imaginary offer, but rather whether the offer exists or not. In other words, it determines whether the pessimistic worker expects to be screened or pooled in period 2.

The effect of a decrease in e_δ^P on profits, discussed above, depends on the ratio of workers' beliefs, i.e. the overall level of naïveté of the workers population. In particular, the higher the naïveté of the optimistic productive worker, the more he believes to be unproductive. Hence, the lower is the positive effect on $E_\phi(U_i(w_\delta(e)))$ of a decrease in e_δ^P and the stronger is the second channel of exploitation on him. This allows the principal to let (IC_ϕ) bind again via (w_ϕ^U, e_ϕ^U), which does not affect her profits directly. Similarly, if ϕ is high and the productive type is "self-aware" — i.e. $E_\phi(\theta)$ gets closer to ϕ_P —, (w_ϕ^U, e_ϕ^U) has a weaker effect on $E_\phi(U_i(w_\phi(e)))$. Therefore, decreasing e_δ^P becomes more costly. Hence, for the use of the second channel to be optimal, the pessimistic unproductive worker has to be naïve enough.

The above is summarised in Result 1.

Result 1 (Pooling of Pessimistic Workers)**.** *When beliefs and productivity are perfectly positively correlated, if the pessimistic unproductive worker is naïve enough, relative to the self-awareness of the optimistic productive worker, that is:*

$$\frac{\delta}{\phi} \geq \lambda, \tag{2.5}$$

then the employer uses an imaginary offer (w_δ^P, e_δ^P) *for the unproductive type and exploitation is strong. If not, the employer uses no imaginary offer and exploitation is mild.*

To fully understand Result 1 consider the following. Notice that the LHS of condition (2.5) corresponds to the naïveté of pessimistic unproductive workers over the self-awareness of optimistic productive workers. Hence, it can be interpreted as a measure of relative workers' naïveté in the population. The higher it is, the more the unproductive type believes himself to be productive and the more the productive worker believes himself to be unproductive. For the condition to hold, the probability a pessimistic unproductive worker assigns to himself being productive has to be larger than the proportion of productive workers in the population. When condition (2.5) holds, $e^P_{\delta} = 1$ and $w^P_{\delta} = E_{\delta}(\theta)$ and exploitation is strong.

On the contrary, when condition (2.5) does not hold, either the optimistic productive worker is too self-aware or the pessimistic unproductive worker is not naïve enough. The latter's relative naïveté is lower than the proportion of productive workers in the population and it is, therefore, too costly to exploit him through the second channel.

Given the structure of the contracts described so far, the next Result studies workers' welfare.

Result 2 (Exploitation with Perfect Positive Correlation)**.** *When beliefs and productivity are perfectly positively correlated, productive workers enjoy a positive rent while unproductive ones are exploited. The rent enjoyed by the former is larger when exploitation of the latter is strong.*

The intuition behind Result 2 follows from the previous discussion. The employer wants to take advantage of the unproductive worker's naïveté. Since beliefs and productivity are positively, perfectly correlated, however, she is forced to leave some positive surplus to the productive worker. The reason for this lies both in the difference in beliefs between belief-types and in the way the two channels of exploitation affect the two contracts.

On the one hand, the first channel cannot be used to extract surplus from the productive worker. The reason for this is that the latter expects to have a lower productivity than he actually has: he never accepts a contract that extracts full surplus from a productive type because he would expect to get a strictly negative utility from doing so.

On the other hand, the role of the imaginary offer in the contract for the optimistic type is not to extract more surplus from him but rather to provide him with a form of "insurance". In other words, it represents a safe option for the optimistic productive worker to choose in the case he turns out to be unproductive — an event that he deems possible with positive probability. Hence, the employer has no ability to exploit productive workers.

Finally, to understand the intuition behind optimistic workers' surplus, notice that an optimistic worker always assigns a larger probability to obtaining U_δ^P if he chooses $w_\delta(e)$ than the pessimistic worker. Hence, any change to $w_\delta(e)$ that increases U_δ^P keeping $E_\delta(U_j(w_\delta(e)))$ constant, increases $E_\phi(U_j(w_\delta(e)))$. For (IC_ϕ) to bind, therefore, the utility from the contract for the optimistic type has to increase when the imaginary offer is added to $w_\delta(e)$.

Notice that this is in accordance with the intuition behind (2.5). The higher is the general level of naïveté in the labour force population — i.e. (2.5) holds — the stronger is the exploitation of the unproductive worker.

2.4.2 Perfect Negative Correlation

In this section, I study the opposite case to the one of section 2.4.1, namely of $p_P = 1$ and $p_U = 0$. That is, the case where all productive workers are pessimistic and all unproductive workers are optimistic.

To understand the framework I have in mind consider the case of a population of newly graduated students looking for their first job. Grades and degrees can explain a lot about knowledge of the topics and intelligence, but when it comes to innate ability, speed of adaptation, productivity and so on, there is nothing like true practice to give an indication of one's capabilities. Suppose that the students have a degree in financial economics and all look for a job in the financial sector. They all apply for jobs according only to their expectations about their own productivity. Once a job is obtained, however, they learn their true productivity and choose the amount of effort to exert in the job accordingly. Some of the students have a passion for finance, they read the news, understand the mechanics and complexities of markets and would be perfect for a job in the financial sector (productive types). Others, instead, have chosen that specific course of study without having a deep interest in financial markets. Hence, they would be a less perfect match for a financial firm (unproductive types). In this section, I assume that understanding the complexities of the job and the mechanisms of financial markets, without having a clear perception of one's own capability, hurts self-confidence and creates a general pessimistic feeling about one own's success in the market (pessimistic productive workers). A candidate who does not comprehend these complexities, instead, has a relatively "arrogant" attitude. He is convinced that the job will be easy (optimistic unproductive workers).

To derive the solution to this problem, it is easy to follow the same procedure of section 2.4.1 where now, however, the period 2 binding constraints are $(IC_{P,\delta})$ and $(IC_{U,\phi})$. Hence, the problem becomes:

$$\max_{\{w_i^j\}_{i=\delta,\phi,j=P,L}} \lambda(y(e_\delta^P) - w_\delta^P) + (1-\lambda)(y(e_\phi^U) - w_\phi^U) \tag{2.6}$$

$$\text{s.t.} \quad \delta(w_\delta^P - \theta_P e_\delta^P) + (1-\delta)(w_\delta^U - \theta_U e_\delta^U) = 0 \tag{IR_δ}$$

$$\phi(w_\phi^P - \theta_P e_\phi^P) + (1-\phi)(w_\phi^U - \theta_U e_\phi^U) = \phi(w_\delta^P - \theta_P e_\delta^P) + (1-\phi)(w_\delta^U - \theta_U e_\delta^U) \tag{IC_ϕ}$$

$$w_\phi^P - w_\phi^U = \theta_U(e_\phi^P - e_\phi^U) \tag{$IC_{U,\phi}$}$$

$$w_\delta^P - w_\delta^U = \theta_P(e_\delta^P - e_\delta^U) \tag{$IC_{P,\delta}$}$$

$$w_\delta^P - w_\delta^U \leq \theta_U(e_\delta^P - e_\delta^U) \tag{$IC_{U,\delta}$}$$

$$w_\phi^P - w_\phi^U \geq \theta_P(e_\phi^P - e_\phi^U). \tag{$IC_{P,\phi}$}$$

Solving the binding constraints for $w_\delta^P, w_\delta^U, w_\phi^P, w_\phi^U$,

$$w_\delta^U = E_\delta(\theta)e_\delta^U, \tag{2.7}$$

$$w_\delta^P = E_\delta(\theta)e_\delta^U - \theta_P(e_\delta^U - e_\delta^P), \tag{2.8}$$

$$w_\phi^U = (E_\delta(\theta) - E_\phi(\theta))e_\delta^U + E_\phi(\theta)e_\phi^P + \theta_U(e_\phi^U - e_\phi^P), \tag{2.9}$$

$$w_\phi^P = (E_\delta(\theta) - E_\phi(\theta))e_\delta^U + E_\phi(\theta)e_\phi^P, \tag{2.10}$$

and substituting the relevant solutions in the maximisation, I obtain:

$$\max_{\{e_j^i\}_{j=\delta,\phi,i=P,L}} \lambda(y(e_\delta^P) - E_\delta(\theta)e_\delta^U + \theta_P(e_\delta^U - e_\delta^P)) +$$
$$(1-\lambda)(y(e_\phi^U) - (E_\delta(\theta) - E_\phi(\theta))e_\delta^U - E_\phi(\theta)e_\phi^P - \theta_U(e_\phi^U - e_\phi^P)). \tag{2.11}$$

From the maximisation problem, the chosen levels of effort e_δ^P and e_ϕ^U correspond to $y'(e^P_\delta) = \theta_P$ and $y'(e^U_\phi) = \theta_U$. As in section 2.4.1, all workers exert efficient effort levels.

Proposition 2.2 (Full Efficiency)**.** *When beliefs and productivity are perfectly correlated, full efficiency is always achieved regardless of the direction of the correlation. That is, both productive and unproductive workers choose first-best levels of effort.*

The result follows from the assumption of perfect correlation between beliefs and productivities.

If workers were naïve, but the correlation between beliefs and productivity were not to be perfect, then even after updating her beliefs, the employer would not be able to tell precisely the productivity type of a worker. Hence, she would assign positive probability to all possible combinations of belief and productivity types. I derive the equilibrium for this case in section 2.5.[7]

To derive the equilibrium values of imaginary offers in the case of perfect negative correlation, notice from (2.11) that the effect of e^U_δ is always negative and the effect of e^P_ϕ is always positive. Hence, $e^U_\delta = 0$ while $e^P_\phi = 1$.[8]

To see why the effort level for the optimistic productive type has positive effects on profits, simply notice that in this framework optimistic workers are always unproductive. Hence, the second channel of exploitation is more powerful than ever and the employer uses an imaginary offer that grants the largest possible incentive compatible surplus to an hypothetical optimistic productive type. Also note that, since the actual productive worker is always pessimistic, he assigns to this offer a much smaller weight than the optimistic productive worker when evaluating $w_\phi(e)$.

As for the effort level of the pessimistic unproductive type, the intuition is unchanged from section 2.4.1.

Given all the above, I can derive the equivalent of Result 2 for the case of perfectly negatively correlated types.

Result 3 (Exploitation with Perfect Negative Correlation)**.** *When beliefs and productivity are perfectly negatively correlated, pessimistic productive workers obtain zero surplus while optimistic unproductive ones are exploited. Exploitation is always strong.*

[7]Notice that, as explained later, if workers were not naïve at all but simply lacked the knowledge of their ability and formed unbiased expectations about it, the full efficiency result would still stand — as shown by Laffont and Martimort (2002) and Harris and Raviv (1979).

[8]This ensures that contracts are incentive compatibile.

Result 3 shows a peculiar feature of this case: pessimistic productive workers enjoy no positive rent.[9] Their pessimistic naïveté is large enough to allow the employer to extract all their surplus, but not large enough for them to be exploited. This is because in period 1 pessimistic productive workers play the role of the *low type*—they obtain zero expected utility. Hence, the employer can extract full surplus from them by offering a contract that ensures zero utility to both productivity types.

Optimistic unproductive types, on the other hand, can be screened away from $w_\delta(e)$ with the promise of a higher utility in the event of being productive and a lower one in the event of being unproductive. That is, introducing an imaginary offer that grants positive utility to productive types and negative utility to unproductive types.

In other words, while in section 2.4.1 (IC_ϕ) acts as a "proper constraint" on the exploitation level of the unproductive type, here it acts as a means of exploitation. It is through (IC_ϕ) that the employer can separate the unproductive type and take advantage of his naïveté.

To conclude this section I present a corollary to Results 2 and 3 that compares welfare findings.

Corollary 2.3 (Perfect Correlation Welfare)**.** *When beliefs and productivity are perfectly correlated, both productive and unproductive workers obtain lower utility when correlation is negative.*

Corollary 2.3 is quite intuitive. When correlation is negative, naïveté plays a much larger role and the productive worker loses all potential information rent. The employer uses the two channels of exploitation to their maximum effect.

To compare the findings with classical results, notice that if workers were not naïve but formed homogeneous unbiased expectations about their productivity (also known as the "selling of the firm" equilibrium, Laffont and Martimort, 2002; Harris and Raviv, 1979, where the worker becomes the residual claimant of the firm's profits), the employer would not be able to strongly exploit workers, but would still be able to achieve full efficiency. If instead workers were fully informed about their true productivity, the classical screening literature tells us that efficiency would be achieved only at the top and that productive workers would enjoy positive rents.

[9] Notice that pessimistic productive workers and optimistic unproductive ones are the only types present in the population.

Hence, an agent's imperfect information about his productivity allows the employer to exploit agents. Naïveté on its own allows her to use the second channel of exploitation at the cost of full efficiency. The perfect correlation between beliefs and productivity enables her to achieve full efficiency and, if the correlation is negative, to extract all of the surplus from the productive types while still strongly exploiting unproductive workers.

2.5 Imperfectly Informative Beliefs

In this section, I study the general case where beliefs are imperfectly informative. More precisely, both productive and unproductive workers have a positive probability of being either optimistic or pessimistic, i.e. $p_P, p_U \notin \{0,1\}$.[10]

As described in section 2.2, the basic intuition from the literature on contracting with naïve agents when there is no correlation between beliefs and productivity shows that the principal designs contracts that induce efficient effort in states — productivity types in my model — that agents deem less likely (Eliaz and Spiegler, 2006, 2008). Exploitation also takes place in these states. Hence, in my model, efficiency would be at the top in the contract for the pessimistic worker and at the bottom in the one for the optimistic worker.

With the introduction of my assumption regarding type-dependent naïveté, however, this intuition fails to hold for the entire parameter space. As I show in this section, the employer may find it optimal to induce a worker of i-belief and j-productivity type to exert the efficient level of effort not because of the misalignment of beliefs between her and the worker, but rather because of her posterior that a belief type i is indeed a j-productivity type.

When beliefs are imperfectly informative, the employer has a prior that assigns positive probability to each possible combination of beliefs and productivity type. Hence she solves problem (2.2). Period 2 incentive compatibility this time, however, is not as straightforward as before.

First of all, which incentive compatibility constraint is binding depends on the

[10]Notice that the case where the correlation between beliefs and productivity is perfect for only one productivity/belief pair is not analysed in the paper. Solutions for these cases, however, can be derived by combining the findings of this section with the ones of section 2.4. They present no further insights on the employer's optimal contracting behaviour.

posteriors of the employer since workers' have self-selected in period 1.

Second, notice that, as in classical screening problems, if $(IC_{j,i})$ binds, then the contract designed for an i-belief type induces the efficient level of effort in him. Hence, deriving conditions for the IC constraints binding in period 2 also indicates whether efficiency for belief type i is at the top or at the bottom. I start with the contract designed for optimistic workers.

Result 4 (Efficiency for Optimistic Workers)**.** *If the employer has a strong updated belief that optimistic workers are unproductive, or unproductive optimistic workers are naïve enough, efficiency is at the bottom in the contract for optimistic workers. That is, if*

$$\Pr\{\theta_P|\phi\} \le \tfrac{\phi}{1-\phi}\Pr\{\theta_U|\phi\}, \tag{2.12}$$

then $y'(e_\phi^U) = \theta_U$.

Condition (2.12) shows the trade-off the employer faces between inducing efficiency for the most probable productivity type and inducing it for the worker that has the most misaligned beliefs with respect to hers.

The employer has two main objectives: to induce efficient levels of effort in workers — maximising the pie — and to extract as much surplus as she can — taking the pie away from agents. When she has a strong belief that an optimistic worker is unproductive, she wants to induce him to exert efficient effort regardless of the extent to which she can exploit him, i.e. the level of his naïveté. On the other hand, even if the posterior on an optimistic worker being unproductive is not particularly high, i.e. most optimistic workers are productive, she may still want to induce efficient effort in unproductive optimistic workers. This happens when the latter are naïve enough — ϕ is large enough. As shown in section 2.4, in fact, naïveté increases exploitation.

A graphical intuition for Result 4 is represented in Figure 2.1. In the Figure constraints (IC_ϕ), $(IC_{U,\phi})$ and $(IC_{P,\phi})$ are plotted, together with isoprofits, in (w_ϕ^U, w_ϕ^P) space. Condition (2.12) holds in the graph to the left and fails in the graph to the right. Notice that profits increase towards the bottom left in each graph and that incentive compatible (both for period 1 and period 2) contracts lie in the area above (IC_ϕ) and between $(IC_{U,\phi})$ and $(IC_{P,\phi})$.

In the graph to the left, the posterior of the employer on an optimistic worker being unproductive is strong. Hence, an increase in w_ϕ^U bites more on profits than

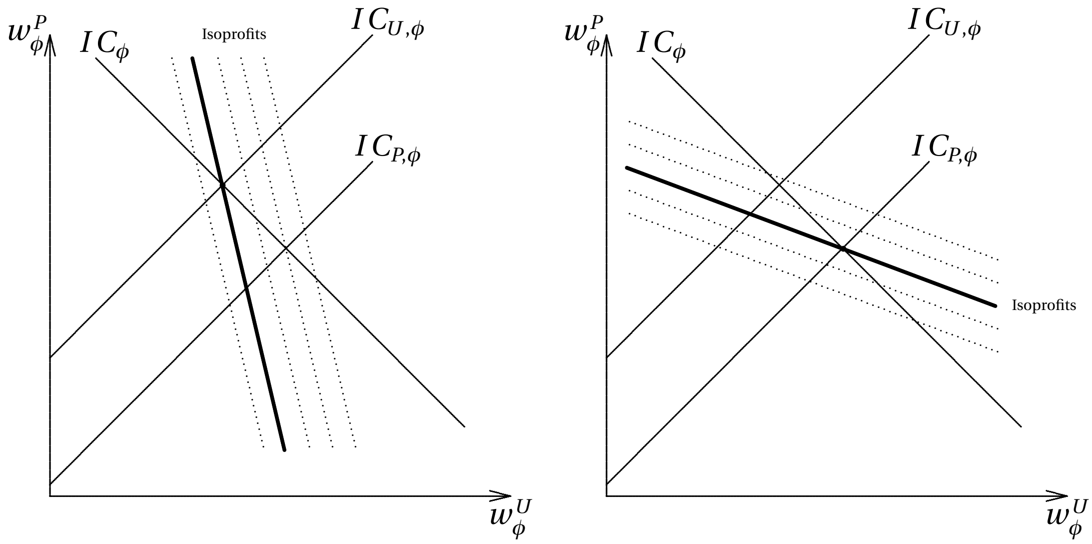

Figure 2.1: Optimistic Workers' Efficiency

In the Figure (IC_ϕ), $(IC_{U,\phi})$, $(IC_{P,\phi})$ and isoprofits are plotted in (w^U_ϕ, w^P_ϕ) space. When condition (2.12) holds — left side graph — isoprofits are steeper than (IC_ϕ). When condition (2.12) fails — right side graph — they are flatter than (IC_ϕ). Profits of the employer increase towards the bottom left of the graphs.

the same increase in w^P_ϕ. Hence, the isoprofits are steeper than the (IC_ϕ) constraints and efficiency is at the top in the contract for optimistic workers. In the graph to the right, the opposite intuition applies.

When beliefs and productivity were perfectly independent, then condition (2.12) would become $\frac{\lambda}{1-\lambda} \leq \frac{\phi}{1-\phi}$, which is true by assumption. In this situation, the employer would gain no information from screening in period 1 and would, therefore, focus on extracting surplus and inducing the most naïve workers in the population to exert efficient effort.

A similar result is true for pessimistic workers:

Result 5 (Efficiency for Pessimistic Workers)**.** *If the employer has a strong updated belief that pessimistic workers are productive, or productive pessimistic workers are naïve enough, efficiency is at the top in the contract for pessimistic worker. That is, if:*

$$\Pr\{\theta_U|\delta\} \leq \tfrac{1-\delta}{\delta} \Pr\{\theta_P|\delta\}, \tag{2.13}$$

then $y'(e^P_\delta) = \theta_P$.

Condition (2.13) is the mirror image of (2.12) for pessimistic workers. Notice that the naïveté of pessimistic productive workers is measured by $\frac{1-\delta}{\delta}$ which is decreasing in δ. The lower δ the larger the naïveté of productive pessimistic workers. Figure 2.2 below shows a similar graphical intuition to the one of Figure 2.1.

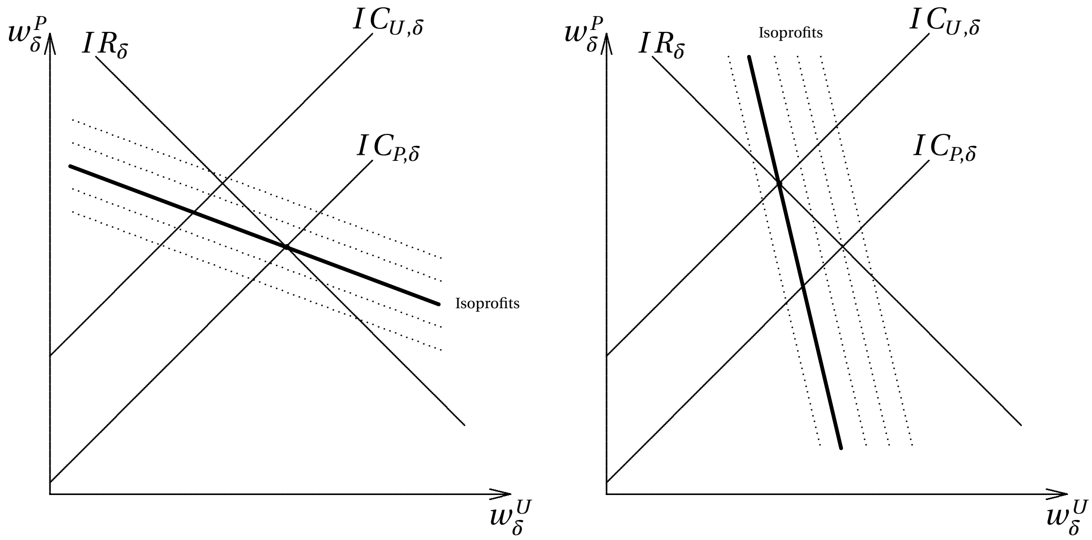

Figure 2.2: Pessimistic Workers' Efficinecy

> In the Figure (IR_δ), $(IC_{U,\delta})$, $(IC_{P,\delta})$ and isoprofits are plotted in (w_δ^U, w_δ^P) space. When condition (2.13) holds — left side graph — isoprofits are flatter than (IR_δ) and efficiency is at the top in $w_\delta(e)$. When condition (2.13) fails — right side graph — they are steeper than (IR_δ) and efficiency is at the bottom. Profits of the employer increase towards the bottom left of the graphs.

Ultimately, conditions (2.12) and (2.13) define the equilibrium of the model. Together they determine the optimal behaviour of the employer and identify which of the two competing effects (exploiting workers' naïveté vs. inducing efficiency in the most common productivity type) dominates. I present this result in Proposition 2.4 and then analyse it in the (p_U, p_P) space for a given δ, λ and ϕ.

Proposition 2.4 (Imperfect Correlation Efficiency)**.** *When both conditions (2.12) and (2.13) hold, efficiency is at the bottom (top) in the contract for optimistic (pessimistic) workers. When both conditions fail, efficiency is at the top (bottom) in the contract for optimistic (pessimistic) workers. If condition (2.12) holds while (2.13) fails, efficiency is at the bottom in both contracts. If condition (2.12) fails while (2.13) holds, efficiency is at the top in both contracts.*

Proposition 2.4 describes the main trade off faced by the employer and states the main contribution of the paper. It follows from combining Results 4 and 5. To understand the Proposition, consider Figure 2.3.

The basic parameters of the model are $\delta, \lambda, \phi, p_P$ and p_U. The first three simply describe the relation between optimistic and pessimistic workers and the proportion of productive types in the population. The last two, instead, are the focus of the paper and determine the level of information about true productivity obtained by knowing a worker's beliefs — i.e. the extent of naïveté type-dependence. In Figure 2.3, I assume $\delta = \frac{1}{3}, \lambda = \frac{1}{2}$ and $\phi = \frac{2}{3}$ and study the equilibrium of the model in (p_U, p_P) space.

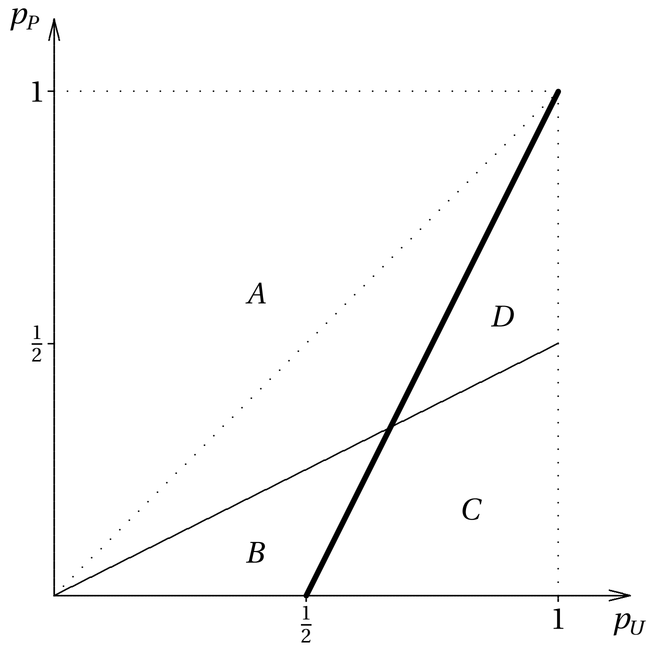

Figure 2.3: Efficiency in Optimal Contracting

In the Figure condition (2.12) and (2.13) are plotted in (p_U, p_P) space. To the left of the bold line, condition (2.12) holds and the contract for optimistic workers features efficiency at the bottom. To the right of it the condition fails and the contract features efficiency at the top, instead. Above the thin line, condition (2.13) holds and the contract for pessimistic workers features efficiency at the top. Below it, the condition fails and the contract features efficiency at the bottom, instead.

The 45° line in the graph separates the area of positive correlation, below the line, and the one of negative correlation, above the line.

In area A the optimal contracts feature efficiency at the top for the pessimistic worker and at the bottom for the optimistic one. Corollary 2.5 shows that this area occupies the entire portion of the parameter space where beliefs and productivity are negatively correlated for every $\delta < \lambda < \phi$. This is perfectly in line with the literature's findings, discussed in section 2.2, and agrees with the intuition that employers induce workers with strong wrongly biased beliefs to exert efficient effort, while distorting offers for more self-aware types. Notice that if the type dimensions were independent, p_P and p_U would be equal. Hence, the 45° line would be the parameter space and area A would characterise all the equilibria. The portion of the parameter space on the 45° line represents the model of Eliaz and Spiegler (2008) with only two belief types of agent.

Corollary 2.5 (Efficiency with Negative Correlation)**.** *When beliefs and productivity are negatively correlated, the contract for pessimistic workers features efficiency at the top while the one for optimistic workers features efficiency at the bottom.*

The Corollary is proven by studying conditions (2.12) and (2.13) when $p_P > p_U$. When beliefs and productivity are negatively correlated, the cost of extracting efficient levels of effort from optimistic unproductive workers and pessimistic productive workers is low. It is so low, in fact, that even when chances to meet such types are low (according to her updated beliefs) the employer still finds it optimal to extract surplus from these types.[11]

The area below the 45° line is divided in four: a portion of area A, area B, where efficiency is at the bottom in both contracts, area C where efficiency is at the top for optimistic and at the bottom for pessimistic workers, and area D where efficiency is at the top for both belief types. This shows the interaction of the two driving forces of the employer's behaviour. On the one hand, when workers are naïve enough, the employer wants to take advantage of their wrong beliefs to exploit the unproductive type without giving up a too large surplus to the productive worker. On the other hand, when workers' naïveté decreases and beliefs become more informative, the employer behaves according to her posteriors, inducing efficiency for the productivity types she deems most probable.

[11] Notice that negative correlation between belief and probability happens in the area of the graph above the 45° line. Hence, however small, there is a strictly positive probability that these types exist.

In particular, in area B, δ has not changed from area A. What has changed, however, is the *expected* naïveté of a pessimistic worker since the probability of meeting a pessimistic worker who is productive is relatively low. When p_P is small the number of pessimistic workers that turn out to be productive is low, hence after the screening of period 1, the employer updates her priors and induces pessimistic unproductive workers to exert efficient effort rather than the pessimistic productive ones. In other words, the chances that a pessimistic worker is productive are so small that the benefits of extracting efficient effort from such a type are negligible.

Area D has the exact opposite intuition. The employer's posterior on facing a unproductive worker given that the latter is optimistic are very small. Hence, she designs a contract with efficiency at the top for optimistic workers.

Finally, area C represents the case where workers have beliefs that are strongly positively correlated with their productivity. Hence, the optimal contracts resemble, in efficiency terms, the ones of section 2.4.1.

In Appendix A.1, I derive the optimal contracts for all the possible cases described. Given the solutions found, the next Result studies the case of bunching of pessimistic types, while the next Corollary studies workers' welfare.

First of all, although in the case of imperfectly informative beliefs there are no imaginary offers, the offers set for workers whose period 2 IC are not binding play a similar role. If the employer wants to exploit the pessimistic unproductive type, for example, through offer $(w_{\delta}^{P}, e_{\delta}^{P})$, she is still capable of doing so. This time, however, offer $(w_{\delta}^{P}, e_{\delta}^{P})$ has a first order direct effect on profits. This is because pessimistic productive workers *do* exist and the employer's priors on a worker being such a type are given by $p_P\lambda$. Hence, designing a contract that screens among differently productive pessimistic workers in period 2 may be suboptimal. Differently from section 2.4.1, this is *regardless* of the direction of the correlation between beliefs and productivities.

The higher the naïveté of pessimistic productive workers, relative to that of optimistic productive workers, the higher is the gain of using $(w_{\delta}^{P}, e_{\delta}^{P})$ for exploitation. When the proportion of optimistic workers is high, however, period 1 separation becomes too costly — in terms of higher expected utility granted to optimistic workers.

Result 6 (Pooling of Pessimistic Workers)**.** *When beliefs and productivity are imperfectly correlated, if the pessimistic productive worker is naïve enough — relative to* *the optimistic unproductive one — or if the proportion of optimistic workers is small, that is:*

$$\frac{\delta}{\phi} \geq \Pr\{\phi\}, \tag{2.14}$$

then the employer separates pessimistic workers on the basis of their productivity. Otherwise, they are bunched together.

Result 6 is reminiscent of Result 1. When the expected utility of $w_\delta(e)$ increases — as a consequence of a rise in the utility granted by (w_δ^P, e_δ^P) — in order to separate the optimistic worker from a pessimistic one, the employer has to increase the expected utility coming from $w_\phi(e)$. This is regardless of the actual productivity of the optimistic worker. Furthermore, since the optimistic worker weights (w_δ^P, e_δ^P) more than a pessimistic one, the increase in expected utility granted to optimistic workers can offset the profit gains of using (w_δ^P, e_δ^P) to exploit the pessimistic unproductive worker. This happens when condition (2.14) fails.

The next Corollary shows that the qualitative results on workers' welfare are common to all four areas.

Corollary 2.6 (Imperfect Correlation Welfare)**.** *When beliefs are imperfectly informative about workers productivity, unproductive workers are always exploited while productive workers enjoy a non-negative surplus.*

Corollary 2.6 follows from the proof of Lemma 2.1. Hence, the result is qualitatively unaffected by the direction and extent of the correlation between beliefs and productivity. Qualitative welfare results are unchanged regardless of the information granted by beliefs on productivity levels.

Since in period 1 the employer has only priors over agents' belief-types, she cannot take advantage of the correlation between type dimensions. In period 2, however, she can update her beliefs and set offers that affect the extent of the exploitation of unproductive workers and the amount of surplus granted to productive ones.

2.6 Concluding Remarks

The purpose of this paper is to study the optimal contracting behaviour of a principal who faces agents with type-dependent nav̈ieté. There are two main implications of this new assumption.

First, when workers' beliefs are perfectly informative of their productivity, full efficiency is achieved. Unproductive workers are always exploited. Productive workers enjoy a positive surplus if their beliefs are positively correlated with their productivity. If the two are negatively correlated, instead, they enjoy no information rent.

The second main result focuses on efficiency of contracts when beliefs are imperfectly informative of their productivity. I show how the employer uses the information gained by screening among different belief types in period 1. She designs contracts that extract efficient effort levels either from the most naïve workers in the population, because it is "cheaper" to do so, or from the type of worker she deems most probable to face given her posteriors.

These findings connect the literature on sequential screening (Courty and Li, 2000, *inter alia*) with the one on contracting with naïve agents (Eliaz and Spiegler, 2006, 2008) by assuming that agents' naïveté depends on the agents' true nature (as in sequential screening problems).

Eliaz and Spiegler (2008) define a *speculative* contract as one that grants an i-belief worker expected utility $\lambda(w_i^P - \theta_P e_i^P) + (1-\lambda)(w_i^U - \theta_U e_i^U) < 0$. In other words, a contract is speculative if it should not be signed by a worker with unbiased beliefs. In the context of this paper, I can prove that an optimistic (pessimistic) worker never (always) signs a speculative contract. This, however, does not save (condemn) him from (to) obtaining zero (negative) surplus.

This result follows from the assumption that one of the period 2 "states of the world", i.e., the two levels of productivity, dominates the other. In other words, for any level of effort e', the utility a productive worker obtains from $(w(e'), e')$ is always higher than the utility obtained by an unproductive worker. This assumption is not present in Eliaz and Spiegler (2006, 2008). The study of what would happen in a framework of type-dependent naïveté with unordered period 2 states is left for future work.

The extension of the present model with a continuum of belief types, and the assumption of heterogeneous distributions of beliefs among equally productive workers are work in progress. Also, of interest for future research is the relaxation of the assumption on the perfect observability of e, introducing of a moral hazard problem in the model.

Chapter 3

Price Discrimination in the Retailing Industry

Chapter Abstract

I study the causes, characteristics and consequences of loyalty schemes in a market with consumers who suffer from self-control problems (Gul and Pesendorfer, 2001). While the literature has mostly focused on loyalty schemes as tools used by firms to compete (Caminal and Claici, 2007) or increase consumers' lifetime value (Caminal, 2012), I look at how a seller can use them to acquire information about consumers' preferences and gain the ability to price discriminate when she has no control over quantity. I show how the more precise is the information acquired, the more the equilibrium converges to first-best, while the effect on consumers' welfare is non-monotonic. Finally, I derive necessary and sufficient conditions for all consumers to optimally disclose their type. These coincide with the conditions for first best to be fully restored.

3.1 Introduction

In this paper, I study how the presence of asymmetric information and consumers who suffer from self-control problems in a market can induce the sellers to implement a loyalty card scheme. These schemes have two main functions: they (i) grant sellers information used to discriminate between consumers and (ii) work as commitment devices for the latter to counter their self-control problems. In the model I assume a continuum of consumers that differ in their level of temptation.

The majority of large retailers offers a loyalty card scheme to their consumers in exchange for some personal information. A good example of such technologies are the loyalty card systems used by large retailers, like Tesco's *clubcard* or Sainsbury's *nectar*.[1] Accepting the loyalty card scheme, the consumer reveals personal information (age, sex, job etc) that help the retailer better understand his preferences. Moreover, to fully enjoy the benefits of his loyalty card, the consumer has to use it (i.e. "scan it" at the till) every time he makes his purchases at the retailer's store. This creates a large amount of data that describes the consumer's purchasing behaviour and can increase even further the ability of the retailer to identify his preferences. Observing this information, she updates her beliefs about consumers' type and sends out personal "discount coupons" to them. Thus, the retailer is capable of price discriminating between consumers who subscribe to the loyalty card scheme. This assumption has been investigated by the empirical literature on "individual pricing" discussed below (Shiller and Waldfogel, 2011; Mikians, Gyarmati, Erramilli, and Laoutaris, 2012, 2013; Hannak, Soeller, Lazer, Mislove, and Wilson, 2014; Waldfogel, 2015; Shiller, 2015).

In general, consumers preferences are heterogeneous. In this paper, I study the case of consumers with self-control problems who differ in their level of temptation. I analyse their behaviour when facing the decision to subscribe or not to a loyalty card scheme. Differently from one without self-control problems, a consumer who suffers from high temptation (and therefore values the good more) may avoid shopping in a store that sells a particularly tempting good. For example, he may be tempted by a good that is too expensive to be bought, yet cheap enough to require self-control effort.

[1] The model can be applied to any market where a retailer offers consumers the opportunity to "register", like Ryanair and Amazon.com.

To see why temptation plays an important role in the kind of markets I study, one can think about the obvious connection between food and self-control. In 2014 the *Financial Times* reported the decision of UK's largest grocer, Tesco, to ban sweets and chocolates from checkout tills in all stores. In 2013 the UK's Department of Health asked grocers to stop inducing "impulses purchases" in their stores.[2] Consumers in the markets I consider are tempted, and retailers take this temptation into account in their selling behaviour.

The model can also be applied to online retailers that are not (necessarily) linked to the grocery industry. Although the connection with these kinds of markets may be less clear, temptation is still an important feature that affects retailers' price-setting behaviour. *Wishlists*, for example, are lists of products that the (registered) consumer can save in his own online profile. Every time he logs onto the retailer's webpage he has his wishlist in front of him and the purchase of all his *wished* products "one click away". This practice has been used by online retailers to tempt consumers with a purchase that they decided to postpone at first. It has the purpose of reminding them of the existence of the product and of making it part of every decision they are making.

Although retailers may find it profitable to tempt consumers, in this paper I show how consumers' self-control problems may also hurt a seller's profit driving them away from her store when she sells tempting goods. The solution to this problem is one of the uses loyalty cards have in this paper.

By subscribing to a loyalty card scheme, consumers grant the seller information about their temptation level (which describes their willingness to pay). This information, however, can be more or less precise depending on the quality of the loyalty scheme technology. On one end of the spectrum, is the case where by accepting the loyalty card, the consumer reveals to the seller only that he is willing to do so. In other words, the seller observes no further information about the consumer's temptation level at all. On the other end, instead, is the case where, by accepting the loyalty card, the consumer fully disclose his level of temptation for the seller to observe. In this paper, I study these two limit cases and all relevant ones in the middle.

Given the information she observes, the seller, in turn, can grant consumers spe-

[2] "*Tesco to ban sweet temptations at tills*", Duncan Robinson, *Financial Times online*, May 22nd, 2014.

cific discounts on certain products in order to make them willing to enter her superstore. This may be achieved, for example, by making tempting goods less expensive, so that it becomes optimal to buy them, or by decreasing the price of non-tempting goods that the consumer buys, so as to compensate him for his self-control cost. A model with "classical" consumers would ignore these features, as shown in the Appendix.

To understand the kind of environment the model is motivated by, consider, for example, the case of a seller that owns a large store (a "superstore") in a specific region. In this region, the superstore is the only place where consumers can buy the latest, more expensive, version of a technological product. All other stores in the area sell the relatively old version which is now well supplied by the market. In this and many other possible examples, the seller enjoys a monopolistic position on a product that has imperfect substitutes sold in the market (as well as in the superstore). In particular, consider the case of this product being a tempting one. A product that makes consumers exert self-control cost is on sale in the store they usually go to do their shopping.

In this framework, I address the following main questions: are loyalty cards a way for a multi-product seller to attract consumers when temptation levels are private information? Can loyalty cards be used to discriminate between different types of consumers when the seller can only control the price? Can this discrimination help consumers deal with their temptation? The answer to all these questions is yes. The results of the paper show how loyalty card schemes induce consumers to *voluntarily* disclose information about their preferences and help the seller achieve the first best outcome. They show how consumers may reject the loyalty card to commit themselves to face a high price for the tempting good, hence, decreasing their self-control cost of shopping in a store. This, can potentially increase overall consumer welfare in the market as I show in Section 3.6.6.

This paper contributes to different branches of the economics literature, analysed in detail in Section 3.8. First is the literature on temptation and self-control (Gul and Pesendorfer, 2001) in contract theory. Second is the rather broad and multi-disciplinary literature on loyalty card schemes. Third, to my knowledge, I provide one of the first theoretical models that approaches the issue of "individual pricing" with consumer tracking. Finally, my findings have aspects in common with the costly state verification literature started by Townsend (1979). The main difference between this work and other applications of Gul and Pesendorfer (2001) is that here the seller is thought of as a retailer, able to alter the price of the good but neither its quality, because she does not technically produce the good, nor its quantity, because consumers are interested in purchasing one unit of the good only. In the applied literature that I analyse later on (Esteban and Miyagawa, 2005, 2006; Esteban, Miyagawa, and Schum, 2007; Foschi, 2014), instead, the seller has two control variables, the price and the quantity/quality of the good, hence she is able to create incentive compatible menus of offers that may not require the use of loyalty cards.

The theoretical literature on loyalty cards takes Competition Policy and Industrial Organisation perspectives to study loyalty cards as business-stealing tools (Caminal and Claici, 2007), as bundled loyalty discounts (Greenlee, Reitman, and Sibley, 2008), as "bribes" to agents that buy products with principal's money (Basso, Clements, and Ross, 2009), as collusion tools under the form of second period discount coupons (Ackermann, 2010) and in dynamic environments as tools to increase consumers' participation and expenditure (Chen and Pearcy, 2010; Caminal, 2012). The purpose of this paper is to take a different perspective, filling the gap between loyalty card schemes, asymmetric information and self-control. In particular, I focus on loyalty cards as a mean of *exchange* to acquire information about consumers' preferences.

The assumption that sellers use large data sets ("Big Data", previously unavailable and only recently become tractable by modern computers,) to study consumers' behaviour is referred to, in the literature, as individual pricing with consumer tracking. Empirical papers (Mikians, Gyarmati, Erramilli, and Laoutaris, 2012, 2013; Hannak, Soeller, Lazer, Mislove, and Wilson, 2014) have found evidence of individual pricing in online markets like Amazon and Netflix. Shiller and Waldfogel (2011) and Waldfogel (2015) consider individual pricing in the market for music and professional higher education, but not via consumers' tracking. Shiller (2015) shows through a simulation study that personalized pricing using individual customers' web browsing data can increase profit by a substantial amount, while decreasing consumers surplus. The issue of consumers' tracking has also attracted the attention of media and antitrust institutions in the recent years.[3] In particular, the focus

[3] *"Websites Vary Prices, Deals Based on Users' Information"*, Jennifer Valentino-Devries et al., *The Wall Street Journal Online*, Dec 24^{th}, 2012.
"How Advertisers Can Use Your Personal Information To Make You Pay Higher Prices", Tarun Wadhwa., *The Huffington Post*, Apr 2^{nd}, 2014.

of the discussion has been on why are sellers allowed to do that and if they should be prevented from using this practice.

The paper is organised as follows. In Section 3.2, I describe the nature of consumers' preferences and temptation. In Section 3.3, I introduce the role of the sellers. Section 3.4 describes the timing of the game and the order of play. In Section 3.5, I describe the benchmark, full information case and the private information equilibrium. In Section 3.6, I introduce the technology of loyalty card schemes. In Section 3.7, I discuss my findings. I analyse the relevant existing literature in Section 3.8. Finally, I conclude the paper in Section 3.9. Proofs of Lemmas, Results and Propositions are all relegated to an Appendix. Also in the Appendix is the solution for consumers without self-control problems.

3.2 Consumers' Preferences

I consider a continuum of consumers that differ in their level of temptation. They are interested in buying only one unit of one of the goods in the market. Representative consumer i (he) suffers from self-control problems and has preferences à la Gul and Pesendorfer (2001). He, therefore, distinguishes between an *offer* (a, p_a), where p_a is the price he pays to purchase a unit of good a, and a *menu* M, that is a set of offers among which he can choose one to purchase. Following Gul and Pesendorfer (2001), tempted consumers evaluate a menu with their *self-control preferences*, represented by utility function $W_i(M)$, and an offer with the sum of their *commitment and temptation utilities*, $U(a, p_a) + V_i(a, p_a)$.

At the moment of purchase, when choosing an offer from a given menu M, consumer i considers his preferences for commitment and temptation and decides which offer to buy according to:

$$\max_{a \in M}[U(a, p_a) + V_i(a, p_a)] \tag{3.1}$$

Function U is called the (net) *commitment utility* while function V is called the (net) *temptation utility*. To understand the difference between these two functions, consider U as the base utility that the individual obtains from the good, free of temptation. Function V, instead, measures the impulses of the individual at the moment of purchase. When given a menu, he considers both his commitment and his temptation and takes his decision.

Choosing from a set of different menus $\{M_1, M_2, M_3, ..., M_N, \emptyset\}$ (where $\emptyset$ represents the rejection of all menus, and $W(\emptyset)$ is normalised to 0), instead, consumer i selects the one that maximises his self-control preferences:

$$W_i(M_j) = \max_{a \in M_j}[U(a, p_a) + V_i(a, p_a)] - \max_{a \in M_j} V_i(a, p_a) \quad j = 1, ...N. \tag{3.2}$$

This is the utility the consumer enjoys facing menu M_j. It is composed of the utility he obtains by making the purchase minus the temptation utility that he is foregoing because he is exerting self-control effort, represented by the offer that would maximise his temptation utility, $\max_{a \in M_j} V_i(a, p_a)$. To understand the intuition behind (4.4), notice that if offer γ maximises $U + V$ and ω maximises V, then $W(\cdot) = U(\gamma) - [V(\omega) - V(\gamma)]$ where $V(\omega) - V(\gamma)$ is known as the *self-control cost* of choosing γ over ω. If the latter is too high, the consumer will not accept the menu in the first place.

In this paper, I assume quasi-linearity of utility functions:

$$U(a, p_a) = u(a) - p_a \tag{3.3}$$

$$V_i(a, p_a) = \phi_i v(a) - p_a. \tag{3.4}$$

A good a is *tempting,* independently of consumers' type, if it grants more temptation than commitment (gross) utility, $v(a) > u(a)$. Heterogeneity of preferences is captured by parameter ϕ_i, which is distributed uniformly in $[0, 1]$. It represents the level of temptation of consumer i and, in particular, it captures the idea that consumers value the tempting features of a in different ways. Hence, it defines a consumer's type.[4]

In my model, I say that $(a, p_a) \succ_T (b, p_b)$ for i, i.e., offer for good a is *more tempting* than offer for good b, if and only if $V_i(a, p_a) > V_i(b, p_b)$.

[4]Notice that, following the existing literature on temptation models, I assume the negative part of utility the consumer gets from paying tariff t is equal to the actual transfer t itself. This is true also for the temptation utility. This results in an ex-post utility $U(x) + V_i(x) = u(q) + v_i(q) - 2t$ even if the transfer is made only once. This t cancels out in (4.4) above, if $\arg\max[U(x) + V_i(x)] = \arg\max V_i(x)$, since $W = u(q) + v_i(q) - 2t - \left(v_i(q) - t\right) = u(q) - t$. The peculiarity of this approach is that, if $\arg\max[U(x) + V_i(x)] \neq \arg\max V_i(x)$, the price of an offer which is not chosen enters the utility function as well.

3.3 Firms

A seller (she) sells two imperfect substitutes, x and z, to the consumers. Good x is also sold by other sellers in a monopolistically competitive market (also called *competitive fringe*). Its price p_x is fixed by the market competition and cannot be altered. Good z, instead, is sold only by the first seller, in her *superstore*, who then enjoys monopoly power (hence, I will also refer to her as "the monopolist"). The costs of z and x are, respectively, c_z and c_x. Throughout the paper I assume for simplicity that $c_z \leq c_x$. The idea behind this assumption is that since good z is sold only by the monopolist it can be thought of as an innovative product with a lower production cost. On the other hand, one could argue that new, innovative products are more costly to produce. This is analysed in an Appendix.[5]

The profit of the monopolist when a consumer buys good $a = \{x, z\}$ is given by:

$$\pi(a, p_a) = p_a - c_a. \tag{3.5}$$

Let A be the subset of consumers buying good a. Then, total profits are given by:

$$\Pi(a, p_a) = \int_A (p_a - c_a) d\phi. \tag{3.6}$$

I show below how my regulatory assumptions imply that every type has a willingness to pay for z at least equal to p_x. Hence, the monopolist will sell z at $p_z \geq p_x$.

All other sellers are "*passive*" players and they have no choice variable in the game.

3.4 Order of Play

The most relevant connection between Gul and Pesendorfer (2001) and the markets I study is the timing of consumers' decision. In Gul and Pesendorfer (2001), tempted consumers are implicitly assumed to take two different decisions at two different times. Formally these decisions are assumed to be taken in two different periods: period 1, the *ex-ante stage*, and period 2, the *ex post stage*. In the first, consumers

[5]Assuming $c_z > c_x$ would affect the results qualitatively in terms of the set of consumers screened out of the market in equilibrium, when information is private. This, however, makes the solution computationally heavier without affecting the results in an interesting way.

decide whether to accept one of the menus they have access to, or none. In the second, they choose an offer from the menu they have chosen ex-ante, if any. Since by *menu* here is intended the entirety of the goods sold by a seller, *accepting a menu* means to be willing to "*enter the store*" of a specific seller to choose among the goods on sale. I will use this terminology henceforth.

Notice that a two-period decision represents well the environments that I study here. Consumers first decide whether to enter or not into the retailer's store (or her website) and then decide how to make their purchases from the "menu" they have in front of them. In the case of the loyalty cards, as I explain in Section 3.6, an extra consumers' decision is introduced before all others, the one of accepting or not the retailer's loyalty card (or *register* online on her website). The outcome of this decision affects the menus faced by the consumers.

In this model, the monopolist moves first at the start of the ex-ante stage and sets price p_z, defining in this way menu $M_m(p_z)$ that contains both x and z at prices p_x and p_z. Other sellers all offer menu M_s with only x offered at price p_x. As in the general literature on self-control and temptation models, I assume that a consumer is always free to enter a store and leave having bought nothing. I represent this by defining the *null offer* 0. If a consumer chooses the null offer from a menu, he obtains 0 and pays 0. The menus at his disposal at the start of the game are, therefore:

$$M_m(p_z) = \{(x, p_x), (z, p_z), 0\}, \quad M_s = \{(x, p_x), 0\} \tag{3.7}$$

In the ex-ante stage, consumer i is outside the stores and has to decide whether to (i) enter the monopolist's superstore, (ii) enter one of the other retailer's smaller stores, or (iii) simply walk past the stores. In taking this decision he uses his self-control preferences, represented by the *ex-ante utilities* $W_i(M_m(p_z))$ and $W_i(M_s)$, from (4.4), and the ex-ante utility of entering no store, $W_i(\emptyset) = 0$.

In the ex-post stage, provided he entered a store ex-ante, consumer i decides which offer to pick, and hence which good to buy. He takes this decision according to (4.1), henceforth called the *ex-post utility*.

Two regulatory assumptions are made. The first is about consumers' preferences and the attractiveness of menu M_s and good x, the second is about products' base utility levels.

Assumption 1. *Every consumer is always willing to buy x as opposed to buying noth-*

ing, that is $U(x, p_x) + V_i(x, p_x) \geq 0$*, for all* ϕ_i.

Assumption 1 affects the outside option of the consumer when deciding whether or not to accept the monopolist's menu, and when choosing the offer from the menu. Notice that it implies that $p_x \leq \frac{u(x)}{2}$. That is the price of good x is lower than the half of the commitment utility obtained by good x. It also states that even the non tempted consumer ($\phi_i = 0$) finds it optimal to buy good x. I show in the Appendix how this assumption also implies that every consumer has the incentive to enter one of the small stores that offers menu M_s, as opposed to walking past the stores. This creates the interesting competing aspect between the monopolist and the competitive fringe.

Assumption 2. $v(z) > v(x) \geq u(x) = u(z) \geq u(0) = 0$.

Assumption 2 ranks the gross utility levels obtained by good x and z. It describes the exogenous nature of products and it is independent of consumers' type. The first inequality is obvious since I define z to be the most tempting good in the model. The second is reasonable if one recalls that the goods are substitutes. Hence, it would seem implausible to think that x features no tempting aspects at all. The last inequality and equality are just simplifying assumptions. What is really crucial is the assumption that the two goods yield the same commitment utility. This assumption implies that each consumer values z at least as much as x. Also, it implies that good z is tempting in a *good* way, since it does not yield a smaller commitment utility than x. In many applications, in fact, u is considered as the utility obtained by "sticking to a plan" — a diet for example. The tempting good usually decreases the commitment utility as it "ruins the plan" — a highly calorific burger. Here there is no "plan", z is harmless and simply features some extra perks which are more or less valuable and tempting to different consumers.

Given consumers' preferences, if the monopolist wants to sell good z to consumer i in period 2, she has to set p_z such that two constraints hold. The first is an ex-ante participation constraints that ensures that consumer i enters the monopolist's superstore ex-ante, as opposed to entering one of the smaller stores (his outside option). The second is an ex-post incentive compatibility constraint that ensures that consumer i buys z and not x once inside the store:

$$W_i(M_m(p_z)) \geq W_i(M_s) \tag{PC}$$

$$U(z, p_z) + V_i(z, p_z) \geq U(x, p_x) + V_i(x, p_x). \tag{IC}$$

In the following Section, I use these two constraints and present the results of the full and asymmetric information case.

3.5 Optimal Pricing Scheme and Equilibrium

In the full information case the seller sets a value of p_z for each consumer, given his ϕ_i. When information becomes private, she relies on her prior belief ($\phi_i \sim U[0,1]$) and decides whether to screen out some types or not.

The seller wants to sell good z to a consumer with temptation ϕ_i if $\pi(z, p_z) \geq \pi(x, p_x)$, where p_z is the solution to:

$$\max_{p_z}[p_z - c_z] \tag{3.8}$$

$$\text{s.t.} \quad W_i(M_m(p_z)) \geq W_i(M_s) \tag{PC}$$

$$U(z, p_z) + V_i(z, p_z) \geq U(x, p_x) + V_i(x, p_x). \tag{IC}$$

The solution to (3.8), depends on whether ϕ_i is commonly known or not.

3.5.1 Full Information (First-Best).

If temptation levels are commonly known, an optimal pricing scheme p_z^* solves (3.8) for every ϕ_i.

Although I assume that consumers are always happy to buy good x, this does not imply that the offer for good x is more tempting than the null offer. The problem then divides in two, one for consumers that are tempted by good x, i.e, $(x, p_x) \succ_T 0$, and one for those who are not, i.e, $0 \succ_T (x, p_x)$. Both cases are solved in the proof of Result 7. The following Result describes the equilibrium under full information.

Result 7. *If temptation levels are commonly known, the optimal pricing scheme for* *z is:*

$$p_z^*(\phi_i) = \begin{cases} \frac{1}{2}\phi_i\left(v(z) - v(x)\right) + p_x & if \phi_i < \phi^* \\ 2p_x - \phi_i v(x) & if \phi_i \in \left[\phi^*, \frac{p_x}{v(x)}\right] \\ p_x & if \phi_i > \frac{p_x}{v(x)} \end{cases} \tag{3.9}$$

$$where \quad \phi^* \equiv \frac{2p_x}{v(z)+v(x)} \tag{3.10}$$

is the type of consumer with the highest willingness to pay. The monopolist is willing to sell z to every type of consumer, since $\pi\left(z, p_z^*(\phi_i)\right) \geq \pi(x, p_x)$ for all ϕ_i.[6] *Total profits are:*

$$\Pi\left(z, p_z^*(\phi_i)\right) = \int_0^1 \left(p_z^*(\phi_i) - c_z\right) d\phi \tag{3.11}$$

Figure 3.1 shows the optimal level of p_z^* for each value of ϕ_i — i.e. the ex-ante willingness to pay of each consumer in the market.[7] Low types do not value the tempting aspects of z enough for the monopolist to charge them a high price. As ϕ_i increases, their valuation of the good increases, the IC becomes slack and the monopolist can raise p_z to make the constraint bind again. This is true for all types below ϕ^*. I define these consumers as *weakly tempted.* Consumers beyond ϕ^*, instead, are so tempted by z that its purchase becomes ex-ante sub-optimal. That is, their ex-ante willingness to pay for good z is decreasing in their temptation level, ϕ_i. I define these consumers as *strongly tempted.* Finally, notice that among these types there are also the most tempted consumers in the market, $\phi_i > \frac{p_x}{v(x)}$. Ex-ante, they ignore completely the tempting features of good z in order to resist to it. That is, by exerting self-control, they value good z exactly as good x and their willingness to pay is constant and equal to p_x.

3.5.2 Private Information.

When information about the temptation level is private, the monopolist believes that ϕ_i is uniformly distributed in the interval $[0,1]$ and hence, she faces a screening problem. Unlike classical second-degree price discrimination, however, she does not control the quantity of the good and cannot, therefore, implement a classical

[6] This changes when $c_z > c_x$ since $(p_x - c_z) < (p_x - c_x)$ and it is not profitable anymore to sell good z to some consumers in the market.

[7] In the Figure, for simplicity I assume that $v(z) = 3v(x)$ so that the absolute value of the slope is the same before and after ϕ^*. This assumption has the only purpose of making the graph pleasant to the eye. I assume this in all the Figures of the paper.

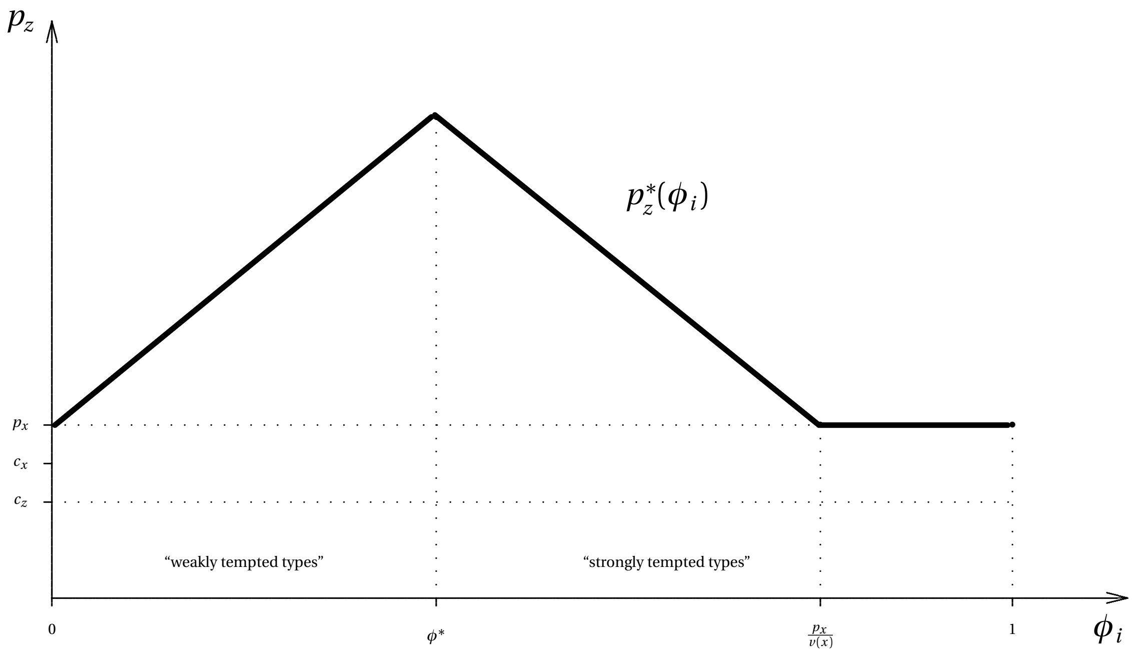

Figure 3.1: First Best Pricing and Willingness to Pay

The thick curve in the Figure shows the willingness to pay, $p^*_z(\phi)$, for each consumer when the seller is able to perfectly observe ϕ_i. I will refer to this curve as the "willingness to pay curve".

truth-telling mechanism. She can choose one of the following: (i) charge a price $\underline{p}_z$ that attracts all consumers in her store, (ii) extract all the surplus of the consumer with the highest willingness to pay charging the highest possible price, $\bar{p}_z$, or (iii) charge an intermediate price $p_z \in \left(\underline{p}_z, \bar{p}_z\right)$ that attracts only a subset of consumers into her superstore. Since (ii) would yield zero profits, only cases (i) and (iii) are possible optimal strategies. The next Lemma identifies the interval $\left[\underline{p}_z, \bar{p}_z\right]$ and the subset of consumers entering the superstore when an intermediate price is charged.

Lemma 3.1. *When the seller does not observe consumers' type, she charges a price p_z in:*

$$\left[\underline{p}_z, \bar{p}_z\right] = \left[p_x, \phi^* v(z)\right]. \tag{3.12}$$

Every consumer enters the superstore at $\underline{p}_z$ while only type ϕ^* enters at $\bar{p}_z$. If instead the seller charges an intermediate price, only types in interval $\Phi(p_z)$ below enter her *superstore:*

$$\Phi(p_z) \equiv \left[\underline{\phi}(p_z), \overline{\phi}(p_z)\right] \equiv \left[\frac{2(p_z - p_x)}{v(z) - v(x)}, \frac{2p_x - p_z}{v(x)}\right]. \tag{3.13}$$

Lemma 3.1 states that to attract all consumers into the superstore the seller has to charge the same price for x and z. This is because some types face a self-control cost so high that they value the two goods equally ex-ante. As for the interval $\Phi(p_z)$, notice that as p_z rises, some consumers are not interested in z anymore since the price is above their valuation ($\phi < \underline{\phi}(p_z)$), while some others ($\phi > \overline{\phi}(p_z)$) find (z, p_z) too tempting not to fall to temptation ex-post, but, at the same time, find the price too high to do so. This difference is crucial for the principal's maximization problem. Weakly tempted consumers still have a (weak) incentive to enter the superstore. Strongly tempted consumers, on the other hand, strictly prefer not to, since they would be tempted by offer (z, p_z). Hence, when decreasing (increasing) the price of z marginally, the seller obtains (loses) $\pi_z = (p_z - c_z)$ from the strongly tempted consumers that are now (not) willing to enter the store, and $\pi_z - \pi_x = (p_z - c_z) - (p_x - c_x)$ from the weakly tempted ones that would have entered anyway, but are now buying z instead of x (x instead of z).

Given the above, the seller finds the most profitable intermediate price to charge to consumers:

$$p_z' = arg\max_{p_z} \int_{\Phi(p_z)} \pi_z d\phi + \int_0^{\underline{\phi}(p_z)} \pi_x d\phi = p_x + \tfrac{1}{2}c_z - \frac{c_x v(x)}{v(z) - v(x)}. \tag{3.14}$$

When charging p_z', she obtain profits:

$$\Pi(p_z') = \int_{\Phi(p_z')} (p_z' - c_z) d\phi + \int_0^{\underline{\phi}(p_z')} (p_x - c_x) d\phi. \tag{3.15}$$

She then compares (3.15) with the profits of letting all consumers in the store to buy z. This leads to the formulation of the first Proposition:

Proposition 3.2. *If the seller cannot observe consumers' types in any way, she compares the expected profits of letting every consumer into the superstore with the ones of charging a higher price and attracting only a subset* $\Phi(p_z)$*. This is captured by the following condition:*

$$\Pi(p_z') \geq (p_x - c_z). \tag{3.16}$$

If (3.16) holds, in equilibrium, the monopolist charges price $p'_z = p_x + \frac{1}{2}c_z - \frac{c_x v(x)}{v(z)-v(x)}$. *Consumers in the interval* $\Phi(p'_z) = \left[\underline{\phi}(p'_z), \overline{\phi}(p'_z)\right]$ *enter the superstore and buy good* z. *Weakly tempted consumers outside* $\Phi(p'_z)$ *have the (weak) incentive to enter the superstore to buy* x. *Strongly tempted consumers outside* $\Phi(p'_z)$, *instead, enter one of the smaller stores to buy good* x. *If (3.16) fails, in equilibrium, the seller charges price* $p_z = p_x$. *All consumers enter the superstore and buy* z.

Proposition 3.2 is represented in Figure 3.2. It shows how the monopolist has the option to *screen out* some consumers from the market for z, if she wants to exploit the high willingness to pay of intermediate types. This is because she is unable to create an incentive compatible screening mechanism, since she cannot control the quantity. Hence, an interesting insight of models with self-control preferences: in equilibrium, when a seller excludes from the market for z some consumers, she is forced to exclude consumers with a low valuation, but also the ones with a very high valuation. These latter suffer from strong self-control problems because of their high level of temptation. They value the features of good z a lot, but, because of their self-control, they ignore them ex-ante. This creates a further negative effect on the seller's profits. While the weakly tempted types still have (weak) incentives to enter the superstore and buy x, the strongly tempted consumers are driven away from the superstore and the seller will not be able to sell them x either.

Notice that the equilibrium described here is one in which the seller has relatively poor tools to discriminate between consumers. Not being able to affect the quantity of the good she is forced to either attract every type in her store or to exclude some of them. In the next Section, I show how a seller can use loyalty schemes to "regain" the ability to price discriminate.

3.6 Loyalty Card Schemes

Suppose now that the seller has the ability to offer a loyalty card scheme to consumers. If consumer i accepts (rejects) he faces price p_A (p_R) for good z. When he accepts, he reveals an exogenous amount of information about his willingness to pay to the seller. I model this by assuming that the loyalty card technology applies an exogenous partition $\mathscr{P}$ to the type space, with each element being an interval $\left[\phi^l, \phi^r\right] \subseteq [0,1]$. By accepting the loyalty card, consumer i reveals the belonging of

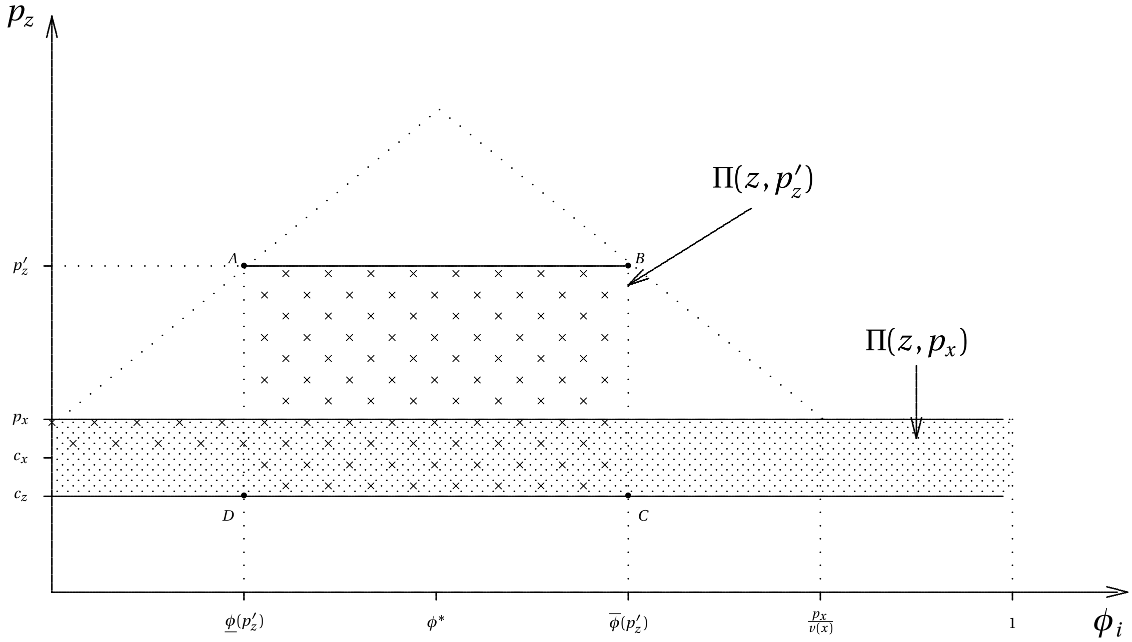

Figure 3.2: Private Information Equilibrium

The shaded areas in the Figure represent the expected profits the monopolist obtains by setting p_x (dotted shaded area) and the ones she obtains setting p_z' (the "×" shaded area). The seller compares these two areas and decides which price to charge. If she sets price p_z', all consumers outside interval $\Phi(p_z') = \left[\underline{\phi}(p_z'), \overline{\phi}(p_z')\right]$ are excluded from the market for good z. Consumers in $(\overline{\phi}(p_z'), 1]$ are driven away form the superstore and buy x from one of the smaller stores.

his ϕ_i to an element of $\mathscr{P}$, without possibility of lying. Hence, the seller can set a specific value of p_A for each element of $\mathscr{P}$. I consider partitions with distinct and ordered elements. This captures the idea that the seller can position the consumer on the $[0, 1]$ spectrum, by exploiting the information granted by the loyalty card.[8]

Formally, a "period 0" is added to the order of play before the ex-ante stage. At the start, the seller sets p_A and p_R, then consumer i decides whether to accept or reject the loyalty card. If he accepts, the monopolist observes the element of the partition to which ϕ_i belongs, this ends period 0. She takes no further decisions in the game and the latter goes on as described in Section 3.4. Notice that if the consumer

[8]To get an idea of the situations is this assumption trying to replicate, consider the case of a seller dividing consumers in groups: "non-tempted", "tempted", "very tempted" etc...

accepts (rejects) the loyalty scheme, he faces menus M_s and $M_m(p_A)$ ($M_m(p_R)$).

In this paper I study all relevant levels of fineness of $\mathscr{P}$, starting from the trivial partition $\mathscr{P}_1 = \{[0,1]\}$. In the latter case, the seller observes only that a consumer who accepts the loyalty card is *willing* to do so. Hence, I call this type of scheme *ϕ-uninformative*. In equilibrium, she adopts this technology in order to attract to her store all the high types that she "loses" when information about ϕ_i is private. In fact, p_R is so high that rejecting the loyalty card removes all the temptation of buying the offer for good z. Consumers then can enter the superstore to buy x without suffering from any self-control cost, using the rejection of the loyalty card as a commitment device.

The other extreme case is to assume that once a consumer accepts the loyalty card he fully discloses his type for the monopolist to see. In this case $\mathscr{P}$ is composed of uncountably many singletons elements, each containing a single ϕ_i. I show how, in equilibrium, every consumer is willing to accept a loyalty card.

In between these two cases there are many possible refinements of the trivial partition $\mathscr{P}_1$.[9] As $\mathscr{P}$ becomes finer, the number of subintervals of $[0,1]$ (its elements) increases. That is, acceptance of the scheme reveals more precise information about ϕ_i. When consumer i accepts a loyalty card, the seller updates her beliefs. She knows that his ϕ_i distributes as a uniform in the element of the partition to which it belongs.

Let $\mathscr{P}_n = \left\{[0,\phi^1),[\phi^1,\phi^2),...,[\phi^{n-2},\phi^{n-1}),[\phi^{n-1},1]\right\}$ be a partition of fineness n (with n elements), then the seller sets $n+1$ prices for z. Namely: p_R, the price of rejection, and the n values of p_A, one for each element of $\mathscr{P}_n$. Formally:

$$p_z = \begin{cases} p_A = \left\{ \begin{array}{ll} p_n & \text{if}\, \phi_i \geq \phi^{n-1} \\ p_{n-1} & \text{if}\, \phi_i \in \left[\phi^{n-2},\phi^{n-1}\right) \\ p_{n-2} & \text{if}\, \phi_i \in \left[\phi^{n-3},\phi^{n-2}\right) \\ ... & \\ p_2 & \text{if}\, \phi_i \in \left[\phi^{1},\phi^{2}\right) \\ p_1 & \text{if}\, \phi_i < \phi^1 \end{array} \right\} & \text{if } i \text{ accepts} \\ p_R & \text{if } i \text{ rejects.} \end{cases} \tag{3.17}$$

In the following subsections I first derive the equilibrium for the case of the triv-

[9]Partition X of a set is a refinement of partition Y if every element of X is a subset of some element of Y.

ial partition. Then I provide a general mechanism to find the equilibrium for any relevant refinement of $\mathscr{P}_1$. Finally, I derive the necessary and sufficient condition on $\mathscr{P}_n$ for the equilibrium of the game to fully replicate first best.

3.6.1 ϕ-uninformative Loyalty Schemes — Trivial Partition

When the seller has access to ϕ-uninformative loyalty schemes, she sets:

$$p_z = \begin{cases} p_A & \text{if } i \text{ accepts} \\ p_R & \text{if } i \text{ rejects} \end{cases} \quad \text{for all } i.$$

As Figure 3.1 shows in Section 3.5.1, strongly tempted consumers are tempted by the first best offer set for types to their left. To see this, consider type ϕ^j and ϕ^{j+h} such that $\phi^* < \phi^j < \phi^{j+h}$, then, for type ϕ^{j+h}, $\big(z, p_z^*(\phi^j)\big) \succ_T (x, p_x)$ and:

$$W_{j+h}\big(\{(x,p_x),(z,p_z^*(\phi^j)),0\}\big) < W_{j+h}\big(\{(x,p_x),0\}\big).$$

When $p_z^*(\phi^j)$ is charged, type ϕ^{j+h} strictly prefers not to enter the superstore.

Now, define p_{max} as the smallest price satisfying $U(z,p_{max})+V_1(z,p_{max})<0$, that is, the smallest price that tempts no types in the market. I show in the proof of Result 8 that if the seller were to set p_{max} as the price of rejection (or acceptance), (z,p_{max}) would not tempt strongly tempted consumers. Hence, they could be attracted back into the store to buy x.

This means that, compared to the case of asymmetric information, increasing marginally the price of good z has now the same positive effect, but a milder negative effect on the seller's profits. When giving consumers the chance to face p_{max} by rejecting the loyalty scheme, the seller does not lose the chance to sell x to the strongly tempted consumers. Hence, the optimal price set for z is not given by (3.14) any longer, but by:

$$\begin{aligned} &arg\max_{p_z} \int_{\Phi(p_z)} \pi_z d\phi + \int_0^{\underline{\phi}(p_z)} \pi_x d\phi + \int_{\overline{\phi}(p_z)}^1 \pi_x d\phi \\ &\Rightarrow p_z^\dagger = \tfrac{1}{2}(\phi^* v(z) + p_x + c_z - c_x) \end{aligned} \tag{3.18}$$

Result 8. *When loyalty cards are ϕ-uninformative in equilibrium the seller sets:*

$$p_z = \begin{cases} p_A = p_z^\dagger & \text{if } i \text{ accepts} \\ p_R \geq p_{max} & \text{if } i \text{ rejects.} \end{cases} \tag{3.19}$$

Consumers in:

- $\left[0, \underline{\phi}(p_z^\dagger)\right]$ *are indifferent between accepting and rejecting the loyalty card and have a (weak) incentive to enter the superstore to buy x,*

- $\Phi(p_z^\dagger)$ *are strictly better off by accepting the loyalty card and buy z from the superstore,*

- $\left[\overline{\phi}(p_z^\dagger), 1\right]$ *strictly prefer to reject the loyalty card. They have a (weak) incentive to buy x from the superstore.*

Notice that the reversed case $p_A \geq p_{max}$, $p_R = p_z^\dagger$ is also an equilibrium, as shown in the proof.

As anticipated, some strongly tempted consumers now reject the loyalty scheme. By doing so, they have the chance to commit themselves to face (z, p_R) — that does not tempt them — once in the superstore, using the rejection of the loyalty card as a commitment device. In this way, the seller is able to attract them back into her superstore to buy x. Hence, the seller is *optimally* inducing some consumers to reject the loyalty scheme.

3.6.2 ϕ-informative Loyalty Schemes

Suppose now that the partition has fineness $n > 1$. As opposed to $\mathscr{P}_1$, loyalty cards are now ϕ*-informative* and the seller can observe to which element of $\mathscr{P}_n$ the ϕ_i of accepting consumers belongs. Figure 3.3 shows an example for $\mathscr{P}_6$.

In order to define the equilibrium for all relevant refinements of $\mathscr{P}_1$, I derive an algorithm that identifies p_A for every level of fineness $n \in [2, \infty)$. For now I rule out the possibility for elements of $\mathscr{P}_n$ to be singletons, for this see Section 3.6.3.

Notice that the optimal pricing scheme set by the monopolist will be of the type described in (3.17). In equilibrium p_A is a discontinuous function of ϕ_i and features n different prices, one for each element of the partition. An equilibrium is therefore defined by a function p_A, a price p_R and consumers' behaviour.

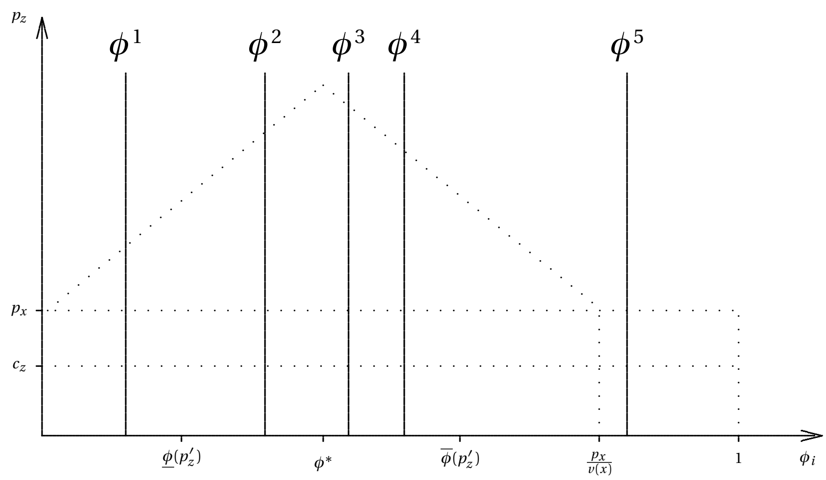

Figure 3.3: Partition Example

The graph shows an example for of $\mathscr{P}_6$. If a consumer accepts the loyalty card, the seller can observe if his ϕ_i is in $[0,\phi^1]$, $(\phi^1,\phi^2]$, $(\phi^2,\phi^3]$, $(\phi^3,\phi^4]$, $(\phi^4,\phi^5]$ or $(\phi^5,1]$.

To identify the algorithm, I divide the problem at hand into n smaller subproblems, one for each interval. The monopolist sets a value for p_A for each subproblem and a value for p_R for all of them. Since consumers cannot lie about their type, the value for p_A set in any interval is irrelevant for all consumers in the other intervals while the value of p_R affects all n of them.

In order to price discriminate between consumers, the seller has the incentive to induce them to accept the loyalty card. This happens under two circumstances: either $p_A > p_R$, and they are tempted by (z, p_R), but it is ex-ante suboptimal to choose this offer ex-post, or $p_R > p_A$, and they would like to choose offer (z, p_A) ex-post. However, since the monopolist wants *different* consumers to buy at *different* prices, it is possible to show that $p_R \geq p_{max} > p_A$ and only consumers that accept the loyalty card buy z.

To see this, consider the case of type ϕ_a, of element a, buying at p_R, after rejecting the card. Notice that (z, p_R) is tempting for all types $\phi_i > \phi_a$. Take, now, consumer $\phi_b > \phi_a$ belonging to element b. If he does not buy z at the p_A set for element b, but is tempted by (z, p_A), he would like to reject the loyalty card and face a high p_R, in order to enter the superstore and buy x free of self-control cost. Since, however, p_R is such that he is tempted also by (z, p_R) he has the incentive to enter one of the smaller stores. The monopolist, therefore, "loses" type ϕ_b. Alternatively, it may be that p_R is so low that ϕ_b is actually willing to reject the loyalty card in order to enter the superstore and choose (z, p_R). If this were to happen, however, the monopolist would be selling z at the same price to consumers belonging to different elements of $\mathscr{P}_n$, losing the opportunity to price discriminate between them.

Setting $p_R \geq p_{max}$ solves both the issues above and does not constrain the monopolist to a suboptimal situation. Consumers in any interval that buy good z will accept the loyalty card while consumers that do not buy z, but are tempted by (z, p_A), can reject the loyalty card to save their self-control cost. Given all this, I will now discuss the equilibrium for a general interval, i.e., a general non-singleton element of a general $\mathscr{P}_n$.

Consider element $\left[\phi^l, \phi^r\right]$ of $\mathscr{P}_n$ where ϕ^l, for "left", and ϕ^r, for "right", represent the endpoints. For each of these elements, the seller faces a problem analogous to the one of Section 3.5.2 but constrained to $\left[\phi^l, \phi^r\right]$. Hence, she first identifies the subinterval of types that enter the store for a given p_A, $\Phi(p_A)$, and then sets the optimal p_A accordingly.

Consumers, instead have an extra choice. If they reject the loyalty card, the ex-ante utility from entering any store in the market is the same since the presence of (z, p_{max}) does not affect their ex-ante utility, from the definition of p_{max}. If they accept the loyalty card and enter the superstore, instead, they face menu $M_m(p_A) = \left\{(x, p_x), (z, p_A), 0\right\}$.
Their decision depends on whether $W_i(\left\{(x, p_x), (z, p_A), 0\right\}) \geq W_i(\left\{(x, p_x), 0\right\})$ or not.

Consider $\phi_i \in \Phi(p_A)$. Since the monopolist sets p_A to sell z to these consumers only, for every $\phi_i \in \Phi(p_A)$, $W_i(\left\{(x, p_x), (z, p_A), 0\right\}) \geq W_i(\left\{(x, p_x), 0\right\})$. The latter holds with strict inequality for types in the interior of $\Phi(p_A)$. They have a strong incentive to accept the card and enter the monopolist's superstore to buy z.[10]

Consumers in $\left[\phi^l, \phi^r\right] \setminus \Phi(p_A)$, instead, have a lower willingness to pay and are of "no interest" for the monopolist. Their equilibrium decision changes depending on whether they are strongly or weakly tempted. To understand why, consider Figure 3.4 below.

Take the case of $\left[\phi^l, \phi^r\right]$ composed of weakly tempted consumers in the Figure. The willingness to pay for z is increasing in ϕ_i in the interval considered. If

[10]These types always exist, since profits from (z, p_A) are zero for any p_A such that $\Phi(p_A)$ is a singleton.

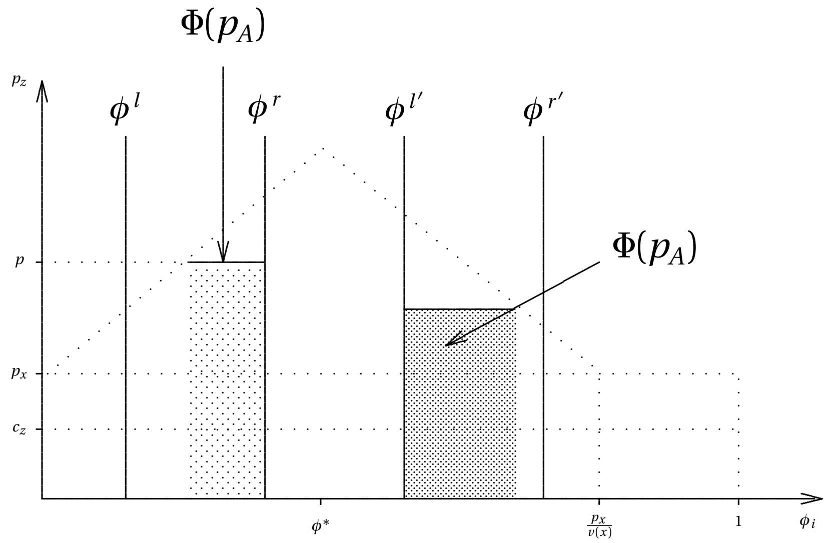

Figure 3.4: Consumers' Acceptance

Two examples of behaviour of consumers of a general element of $\mathscr{P}_n$. When the interval is composed only of weakly tempted consumers, e.g. $[\phi^l, \phi^r]$, the light shaded area represents the set of consumers accepting the loyalty card and entering the superstore. Consumers to the left of the area are indifferent between accepting and rejecting the loyalty card. When the interval is composed only of strongly tempted consumers, e.g. $[\phi^{l'}, \phi^{r'}]$, the dark shaded area represents the set of consumers accepting the loyalty card and entering the superstore. Consumers to the right of the area strictly prefer to reject the loyal card.

the monopolist wants to attract a subset of these consumers into the store, she will set a price p_A in the range $\left[p_z^*(\phi^l), p_z^*(\phi^r)\right]$. From the willingness to pay curve, the least tempted consumer to enter the store accepting the loyalty card is the ϕ_i such that $p_z^*(\phi_i) = p_A$. All consumers to his right accept the loyalty card and buy z. All consumers to his left are indifferent between accepting or rejecting the loyalty card and have a (weak) incentive to enter the retailer's superstore — as Proposition 3.2 describes. Since the seller has no information inside the interval to discriminate between these types and consumers that buy z at p_A, she can only screen them out of the market for z.

Consider now the case of interval $\left[\phi^{l'}, \phi^{r'}\right]$ of strongly tempted consumers. The same logic explained above applies. However this time, given a price p_A, the *most* tempted consumer to accept the card is the ϕ_i such that $p_z^*(\phi_i) = p_A$. All consumers to his left accept the card and buy z. All consumers to his right, instead, are tempted by (z, p_A). Hence, as in the trivial partition case, they use the rejection of the loyalty card as a commitment device, not to be tempted ex-post. In the second stage of the game they have a (weak) incentive to enter the superstore.

To summarise, the solutions to each of the n subproblems of intervals $\left[\phi^l, \phi^r\right]$ have a common structure. The seller sets $p_R \geq p_{max}$ and p_A according to set $\Phi(p_A) \subseteq \left[\phi^l, \phi^r\right]$. Consumers in $\Phi(p_A)$, indeed, accept the loyalty card and enter the superstore to buy z. Consumers in $\left[\phi^l, \phi^r\right] \setminus \Phi(p_A)$ behave as explained above. The combination of all the n solutions describes the equilibrium of the game for a general partition $\mathscr{P}_n$. The next set of Lemmas provides an algorithm that identifies function p_A and Proposition 3.9 identifies the equilibrium for every $n \in [1, \infty)$. Define $p_1, p_2, ..., p_n$ where p_1 is the value for p_A in the first element of $\mathscr{P}_n$, $[0, \phi^1)$, p_2 the one for the second, $[\phi^1, \phi^2)$, and so on. The algorithm changes depending on the position and size of the n elements of $\mathscr{P}_n$.

First of all, consider a general element $\Phi_i = \left[\phi^l, \phi^r\right]$, where $\phi^l, \phi^r \in [0,1]$. If $\Phi(p_z^\dagger) \subseteq \Phi_i$, the seller sets $p_z = p_z^\dagger$ as in Section 3.6.1. The intuition behind this is that consumers outside $\Phi(p_z^\dagger)$ are excluded from the market for z when the price is set for the entire $[0,1]$ interval. Hence, restricting the set of consumers to Φ_i with $\Phi(p_z^\dagger) \subseteq \Phi_i \subseteq [0,1]$, adds no relevant information to the seller's optimization. Lemma 3.3 states this concept formally.

Lemma 3.3. *If an element of the partition is composed of at least all consumers that buy z in the case of ϕ-uninformative loyalty schemes, then the price of acceptance is set to $p_A = p_z^\dagger$.*

The rest of the algorithm, therefore, describes the rules for the case of $\Phi(p_z^\dagger) \nsubseteq \Phi_i$. For simplicity, I being by studying the first, p_1, and the last, p_n, prices of the algorithm and then move to all the ones in the middle. I start with Lemma 3.4, that describes the value of p_1.

Lemma 3.4. *If $\Phi(p_z^\dagger) \nsubseteq \Phi_i$, the value of p_1 depends on the composition of the first element of $\mathscr{P}_n$, $\Phi^1 = [0, \phi^1)$.*
When element Φ_1 is composed only of weakly tempted consumers, i.e. $\phi^1 < \phi^*$, the

value of p_1 is given by:

$$\Rightarrow p_1 = \max\left\{p_x, \arg\max_{p_z} \int_{\underline{\phi}(p_z)}^{\phi^1} \pi_z d\phi + \int_0^{\underline{\phi}(p_z)} \pi_x d\phi\right\}, \tag{3.20}$$

When element Φ_1 is composed of both weakly and strongly tempted consumers, i.e. $\phi^1 \in [\phi^*, \overline{\phi}(p_z^\dagger))$, the value of p_1 is given by:

$$\Rightarrow p_1 = \max\left\{p_x, \min\left\{p_z^*(\phi^1), \arg\max_{p_z} \int_{\underline{\phi}(p_z)}^{\phi^1} \pi_z d\phi + \int_0^{\underline{\phi}(p_z)} \pi_x d\phi\right\}\right\}, \tag{3.21}$$

The maximisation problem in (3.20) and (3.21) is analogous to (3.14). The seller may screen out some consumers to extract the highest willingness to pay of others. As the proof of the Lemma shows, however, only the left boundary of $\Phi(p_A)$ now depends on p_z. Of course, as mentioned above, all weakly tempted consumers excluded from the market for z now buy good x.

To understand the intuition behind (3.21) notice that the maximisation problem $\arg\max \int_{\underline{\phi}(p_z)}^{\phi^1} (p_z - c_z) d\phi$ ignores the negatively slope portion of the willingness to pay of consumers between ϕ^* and $p_z^*(\phi^1)$. Given this, its solution may very well be larger than $p_z^*(\phi^1)$. This, however, would imply that the derived p_z should be a local maximum also when loyalty schemes are ϕ-uninformative. Result 8 shows that this is not the case.

Before moving to the more complicated derivation of $p_2, ..., p_{n-1}$, I identify p_n, the price the monopolist charges to consumers that prove to be in the last element of $\mathscr{P}_n$, $\Phi_n = [\phi^{n-1}, 1]$.

Lemma 3.5. *If $\Phi(p_z^\dagger) \nsubseteq \Phi_i$, the value of p_n depends on the composition of the last element of $\mathscr{P}_n$, $\Phi_n = [\phi^{n-1}, 1]$.*
When element Φ_n is composed only of strongly tempted consumers, i.e. $\phi^{n-1} > \phi^*$, the value of p_n is given by:

$$\Rightarrow p_n = \max\left\{p_x, \arg\max_{p_z} \int_{\phi^{n-1}}^{\overline{\phi}(p_z)} \pi_z d\phi + \int_{\overline{\phi}(p_z)}^{1} \pi_x d\phi\right\}, \tag{3.22}$$

When element Φ_n is composed of both weakly and strongly tempted consumers, i.e.

$\phi^{n-1} \in [\underline{\phi}(p_z^\dagger), \phi^*)$*, the value of* p_1 *is given by:*

$$\Rightarrow p_n = \max\left\{p_x, \min\left\{p_z^*(\phi^{n-1}), \arg\max_{p_z} \int_{\phi^{n-1}}^{\overline{\phi}(p_z)} \pi_z d\phi + \int_{\overline{\phi}(p_z)}^{1} \pi_x d\phi\right\}\right\}, \tag{3.23}$$

The intuition behind Lemma 3.5 is similar to the one described for Lemma 3.4.

I now identify $p_2, ..., p_{n-1}$. I look at the equilibrium p_i for the general element $\Phi_i = \left[\phi^l, \phi^r\right)$ of $\mathscr{P}_n$ where this time $\phi^l, \phi^r \in (0,1)$. The value of p_i is affected, in particular, by the composition of Φ_i, i.e., on what kind of temptation the consumer inside it suffer from. Lemma 3.6 describes the case of Φ_i to be composed only of weakly or strongly tempted consumers, while in Lemma 3.7 both types of tempted consumers can be part of Φ_i.

Lemma 3.6. *When element* Φ_i *is composed only of weakly tempted consumers (i.e.* $\phi^r < \phi^*$*), the value of* p_i *is given by:*

$$p_i = \max\left\{p_z^*(\phi^l), arg\max_{p_z} \int_{\underline{\phi}(p_z)}^{\phi^r} \pi_z d\phi + \int_{\phi^l}^{\underline{\phi}(p_z)} \pi_x d\phi\right\} \tag{3.24}$$

When element Φ_i *is composed only of strongly tempted consumers (i.e.* $\phi^l > \phi^*$*), the value of* p_i *is given by:*

$$p_i = \max\left\{p_z^*(\phi^r), arg\max_{p_z} \int_{\phi^l}^{\overline{\phi}(p_z)} \pi_z d\phi + \int_{\overline{\phi}(p_z)}^{\phi^r} \pi_x d\phi\right\} \tag{3.25}$$

The intuition behind Lemma 3.6 is similar to the ones discussed for the previous Lemmas. Notice, however, that for the case of only weakly tempted consumers, p_x is never the price of equilibrium for z in the interval, since all consumers in it have a willingness to pay strictly larger than p_x. Hence, the lower bound for the optimal price is not p_x anymore but rather $p_z^*(\phi^l)$.

In the case of an interval with only strongly tempted consumers, instead, p_i can indeed be equal to p_x. Setting $p_i = p_x$ may become optimal if there are consumers in Φ_i that, ex-ante, value good z as good x (i.e. $\phi^r > \frac{p_x}{v(x)}$).

To conclude the algorithm, Lemma 3.7 analyses the case of an interval that is not the superset of $\Phi(p_z^\dagger)$, but it is composed of both weakly and strongly tempted consumers, i.e., $\phi^* \in \Phi_i$. Notice that in this case knowing the proportion of strongly vs. weakly tempted consumers in the interval is not enough to identify the maximisation problem that provides the optimal price. To see why, let the decreasing portion of the willingness to pay curve be flatter than the increasing portion. Assume the case of an interval composed mostly of strongly tempted consumers, i.e., $(\phi^r - \phi^*) \geq (\phi^* - \phi^l)$. If the decreasing portion of the curve is flat enough, the willingness to pay of ϕ^r is larger than the one of ϕ^l. In this case, the optimal p_i belongs to $[p_z^*(\phi^l), p_z^*(\phi^r)]$.

Lemma 3.7. *Let $\Phi(p_z^\dagger) \nsubseteq \Phi_i$. When element Φ_i is composed of both strongly and weakly tempted consumers (i.e. $\phi^* \in \Phi_i$), the value of p_i depends on its composition.*
If the least tempted consumer in the element has the lowest willingness to pay then:

$$\Rightarrow p_i = \max\left\{p_z^*(\phi^l), \min\left\{p_z^*(\phi^r), arg\max_{p_z} \int_{\underline{\phi}(p_z)}^{\phi^r} \pi_z d\phi + \int_{\phi^l}^{\underline{\phi}(p_z)} \pi_x d\phi\right\}\right\} \quad (3.26)$$

If the most tempted consumer in the element has the lowest willingness to pay then:

$$\Rightarrow p_i = \max\left\{p_z^*(\phi^r), \min\left\{p_z^*(\phi^l), arg\max_{p_z} \int_{\phi^l}^{\overline{\phi}(p_z)} \pi_z d\phi + \int_{\overline{\phi}(p_z)}^{\phi^r} \pi_x d\phi\right\}\right\} \quad (3.27)$$

Lemma 3.7 is a combination of the previous ones and completes the algorithm.

In Figure 3.5, I show an example of how to apply these rules. Notice, however, that in this Section I have not considered the case of singleton elements of $\mathscr{P}_n$. That is not because the algorithm described here does not apply in the case of singleton elements. Indeed it does. However, a much easier and direct solution can be derived.

3.6.3 Singleton Intervals and Full Disclosure

To conclude the analysis of all relevant refinements of $\mathscr{P}_1$, in this Section I consider partitions $\mathscr{P}_n$ with at least one singleton element. If consumers with types belonging to such an element accept the loyalty card, they perfectly disclose their type to the seller.

Result 9. *If by accepting the loyalty scheme, consumer i perfectly disclose his type to the seller, the equilibrium value of p_A is a function of ϕ_i and coincides with the first* *best price:*

$$p_A = p_z^*(\phi_i). \tag{3.28}$$

Every consumer with this option (i.e. he belongs to a singleton element of $\mathscr{P}_n$) accepts the loyalty scheme and enters the store to buy good z.

The Result implies the following Corollary, one of the main points of the paper:

Corollary 3.8. *A sufficient condition for first best to be fully restored is for all consumers that differ in their ex-ante willingness to pay to be able to perfectly disclose their type by accepting the loyalty scheme.*

In other words, if $[1, \frac{p_x}{v(x)}]$ is partitioned in uncountably many singleton elements by $\mathscr{P}_n$, first best is fully restored.

To see why this is true, simply notice that consumers beyond $\frac{p_x}{v(x)}$ have the same ex-ante willingness to pay. Hence, there is no need for the monopolist to discriminate among them as long as he can separate them from the rest. In equilibrium, she charges (3.28) to consumers in the infinite portion of $\mathscr{P}_n$ and $p_z = p_x$ to the rest.

3.6.4 Equilibrium

The subgame perfect equilibrium for any possible partition $\mathscr{P}_n$ is summarised in the next Proposition.

Proposition 3.9. *Given partition $\mathscr{P}_n$ of the type space, suppose the seller can offer loyalty schemes to consumers. When consumer i accepts, he reveals to the seller that his type belongs to a specific element of $\mathscr{P}_n$. She, in turn, charges the consumer p_A instead of p_R (which he faces when rejecting) for z. In equilibrium she sets price $p_R \geq p_{max}$ for every level of fineness n while the value of p_A depends on the structure of $\mathscr{P}_n$.*

- *If $n = 1$, then p_A follows (3.19).*
- *If $n > 1$, then p_A follows (3.20)-(3.27).*
- *If $\mathscr{P}_n$ has singleton elements, then p_A for all singleton elements equals the price of first best, (3.28).*

Given p_A, all consumers in $\Phi(p_A)$ accept the loyalty card, enter the superstore and buy good z. As for the ones outside $\Phi(p_A)$, weakly tempted consumers are indifferent *between accepting and rejecting the loyalty card while strongly tempted consumers strictly prefer to reject it. All have a (weak) incentive to enter the superstore to buy x.*

The above expresses the second main result of the paper and it implies the following:

Corollary 3.10. *When consumers' type are not commonly known, the seller uses a loyalty card scheme to acquire information about their preferences. Some consumers accept the loyalty card,* voluntarily *disclosing their type to the seller. Others prefer to reject the loyalty card in order to face a high price for the tempting good. The seller uses this information to price discriminate between consumers. Generally, the more precise the information acquired through the scheme, the more the equilibrium replicates first best.*

Proposition 3.9 also shows that interval $\Phi(p_A)$ does not need to be continuous since p_A is not a continuous function when $n \geq 2$. It also shows that p_A is non-monotonic in ϕ_i. This follows from the fact that consumers' willingness to pay is non-monotonic in their valuation of the good, as explained by the ex-ante participation constraint (PC) described in Section 3.4. This is specific of self-control models.

3.6.5 Example

Let $p_x = 2$, $c_x = 1$, $c_z = \frac{1}{2}$, $v(z) = 21$ and $v(x) = 3$. To describe the results of the paper, I provide an example in Figure 3.5. Consider an exogenous partition of the following type:

$$\tilde{\mathscr{P}}_n = \Big\{[0,0.01),[0.01,0.14),[0.14,0.25),[0.25,0.45),[0.45,0.55),\{j\}_{j\in[0.55,0.75)},[0.75,1]\Big\}.$$

Such a partition is composed of uncountably many elements. If a consumer in $[0,0.01),[0.01,0.14),[0.14,0.25),[0.25,0.45),[0.45,0.55)$ or $[0.75,1]$ accepts the loyalty scheme, the seller only knows that his type is distributed uniformly in the element he belongs to. If, instead, a consumer in $[0.55,0.75)$ accepts the loyalty scheme, then the seller perfectly observes his type.

For the values set above, $\phi^* = 0.16$, $p'_z = \frac{25}{12}$ and $\phi(p'_z) = [0.01,0.64]$. This generates Figure 3.5, where p_1 is set according to (3.21), p_2 according to (3.24), p_3 according to (3.27), p_4 and p_5 according to (3.25), $p_6, p_7, ..., p_{n-1}$ according to (3.28), p_n according to (3.22) and $p_R \geq 3.36 = \overline{p}_z$.

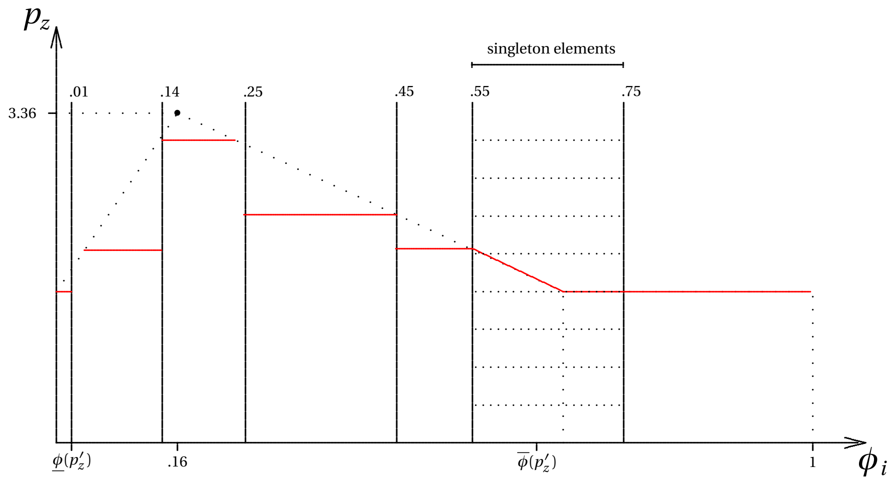

Figure 3.5: Numerical Example

In the Figure the equilibrium for the special case of $\tilde{\mathscr{P}}_n$ is shown. The partition features uncountably many singleton elements in the interval $[0.55, 0.75)$. Price p_A is described by the, discontinuous, function represented by the red line. The function is plotted only for consumers that accept the loyalty card, enter the superstore and buy good z.

The Figure shows the optimal price set for each element (the red line). The smaller are the elements of the partition, the more the equilibrium function p_A resembles the price of first-best. In particular, it is easy to see how in the interval of the type space partitioned by singleton elements, first best is perfectly restored.

3.6.6 Consumer Surplus

I now analyse the effect of loyalty schemes on consumers' welfare. The textbook model of 1^{st}-degree price discrimination describes it as the worst situation for consumers.[11] Prices are set equal to the willingness to pay and consumer surplus is fully extracted by the seller. Shiller (2015) reproduces this result through simulations for a case of individual pricing with consumer tracking.

[11]Notice that first degree price discrimination is the closest relative to individual pricing but the two are not exactly the same. Individual pricing, in fact, does not imply that the seller is able to charge the exact willingness to pay to each consumer.

Carrying out consumers' welfare analysis for models with time inconsistent agents is not a straightforward task. The, to some extent ethical, main question one should address is: which utility function should be considered to measure consumers' surplus? Should the "ex-ante consumer" be given more or less importance than the "ex-post one"?

In this paper, I can circumvent this question and provide predictions on consumers' welfare. To see why this is the case, notice that the effect of a change in price, or in the decision of the good bought, affect the ex-ante and ex-post utility in the same way.

Given Assumption 1, the minimum ex-ante utility consumers can obtain in the model is given by $W(\{(x,p_x),0\})$, while the ex-post is given by $U(x,p_x)+V(x,p_x)$. Consider the case of a consumer buying x when loyalty schemes are not used. If after the introduction of loyalty schemes the consumer still buys good x, then his utilities, and therefore his welfare, are unchanged. If, instead, he switched to buying z then it must be that both his ex-ante and ex-post utility are at least as high as before, since his PC and IC in (3.8) hold. Hence, if a consumer buys x before loyalty schemes are in use, his welfare once they are introduced is at least as high as before.

Consider now the case of a consumer that buys good z before loyalty cards are introduced. Notice that his PC and IC are holding. Assume that after the loyalty schemes are introduced, he switches to x. His utilities, both ex-ante and ex-post are at most the same as in the case of asymmetric information, since PC and IC now fail. If, instead, he still buys z when loyalty schemes are introduced, then the effect on welfare depends solely on price. If the price he pays has increased (decreased) his utilities have decreased (increased).

Ultimately, therefore, loyalty schemes have two opposite effects. First, they lower the price for some "low types" in order to induce them to buy z as well. Second, they increase the price charged to "high types" in order to extract more surplus from them. Hence, the overall change in welfare depends on which one of these effects dominates the other. Proposition 3.11 below summarises the result.

Proposition 3.11. *Consumers that buy good x in the absence of loyalty schemes, face a non-negative change in welfare when the schemes are introduced. Consumers that buy good z in the absence of loyalty schemes and switch to x when the latter are introduced, face a non-positive change in welfare when the schemes are introduced.*

Consumers that buy good z in both cases, face a change in welfare inversely proportional to the change in the price p_z they pay.

This result implies that there exist partitions $\mathscr{P}_n$ such that the introduction of loyalty schemes increases consumers' welfare overall.[12]

The main message behind Proposition 3.11 is that individual pricing per se is not necessarily hurting consumers. It may, in fact, allow consumers with low willingness to pay to afford goods they would not buy otherwise — i.e., at the prices charged to other consumers in the market.

3.7 Discussion

This paper addresses the economic problem of a seller that offers her tempting product, z as a monopolist in a market where an imperfect substitute, x, is sold. She faces a continuum of consumers with self-control preferences à la Gul and Pesendorfer (2001). The problem is analysed under full information, asymmetric information and when the seller makes use of loyalty schemes — as described in Section 3.6.

The benchmark case shows the peculiarities of equilibria of self-control models. As it is easy to see from Figure 3.1, the first-best price, $p_z^*(\phi_i)$, is described by a hill-shaped function. The seller is unable to extract the entire ex-post surplus from strongly tempted consumers. In equilibrium, they enter the superstore to buy z obtaining the same ex-ante utility granted by the outside option — entering a smaller store to buy x. Ex-post, however, they obtain a positive surplus, $U(z, p_z^*(\phi_i)) + V_i(z, p_z^*(\phi_i)) > U(x, p_x) + V_i(x, p_x) \geq 0$. This has two interpretations. First, because of their temptation, consumers are not willing to fall to temptation ex-post, i.e., to buy good z, unless they obtain a substantially large reward ex-post (as for example a consumer on a diet who avoids burger shops unless prices are extremely low). Second, in the ex-ante stage, consumers behave as in a dual-self model (Strotz, 1956). They "play against their future self" and do not enter the superstore unless they know that the choice their future self makes ex-post is one that is optimal for them as well. This makes high types the most valuable in the market

[12] Of course, it is easy to see that this does not hold if loyalty schemes are capable of perfectly restoring first best.

no more and distinguishes the model from one without self-control preferences.

Following this intuition, the asymmetric information equilibrium shows how high types are screened out of the market together with the low types. These latter are not valuable *enough* for the seller while the former are "*too* valuable" and their self-control problem makes their willingness to pay decrease so much that the seller is better off by excluding them from the market for z in the first place. While the first ones are still willing to buy x from the seller, however, the strongly tempted are driven away from the superstore.

In equilibrium, the seller only attracts in the superstore *averagely* tempted types that value the good enough to be willing to pay a high price, but not so much as to suffer a large self-control cost of falling to temptation.

So far, the monopolist faces two problems when information about temptation levels becomes private. First, she is unable to price discriminate because she lacks control on the quantity sold. Second, she "loses" some consumers that are too tempted to enter her store. I show how, using loyalty card schemes that partition the type space as in Section 3.6, she is able to acquire information about consumers' preferences and (at least partially) solve these problems, increasing her profits. In particular, the more precise is the information acquired through the loyalty card, the more the equilibrium replicates first best.

This equilibrium convergence can be seen graphically in the discussion about Figure 3.5. Its intuition, however, is much simpler to visualise. When the information about consumer i's level of temptation becomes more precise after the latter accepts, the seller is able to tailor a more personalised pricing scheme. As the level of fineness rises she can target narrower groups of consumers. As these groups shrink in size and increase in number they converge to singletons (containing only one type of consumer). At the limit, the seller can target a single consumer with an individually specific price. By charging $p_R \geq p_{max}$ when consumers reject, she removes the possibility to have an outside option of buying z. This implies that first best is restored in all the intervals of the type space that constitute a singleton element of the partition. When this holds for all consumers that differ in their willingness to pay, the equilibrium of the game is equivalent to the one of first best and the seller behaves as in a first-degree price discrimination (or individual pricing) model.

The second main point of the paper, studying why would a consumer want to avoid getting personalised discounts. Proposition 3.9 also states that as long as there exist some non-singleton element of strongly tempted consumers in the partition, there may also be types that strictly prefer to reject the loyalty card. These consumers are the strongly tempted ones that do not buy at the value of p_A for their interval. They find this price low enough to tempt them but high enough to make it ex-ante suboptimal to choose (z, p_A). Hence, by rejecting the loyalty card they are sure of not being tempted once in the superstore (from the definition of p_{max}). Ultimately then, these consumers use the rejection of the loyalty card as a commitment device to remove the tempting offer from the menu saving all the self-control cost. The seller, in turn, is willing to offer them such a device since in this way she can attract them back into her superstore to buy good x.

3.8 Existing Literature

This paper contributes to the literature on optimal pricing when consumers suffer from bounded rationality. It enters the very vast literature that studies loyalty schemes that goes from Marketing and Retailing to Industrial Organisation. In this Section, I discuss the closest papers to this one.

The seminal paper for self-control preferences and temptation is Gul and Pesendorfer (2001). In Section 3.2 I explain the approach they take on these topics. Applications of this model have been studied under several different settings. Esteban and Miyagawa (2005) consider two types of tempted consumers and show how asymmetric information does not, always, impede first best allocation in these type of models. Esteban and Miyagawa (2006) and Esteban, Miyagawa, and Schum (2007) extend the model to, respectively, a market with perfect competition and a continuum of consumers. In Esteban, Miyagawa, and Schum (2007), however, the monopolist is forced to offer a single menu of offers ex-ante. In Foschi (2014), I study the issue of commitment and self-control. There, the monopolist cannot commit ex-ante to the prices and quantities offered ex-post. The main difference between these papers and the present one — besides the loyalty card schemes — is that in the literature cited above sellers control also the quantity of the good. This allows them to create incentive compatible menus of offers that may not need the use of loyalty cards.

Loyalty cards have been studied both from a theoretical and from an empirical point of view. Empirically studies have focused on the effect that loyalty schemes have on consumers' lifetime value, loyalty enhancement, duration and retention (among others, Byrom, 2001; Lewis, 2004; Meyer-Waarden, 2007; Gomez, Arranz, and Cillàn, 2006).

Theoretical studies on loyalty schemes contribute to the Competition Policy and Industrial Organisation literatures. To my knowledge, however, the literature lacks a study that models loyalty cards as a tool for price discrimination. They have been studied as business-stealing tools (Caminal and Claici, 2007), as bundled loyalty discounts (Greenlee, Reitman, and Sibley, 2008), as "bribes" to agents that buy products with principal's money (Basso, Clements, and Ross, 2009), as collusion tools (Ackermann, 2010) and in dynamic environments as tools to increase consumers' participation and expenditure (Chen and Pearcy, 2010; Caminal, 2012). Ackermann (2010)'s loyalty card schemes offer discounts similar to the ones analysed in this paper. His findings focus on the competition aspect of loyalty cards. Caminal (2012) deals with a model closer to the one described in this paper, but focuses on the aspect of consumers' preference dependance across two periods. The intuition and idea of loyalty schemes is fundamentally different from the one described in this paper. There, they are rewards and tools used to *continue* the purchasing relationship. Here, they are a means of *exchange* that the monopolist uses to acquire information about consumers' level of temptation.

This interpretation follows the literature on Big Data and individual pricing. Evidence of the existence of individual pricing in markets where sellers have access to large and comprehensive data sets (Big Data) — like browsing data — has been observed by Mikians, Gyarmati, Erramilli, and Laoutaris (2012, 2013) and Hannak, Soeller, Lazer, Mislove, and Wilson (2014). Shiller and Waldfogel (2011) and Waldfogel (2015) study individual pricing in the music market and in the market for professional higher education. Both papers, however, do not study consumer tracking. Shiller (2015) simulates the effect on profits of a seller that performs consumers tracking via web browsing data. To my knowledge, my model is one of the first to provide a formal theoretical framework of individual pricing with consumer tracking and to study its effects on welfare.

3.9 Conclusions

The contribution of this paper is to model the seller-consumer relationship in the presence of self-control preferences and loyalty card schemes. I provide a theoretical framework able to analyse the effect of individual pricing in the presence of consumer tracking. It answers two main questions: how can a seller price discriminate between heterogenous consumers when she has no control over quantity? How can she sell her products to consumers that avoid her store to save the self-control cost? I show in the previous sections that loyalty card schemes are a way for the seller to address both these issues. I model them as a means of exchange rather than as a competition tool. Through the scheme, the seller offers consumers personalised discounts in exchange for (unalterable) information about their preferences. As this information becomes more precise, the first best may be restored. Moreover, by setting a very high price of rejection, the seller is able to convince strongly tempted consumers that their temptation will not be exploited once they enter the store. The latter, therefore, are happy to carry out their shopping free of self-control costs.

I show how the welfare effects of such a technology are non-trivial and depend on the amount of information the seller can obtain via the scheme. In particular, I argue that there are conditions under which loyalty schemes and individual pricing can increase consumer surplus overall.

Current and future research is now in progress on endogenous, costly partition setting. Assuming the seller has the option to set up the partition with a fixed amount of (costly) elements, how would the optimal partition be structured? This and under interesting follow-up questions are now under study.

Chapter 4

Asymmetric Information, Commitment and Self-Control

Chapter Abstract

I study a two period model where the consumer suffers from self-control problems and his level of temptation is private information. I derive the optimal behaviour of a seller that offers her product to a consumer. In period 1, the consumer decides whether or not to "enter the store" based on the prices posted by the seller. In period 2 he decides how much of the product to buy, if any. Differently from the existing literature, I assume that the seller cannot commit to the prices posted in period 1. I show how, under this framework, the presence of tempted consumers and asymmetric information can explain the existence of entry bonuses paid by the seller to the consumer in exchange for entering the store. In contrast with classical contract theory, I show that the relatively untempted consumer (the "low type") can be better off when information about his type is private than when the seller is fully informed. Moreover, the presence of self-control may induce the seller to exclude the relatively strongly tempted consumer (the "high type") from the market.

4.1 Introduction

I study markets with heterogonously tempted consumers who suffer from self- control problems. Following the existing literature on temptation models with asymmetric information, I examine a two period game where the purchase takes place in period 2. Differently from other contributions (see Esteban and Miyagawa, 2005, 2006; Esteban, Miyagawa, and Schum, 2007, among others), however, I assume a seller that cannot commit in period 1 to the prices and quantities she sets in period 2. This generates two main results. First, I show why these markets are often characterised by entry bonuses that the seller pays to the buyer before the latter makes his purchase decision. Second, I highlight how a consumer that values the good more than others may suffer from higher self-control problems. In period 1, he anticipates his self-control cost, which decreases his willingness to pay. In period 2, he faces the good while prey to temptation, which rises his willingness to pay. The opposite is true for consumers with low valuations of the good. In contrast with classical price discrimination problems results (e.g. Spence, 1977), this causes a role "swap" between "high" and "low" types of consumer. Under some conditions, the former is excluded from the market and the latter obtains a positive surplus.

Nevertheless, self-control is not the only way to fight temptation. Sometimes we are able to anticipate our later decisions and understand the self-control cost we may bear. In such cases we might decide to avoid the temptation in the first place by, for example, not entering an ice-cream shop if we are on a diet or not registering to an online casino if we cannot resist gambling.

Consider the following case of a seller and a buyer. The seller runs a store offering a specific good. The buyer is in the street in front of the store and is willing to purchase the good. Before entering the store, he has a clear idea of the quantity of the good he wants to buy. He knows, however, that, once in the store, he will face temptation, that is, he will be willing to buy more (*upward* temptation), or less (*downward* temptation), than he was willing to buy when he was still outside the store. Therefore, if the buyer enters the store he will need to exert self-control effort to resist temptation. He then can either decide to bear this cost, and enter the store, or to avoid temptation completely, by not entering the store and walking past. Throughout the rest of my analysis, I call markets for products or services where temptation is an issue *temptation markets*.

Temptation markets have been studied extensively in the literature, and I analyse the closest papers to this in Section 4.2. In particular, Esteban and Miyagawa (2005, 2006) and Esteban, Miyagawa, and Schum (2007) provide an extensive analysis of temptation markets with asymmetric information, where self-control is modelled à la Gul and Pesendorfer (2001). All these papers, however, make the implicit assumption that the seller is able to completely commit herself in the first period to what she offers in the second period. I claim that such an assumption is not easily justifiable in a large proportion of temptation markets. One pertinent example is online temptation markets, described below.

The increasing prevalence of online markets and online payments over the last ten years has reduced the sellers' cost associated with reaching consumers and exploiting their temptation.[1] A simple click on a banner has replaced the need to physically go into a shop and use the money in your pocket to buy a good, drastically cutting transaction costs. Moreover, the increased safety in online payments has decreased uncertainty and allowed the online market to expand even more. Exactly as in offline (on the street) markets, consumers make two different choices in online markets. On the street they first choose whether to enter the store and then whether to buy the good. Online, instead, they first choose whether to register to the website and then whether to buy the product or the service the website offers.

Some interesting examples of such online temptation markets are online casinos, betting sites, "app-stores" and the "apps"[2] that are sold there.[3] Sellers in these markets require customers to register before making a purchase. They asked to create a personal account linked to a payment method that is registered by the website or app. In this way, every time they want to make a purchase the only effort required is to go online, log in and click a button.[4]

[1] On 10th January 2014, BBC news reported that online trading in December 2013 accounted for the 18.5% of total trading of non-food goods and services in the United Kingdom: an Increase of 2 percentage points from 2012.

[2] An app is a piece of software that runs on mobile devices such as smartphones and tablets. It is sold in online stores which are called "app-stores". A specific feature of these stores is that, most of the time, consumers can buy and download apps for their devices only via the device's producer's app-store and not via the competitors' ones, thus creating market power for the device's producer.

[3] The model could be applied to big online retailers with accounts systems and varying prices — like Amazon — as well.

[4] To get an idea of the size and relevance of such markets: in a report by PricewaterhouseCoopers in 2011 on "the online gaming market evolution to 2015", online casino markets' sales amounted to \$117 billion, 49% of which is in the United States; the IHS screen digest of February 2011 shows that the mobile application market sales of the leading company — Apple — in 2011 amounted to \$1.78

While the link between temptation, self-control, obsessive impulses and gambling is well known and analysed in the psychology literature (see Nower and Blaszczynsky, 2004, and the references therein), it is worth exploring how temptation enters the market for apps. The majority of software sold in app-stores is downloadable free of charge and is characterised by "in-app purchases".[5] An in-app menu of extra offers is available that makes the app better or brings it "to the next level".[6] Often, the menu of in-app purchases changes after the consumer has registered and while he is using the software. Some specific features of the software may become free while others become available upon payment and part of the "in-app purchases"; completely new features can be added to the in-app menu. Therefore, a tempted consumer may succumb to temptation and buy more than he planned when he registered to the app-store and downloaded the free app.

Online temptation markets share two important features. First there is often an "entry bonus". Most of the time, if not always, free-entry online casinos and free apps also give the consumer free credit to spend on goods and services offered by the website. Second, the frequency with which prices change *after* the registration is high. This second point is the main reason why the assumption of a fully committed seller in all temptation markets seems quite strong. In this paper, therefore, I relax the perfect commitment assumption. By doing so I am also able to show that the presence of tempted consumers can motivate the use of entry bonuses.

At first, the reason for setting up entry bonuses may seem trivial: sellers use them to attract consumers ex-ante and exploit their temptation ex-post. Notice, however, that, when types are private information, the seller pays these bonuses to *every* consumer in the market and that nothing prevents consumers from registering, obtaining the entry bonus, spending it, and never logging in to the website ever again. Depending on their level of temptation, this behaviour could be perfectly rational. Hence, the puzzle remains.

This paper relaxes the general assumption of perfect commitment and shows

billion.

[5]An in-app feature is a service or software add-on component that works through, and only through, the app.

[6]A prominent example are smartphone-apps that improve the quality of the built-in camera or add picture-filters to make your picture look old, or to turn it into black and white. These apps are, often, sold free of charge and they come with basic features. If the buyers is willing to acquire the entire services offered by the app — for example, buy the filter to take pictures in black and white — he has to do it via the in-app store.

how a model with tempted consumers who suffer from Self-Control explains the use of this practice in temptation markets. In my model a single seller — therefore also referred to as a monopolist — offers his good to a tempted consumer that can be of two different types, high or low. The high (low) type values the good more (less) and is therefore tempted to buy more (less) of the good sold by the monopolist. Under some conditions, when temptation levels are private information, in equilibrium, the low type obtains a positive information rent ex-ante, and zero surplus ex-post, while the high type enjoys an information rent ex-post, and zero surplus ex-ante.

This last point contrasts with the traditional results of classical screening problems (e.g. Spence, 1977) and suggests an ex-ante role "swap" between high and low types. Consumers with high valuation of the good suffer from high temptation ex-post, which increases their willingness to pay. Ex-ante, however, they are capable of anticipating their self-control cost and are, therefore, willing to pay less. Exactly the opposite happens for consumers with lower valuations, since they do not suffer from milder levels of temptation ex-post. Hence, while ex-ante the low type is the "best" type in the market for the seller — i.e. the one with the highest willingness to pay — ex-post is the high type to be the most valuable customer to sell to. Therefore, when discriminating, the seller leaves positive surplus to the low type ex-ante and to the high type ex-post. This result is discussed extensively in Section 4.4.2.

The paper is organised as follows. In Section 4.2, I position the paper in the existing literature. In Section 4.3, I describe the model. In Section 4.4, I solve the model looking for a perfect Bayesian-equilibrium. In Section 4.5, I provide some comparative statics. A brief conclusion is present in Section 5. All proofs are relegated to the Appendix. Also in the Appendix are extensions and alternative approaches.

4.2 Related Literature

This paper contributes to the literature on temptation models. In particular, I model self-control preferences à la Gul and Pesendorfer (2001) which I describe in the following Section.

The economics literature has studied temptation and self-control in several different frameworks, some of which use a less general multi-self model approach. Broadly, **?** introduce a portion of näive consumers in temptation markets and show that the seller would be willing to educate consumers. Kumru and Thanopoulos (2008) use self-control preferences to study social security systems. Galperti (2015) uses a multi-self model to explain the trade-off between commitment and flexibility in contracts offered by a seller to a consumer with dynamically consistent or inconsistent preferences; he shows how the low type (the consistent one) enjoys an information rent. Christensen and Nafziger (2016) study the optimal packaging of "sin" goods in the presence of consumers that suffer from temptation. In Foschi (2015), I study how consumers with self-control problems may provide a justification for the existence of loyalty schemes in the retailing industry.

The closest papers of this literature to the present one are Esteban and Miyagawa (2005, 2006) and Esteban, Miyagawa, and Schum (2007). The first studies optimal contracting of a monopolist who sells a good to a tempted consumer with private information on his own level of temptation (his type). It shows how the monopolist can replicate first best by offering two separate menus and "decorating" the one designed for the less tempted consumer. They, however, implicitly assume that the monopolist can perfectly commit to specific menus of offers, and that she is unable to change them once the consumer is "in the store". By doing so, they allow the monopolist to set different menus for different customers. Once this assumption is dropped, the monopolist sets a single menu of offers designed according to the ex-post utility. I show how, if this is the case, the result of Esteban and Miyagawa (2005) does not hold any longer. Esteban and Miyagawa (2006) and Esteban, Miyagawa, and Schum (2007) extend the model to, respectively, a market with perfect competition and a continuum of types.

Finally, this paper contributes also to the literature on solutions to the Diamond Paradox (Diamond, 1971). When consumers have to pay a search cost in order to acquire information on the price of a good, firms can create a hold-up problem. Once a consumer is in a store, his willingness to pay raises by the search cost he has to bear if he were to look for the same good in another store. Firms exploit this hold-up problem and raise the price. Diamond (1971) shows that this creates an upward thrust on the equilibrium price that, eventually, reaches the one of joint profit maximisation. Consumers, however, anticipate the firms' behaviour and decide not to "search" for the good in the first place. This leads to complete market break-down (and the paradox). I show how, in temptation models without commitment, the hold-up problem is endogenous and defined by the level of temptation that afflicts the consumer.

One of the many solutions to this paradox studied in the literature is for the seller to commit to a particular price format (Wernerfelt, 1994; Anderson and Renault, 2006).[7] If commitment is impossible, the easiest way for the seller to attract the consumer in the store is to compensate him for the search cost. In this paper I follow a similar logic. The consumer suffers from temptation ex-post. When he enters the store, the monopolist may exploit his temptation and his higher willingness to pay. In period 1, outside the store, the consumer anticipates this behaviour and does not enter the store. Hence, the seller, being unable to commit, compensates the consumer for his negative ex-ante utility by means of a negative entry fee, i.e, an entry bonus.

4.3 The Model

A monopolist (she) sells a good to a tempted consumer (he) in her store. She posts a menu M of offers $x = (t, q) \in \mathbb{R}^2_+$ where t is the transfer the consumer has to make in order to acquire quantity (or quality) q of the good.

There are two periods; in period 1, the *ex-ante stage*, the consumer is "outside of the store". In this stage the monopolist sets an entry fee $F \in \mathbb{R}$. Transfer F takes place if and only if the consumer decides to "enter the store". Given F, which can be positive or negative, the consumer decides whether to enter or not. If he does not enter, the game ends. As in standard mechanism design problems, I normalise the consumer's payoff from the outside option to zero. Hence, if the consumer does not enter the store, both he and the monopolist obtain zero payoff.

If the consumer enters in period 1 the game continues in the *ex-post stage* (period 2). In this stage the monopolist sets a menu M of offers and the consumer chooses which $x \in M$ to buy. Following the existing literature discussed in Section

[7] A few more examples: Varian (1980) assumes the presence of temporal dispersion of information on prices; Burdett and Judd (1983) introduce "noisy" search — which means that consumers may learn two, or more, prices every time they search —; Stahl (1989) assumes the presence of fully informed consumers; Anderson and Renault (1999) establish the relationship between preferences for product differentiation and searching cost; Anderson and Renault (2006) introduce advertisement as a form of partial commitment that the seller can use to disclose an optimal amount of information; Rhodes (2014) builds on Anderson and Renault (2006) but considers the case of multi product retailing where the monopolist creates a "low price image" of himself by advertising low price products only.

4.2, I assume that the monopolist cannot prevent the consumer from leaving the store having bought nothing. That is, offer $0=(0,0)$ is always in the menu. Once the consumer has chosen an offer, payoffs are realised and the game ends.

The main difference between this paper and the existing literature on temptation models is that the monopolist cannot commit to a specific menu ex-ante and, therefore, sets M ex-post.

The consumer's preferences follow Gul and Pesendorfer (2001). He is affected by temptation in the second period, when choosing the offer from the menu, but he is able to anticipate this in period 1 when he is deciding whether or not to enter the store. Therefore, the decision to enter the store or not in the ex-ante stage depends crucially on the menu the monopolist will set in the ex-post stage. For instance, ex-ante, the consumer might be willing to choose offer 0 ex-post but knows that, once inside, he will fall victim to temptation and buy offer $x \neq 0$ instead.

The consumer can be of two types: low (L), with probability β, and high (H). In the ex post stage the consumer chooses an offer from menu M according to:

$$\max_{x \in M}[U(x)+V_i(x)] \qquad i=H,L. \tag{4.1}$$

Function U is called the *commitment (net) utility* while function V is called the *temptation (net) utility*. To understand the difference between these two functions, consider U as the base utility that the individual obtains from consuming the good, free of temptation. Function V, instead, measures the impulses of the individual in period 2. Ex-post, the individual considers both his commitment and his temptation and makes the choice. These utilities are assumed to be quasi-linear and to differ in the scaling of q:

$$U(x)=u(q)-t \tag{4.2}$$

$$V_i(x)=v_i(q)-t \tag{4.3}$$

Functions U, V_H and V_L all satisfy the single crossing property.[8] The temptation

[8]Notice that, following the existing literature on temptation models, I assume the negative part of utility the consumer gets from paying tariff t is equal to the actual transfer t itself. This is true also for the temptation utility. This results in an ex-post utility $U(x)+V_i(x)=u(q)+v_i(q)-2t$ even if the transfer is made only once. This t cancels out in (4.4) below, if $\arg\max[U(x)+V_i(x)]=\arg\max V_i(x)$, since $W = u(q)+v_i(q)-2t-\big(v_i(q)-t\big) = u(q)-t$. The peculiarity of this approach is that, if

(gross) utility of the low type, v_L, values less, with respect to u, the quantity of each offer, while one of the high type, v_H, values it more. All functions u, v_L, v_H are increasing and concave in q. I assume, that they satisfy $u(0) = v_i(0) = 0$.[9] The direction of the temptation is therefore characterised by the slope of v_i relative to u. Following Esteban and Miyagawa (2005), I write $v_L \prec u$ to indicate that v_L is flatter than u and, therefore, I say that the low type is *downward* tempted. For high types instead, I say $v_H \succ u$ to indicate that v_H is steeper than u and, that the high type is *upward* tempted. Hence:

$$v_L \prec u \Longleftrightarrow \left.\frac{\partial v_L}{\partial q}\right|_{q=q'} < \left.\frac{\partial u}{\partial q}\right|_{q=q'} \forall q' \Longleftrightarrow \text{L is downward tempted}$$
$$v_H \succ u \Longleftrightarrow \left.\frac{\partial v_H}{\partial q}\right|_{q=q'} > \left.\frac{\partial u}{\partial q}\right|_{q=q'} \forall q' \Longleftrightarrow \text{H is upward tempted}$$

Clearly, $v_L \prec v_H$. Figure 4.1 illustrates this concept.

In the ex-ante stage, the consumer anticipates his ex-post decision and the menu the monopolist sets ex-post. Hence, Gul and Pesendorfer's preference representation implies an ex-ante utility of the type:

$$W_i(M, F) = \max_{x \in M}[U(x) + V_i(x)] - \max_{x \in M} V_i(x) - F \qquad i = H, L. \tag{4.4}$$

This is the utility the consumer obtains by accepting an entry fee F, anticipating that he will face menu M in the ex-post stage, i.e., the utility he gets by entering the store.[10] Notice that the first part is composed of the utility the consumer obtains in the ex-post stage minus the temptation utility that he is foregoing because he is exerting self-control effort. This is represented by the offer that would maximise his temptation utility, $\max_{x \in M} V_i(x)$.

To understand the intuition behind (4.4), consider the following example. A consumer on a diet is facing menu $M = \{s, h\}$ where s is a healthy salad and h is a (very tasty) hamburger. Suppose there is no entry fee. In this context, s is the offer that

$\arg\max[U(x) + V_i(x)] \neq \arg\max V_i(x)$, the price of an offer which is not chosen enters the utility function as well.

[9]In Appendix C.1, I study an alternative case and show that results are qualitatively similar.

[10]The entry fee F can also be included in U and V_i as an additional tariff, since its value does not depend on the menu.

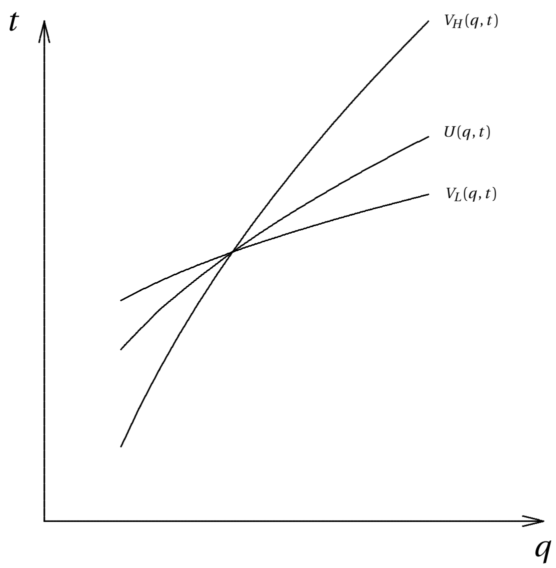

Figure 4.1: Single Crossing Property

In the figure above, general concave temptation (net) utility indifference curves are drawn. The single crossing property implies that they cross only once in $\mathbb{R}^2_+$.

maximises his ex-post utility since, when choosing in the ex-post stage, he considers his commitment to his diet. Offer h, instead, is tempting the consumer, i.e., it maximises his temptation utility. The difference $V(h) - V(s)$ is known as the *self-control cost*. Notice that (4.4) then becomes $W(M) = U(s) - [V(h) - V(s)]$. Therefore, if the self-control cost of choosing the salad exceeds the commitment utility, the consumer will not accept the menu in the ex-ante stage. This is because he understands that in order to obtain utility $U(s) + V(s)$ in period 2 he also has to incur a self-control cost that makes his ex-ante utility negative.[11]

Type i consumer enters the store if and only if his ex-ante participation constraint $\underline{PC}_i$ is satisfied:

$$W_i(M,F) \geq 0. \qquad (\underline{PC}_i)$$

The monopolist's ex-post profit is given by $\pi(x) = t - c(q)$. The function $c(q)$ is

[11]An alternative to this approach is a multi-self model (as in Strotz, 1956). However, as I show in Appendix C.9, a temptation model like the one considered here endogenises the change from a classical model with consumers's preferences $U(x) + V_i(x)$ to a multi-self model with self one's preferences $U(x)$ and self two's preferences $U(x) + V_i(x)$.

the total cost of production. It is assumed to be strictly increasing, convex in q and such that $c(0) = 0$. Following the existing literature, the monopolist faces no cost of adding an offer to a menu.

If there is perfect information about the consumer's type, in the ex-post stage the monopolist maximises her payoff given the consumer's ex-post participation constraint ($\overline{PC}_i$):

$$\max_{x_i} \pi(x_i) = \max_{x_i} \left[t_i - c(q_i) \right] \tag{4.5}$$

$$\text{s.t.} \quad [U(x_i) + V_i(x_i)] \geq 0 \tag{$\overline{PC}_i$}$$

The solution to (4.5) s.t. ($\overline{PC}_i$) is the first-best offer, written x_i^*. A first best optimal menu set by the monopolist in the ex-post stage is, therefore, $M_i^* = \{0, x_i^*\}$.

In the ex-ante stage, the monopolist's profit are given by $\Pi(M, F) = \pi(x) + F$. Therefore, she sets the maximum possible F such that $\underline{PC}_i$ binds:

$$F^* = \{F \mid W_i(M_i^*, F) = 0\}. \tag{4.6}$$

Now consider the case of asymmetric information. In particular, suppose that the monopolist cannot observe the consumer's type, even when the latter is in the store. As mentioned above she knows that the probability the consumer is of low type is β (this constitutes her prior). Unless the monopolist is able to separate types ex-ante, in the ex-post stage, she faces a classical second-degree price discrimination problem (Spence, 1977):

$$\max_{M} \pi(M) = \max_{M} \left[\pi(x_L)\beta + \pi(x_H)(1-\beta) \right] \tag{4.7}$$

$$\text{s.t.} \quad U(x_H) + V_H(x_H) \geq 0 \tag{$\overline{PC}_H$}$$

$$U(x_L) + V_L(x_L) \geq 0 \tag{$\overline{PC}_L$}$$

$$U(x_L) + V_L(x_L) \geq U(x_H) + V_L(x_H) \tag{$\overline{IC}_L$}$$

$$U(x_H) + V_H(x_H) \geq U(x_L) + V_H(x_L), \tag{$\overline{IC}_H$}$$

where M is the menu of all offers set ex-post. Incentive compatibility constraints

$\overline{IC}_L$ and $\overline{IC}_H$ are also introduced. They ensure that, ex-post, type i buys the offer designed for him and not the one set for type j. In the following section, I solve the model backwards, looking for the perfect Bayesian equilibrium.

4.4 Optimal Contracts

I now derive the optimal menu the monopolist sets ex-post and the optimal entry fee she charges ex-ante. Solving the game by backward induction, I first show what is best for the monopolist ex-post and then characterise the optimal entry fee.

4.4.1 Benchmark Case of Full Information

As a benchmark I solve the full information problem finding the first best menu and offers. Suppose the monopolist is fully informed about the consumer's type both ex-post and ex-ante. The ex-post problem then becomes:

$$\max_{x_i} \left[t_i - c(q_i) \right] \tag{4.8}$$

$$\text{s.t. } u(q_i) + v_i(q_i) - 2t_i \geq 0 \tag{$\overline{PC}_i$}$$

$$\text{for } i = L, H.$$

Given the solution to (4.8) the monopolist sets the optimal entry fee according to (4.6). Since full information is assumed, she is able to set two different entry fees, one for each type.

Proposition 4.1. *If the monopolist is capable of perfectly observing the type of the consumer, both types of consumer enter the store and obtain zero ex-ante and ex-post utility. The first best is characterised by an optimal offer $x_i = (q_i, t_i)$ and an ex-ante entry fee F_i for each type i where:*

$$t_i^* = \tfrac{1}{2}\left[u(q_i^*) + v_i(q_i^*) \right], \quad q_i^* = \left\{ q \,\middle|\, \tfrac{1}{2}\left[u'(q) + v_i'(q) \right] = c'(q) \right\}, \tag{4.9}$$

$$F_H^* = \tfrac{1}{2}\left[u(q_H^*) - v_H(q_H^*) \right] < 0, \quad F_L^* = 0. \tag{4.10}$$

It is easy to see that, in equilibrium, both types of consumer get $U(x_i^*)+V_i(x_i^*)=0$ and $W_i(\{0,x_i^*\},F_i^*)=0$. Notice that the high type is upward tempted, in equilibrium he buys more than he should, according to his ex-ante utility. Hence, in order to attract him into the store, the monopolist has to compensate him ex-ante with a negative entry fee.

As expected, both offers in the menu ensure that the marginal expected utility from consuming q_i^* is equal to the marginal cost of producing it.

4.4.2 Asymmetric Information

Assume now, instead, that the monopolist cannot observe the type of the consumer she faces. The model becomes a dynamic game of incomplete information.

At the start of the game, the monopolist is assumed to have a prior $\Pr[i=L]=\beta$, that she updates ex-post. I show, however, that, when focusing on pure strategies, updating is trivial.

Notice that the only tool the monopolist has to separate types ex-ante is the entry fee. Even if the monopolist were to set distinct entry fees, however, both types, being free to choose whichever they like, would enter choosing the lower of them, and without self-selecting themselves, making ex-ante separation impossible. This follows from $W(\cdot,F') > W(\cdot,F'')$ for all $F' < F''$. Hence, the only way to separate types ex-ante is by setting an entry fee such that only one type finds it optimal to enter the store, while the other stays out.

This depends on the assumption that the monopolist is not able to commit herself ex-ante to the menus she sets ex-post. If this assumption is dropped, the monopolist can "force" the consumer to buy from a specific menu when picking a given entry fee ex-ante.

Consider the situation where only the low type enters the store, while the high type stays out. In any equilibrium, the monopolist sets the first best offer given by (4.9) ex-post. Therefore, she can charge no entry fee. Notice, however, that, since $U(x_L^*)+V_L(x_L^*)=0$, then $U(x_L^*)+V_H(x_L^*)>0$. In other words, the high type obtains a positive utility from the offer designed for the low type. Hence, facing menu $\{0,x_L^*\}$, he chooses x_L. Given this, the high type enters the store and buys the offer designed for the low type. He obtains an ex-ante utility given by $W_H(\{0,x_L^*\},0)>0$. This would seem to imply that an equilibrium where the low type enters and the high type does not cannot happen. As shown, however, by Esteban and Miyagawa (2005), the monopolist can "decorate" the menu ex-post in order to tempt the high type and make his ex-ante utility negative. Here, I show one way of decorating the menu.

It is sufficient to show that there exists an offer z that tempts the high type only and makes his ex-ante utility W_H negative. This offer should not be chosen by either of the two types ex-post so as not to affect directly the monopolist's profits. Let $M' = \{0, x_L^*, z\}$ be the decorated menu the monopolist sets ex-post, then z has to satisfy:

$$U(x_L^*) + V_L(x_L^*) \geq U(z) + V_L(z) \tag{4.11}$$

$$V_L(x_L^*) \geq V_L(z) \tag{4.12}$$

$$U(x_L^*) + V_H(x_L^*) \geq U(z) + V_H(z) \tag{4.13}$$

$$V_H(x_L^*) \leq V_H(z) \tag{4.14}$$

$$V_H(z) - V_H(x_L^*) \geq U(x_L^*). \tag{4.15}$$

The first two constraints say that z does not affect low type's behaviour once he is inside the store. If they do not hold, the utility obtained by the low type is affected by the presence of z both ex-ante, if (4.12) fails and the presence of z increases the self-control cost, and ex-post, if (4.11) fails and he chooses z over x_L^*, moving away from equilibrium. The second two say that z tempts the high type when he chooses x_L^* from M' and the last one says that the self control cost of choosing x_L^* from M' is too high for the high type and, therefore, $W_H(M', 0) < 0$.

To see that such an offer exists consider the general case in Figure 4.2.

In the Figure, utility increases towards the bottom right of the graph. Offer z is above $U(x_L) + V_L(x_L) = 0$ and on the right of x_L. Therefore, z is also above $V_L(x_L)$ hence (4.11) and (4.12) hold. Also, x_L and z lie on the same $U + V_H$ indifference curve making (4.13) bind. It is easy to see that (4.14) holds. Finally, since $U(z) = 0$, $U(z) + V_H(z) = V_H(z)$. Hence, $U(x_L) + V_H(x_L) = V_H(z)$ and $V_H(z) - V_H(x_L) = U(x_L)$, which shows that (4.15) binds.

Setting menu M', the monopolist knows that when charging a zero entry fee only the low type enters the store, since $W_L(M', 0) = 0$.

Notice that, because of the lack of commitment from the monopolist to the menu she sets in period 2, this equilibrium depends crucially on the assumption that adding

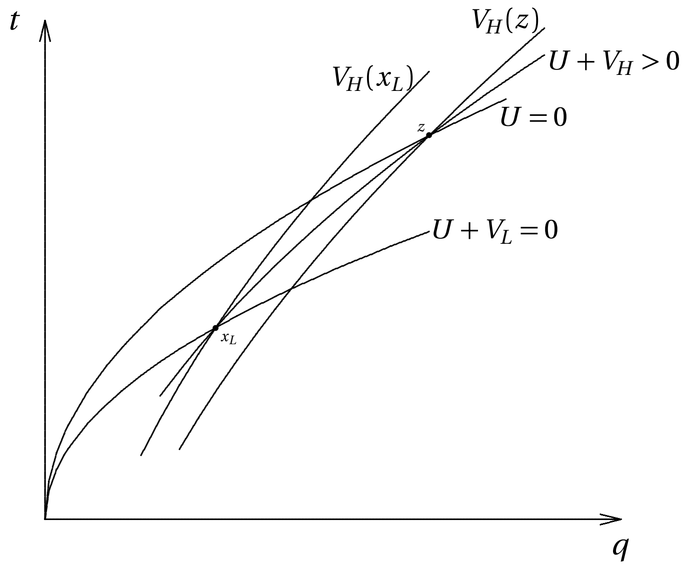

Figure 4.2: Decorated Menu

A "decorated" menu to separate types ex-ante. With a menu $M' = \{0, x_L^*, z\}$ the low type is willing to enter the store while the high type stays out. The presence of the offer z makes the high type's ex-ante utility negative, while not affecting the low type's ex-ante and ex-post utility.

an offer to the menu is costless for the monopolist. If this is not the case, then ex-post the monopolist has no incentive to add an offer like the z described. Ex-ante, she knows that she will not decorate the menu ex-post and cannot, therefore, exclude the high type optimally.[12]

The following result highlights ex-ante profits for this case.

Lemma 4.2. *When the monopolist excludes the high type ex-ante, by setting no entry fee and menu M' ex-post, she earns ex-ante profits:*

$$\Pi^{EH} = \pi(x_L^*)\beta + 0 = \left[\tfrac{1}{2}\left[u(q_L^*) + v_L(q_L^*)\right] - c(q_L^*)\right]\beta \tag{4.16}$$

where EH stands for "exclude the high type".

Consider, now, the situation where only the high type enters the store, while the low type stays out. In any equilibrium, the monopolist is now certain to face a high

[12]This equilibrium does not play a crucial role in the main results of the paper since no entry fee is set and no type obtains positive ex-ante or ex-post utility. Hence, the main message remains unaltered by the presence of this assumption.

type consumer. Hence, she sets the first best offer given by (4.9) ex-post. Therefore, she also has to charge a negative entry fee to induce the high type to enter the store, as in (4.10). Notice, however, that, since $U(x_H^*)+V_H(x_H^*)=0$, then $U(x_H^*)+V_L(x_H^*)<0$. In other words, the low type obtains a negative utility from the offer designed for the high type. Hence, facing menu $\{0, x_H^*\}$, he chooses *0*. Given this, the low type enters the store, obtains the entry bonus, and buys nothing from the store. He obtains an ex-ante utility given by $W_L(\{0, x_H^*\}, F_H^*) > 0$. This implies that there can be no equilibrium where the high type enters and the low type does not.

It is easy to see that the menu cannot be decorated to exclude the low type ex-ante. This is because he suffers from downward temptation, $v_H \succ u \succ v_L$. Hence, there is no way to tempt him without also tempting the high type. Consider offer $x_j=(q_j,t_j)$, since $u \succ v_L$, then $\left[u(q_j)+v_L(q_j)\right] \geq v_L(q_j) \geq 0$ for all $q_j \in \mathbb{R}_+$. Given that in order for this offer to tempt the low type it has to be that $v_L(q_j) > t_j$, the ex-ante utility he obtains from a menu that contains offer *0* is at least $u(q_j)-t_j>0$. This makes it impossible to decorate the menu $\{0, x_H^*\}$ with an offer x_j in such a way that the low type does not find it optimal to enter.

Given this, only one case remains: the one where consumer's types do not separate ex-ante and they both enter (i.e. pooling ex-ante).[13] Since self-selection does not take place ex-ante, the posterior beliefs of the monopolist are unchanged and she believes that the consumer is of low type with probability β. In this case, the problem she solves is a classical second-degree price discrimination, as in (4.7). Hence, she has three options: (i) exclude low types from the market, offering ex-post only the first best offer for high types x_H^*, (ii) set a single offer x^P that both types are willing to buy — pooling —, (iii) set a separating menu $M^S=\{0, x_L^S, x_H^S\}$ that induces types to self-select themselves ex-post. I first work out the optimal contracts in all three cases and then compare ex-post profits to obtain the equilibrium of the ex-post subgame.

Case (i) needs no computations since the monopolist sets a single offer, and therefore a menu $\{0, x_H^*\}$ — recall, from above, that the low type chooses *0* from $\{0, x_H^*\}$. By doing so, the monopolist obtains ex-post profits:

$$\pi^{EL}=\pi(x_H^*)(1-\beta)=\left[\tfrac{1}{2}\left[u(q_H^*)+v_H(q_H^*)\right]-c(q_H^*)\right]\left(1-\beta\right). \tag{4.17}$$

[13]The monopolist could, of course, also exclude them both ex-ante. In this case the game would end and the monopolist ex-ante profits would be equal to zero.

It is important to stress that these are *ex-post* profits, and the entry fee plays no role in the ex-post subgame.

Case (ii) is also easy to derive. The monopolist drops IC_H and IC_L from (4.7) and sets the offer that solves:

$$\begin{aligned} &\max_{x^P} \pi(x^P) = t - c(q) \\ &\text{s.t. } U(x^P) + V_L(x^P) \geq 0 && (PC_L) \\ &\quad\;\; U(x^P) + V_H(x^P) \geq 0. && (PC_H) \end{aligned}$$

Notice that PC_L is binding at x_L^* and that x_L^* maximises $\pi(x)$ subject to PC_L. Also, as in the general case — see Lemma 4.3—, if PC_L binds, PC_H is slack. Therefore, $x^P = x_L^*$ solves the pooling problem. Ex-post profits of pooling are therefore given by:

$$\pi^P = \pi(x_L^*) = \tfrac{1}{2}\left[u(q_L^*) + v_L(q_L^*)\right] - c(q_L^*) \tag{4.18}$$

Case (iii) requires some computations. If the monopolist wants to separate types ex-post, she solves the entire problem (4.7). In the next Lemma, I show how two of the four constraints of problem (4.7) can be ignored. This follows from the well know solution of second degree price discrimination problems.

Lemma 4.3. *When types are private information and both types enter the store ex-ante, the monopolists sets the optimal contracts according to (4.7), where the participation constraints of the low type and the incentive compatibility constraint of the high type are binding. Other constraints are slack.*

The optimal ex-post menu of case (iii) then solves:

$$\max_M \Pi(M) = \max_M \left[\pi(x_L)\beta + \pi(x_H)(1-\beta)\right] \tag{4.19}$$

$$U(x_L) + V_L(x_L) \geq 0 \qquad (PC_L)$$

$$U(x_H) + V_H(x_H) \geq U(x_L) + V_H(x_L) \qquad (IC_H).$$

which yields as a solution:

$$M^S = \{0, x_L^S, x_H^S\} \quad \text{where} \quad x_i^S = (q_i^S, t_i^S) \quad i = H, L \tag{4.20}$$

$$t_H^S = \tfrac{1}{2}\left[u(q_H^S) + v_H(q_H^S) - v_H(q_L^S) + v_L(q_L^S)\right] \tag{4.21}$$

$$q_H^S : \tfrac{1}{2}\left[u'(q) + v_H'(q)\right] = c'(q) \tag{4.22}$$

$$t_L^S = \tfrac{1}{2}\left[u(q_L^S) + v_L(q_L^S)\right] \tag{4.23}$$

$$q_L^S : \tfrac{1}{2\beta}\left[(u'(q) + v_H'(q))\beta - (v_H'(q) - v_L'(q))\right] = c'(q). \tag{4.24}$$

Lemma 4.5 below shows how (4.20)—(4.24) exhibit no distortion at the top and leave no surplus to the low type, as expected. First of all, however, notice that optimal separation is only feasible if $q_L^S \geq 0$, which is discussed by the following Lemma.

Lemma 4.4. *Ex-post optimal separation is possible, i.e., $q_L^S \geq 0$, if and only if:*

$$\beta \geq \underline{\beta} \equiv \frac{v_H'(q_L^S) - v_L'(q_L^S)}{u'(q_L^S) + v_H'(q_L^S)} \tag{4.25}$$

Lemma 4.4 shows that separation becomes possible only when the probability that the consumer is indeed a low type is high enough. The intuition is quite simple: if the consumer is almost certainly a high type, the monopolist decreases the positive surplus left for him. Hence, in order for separation to be possible, she has to sell a smaller quantity to the low type. If β is particularly low, the optimal q_L^S turns negative, making separation impossible.

I now move to study the comparison between this sub-game equilibrium candidate and the one of first best. The following Lemma shows that $q_H^* = q_H^S > q_L^* > q_L^S$:

Lemma 4.5. *When the monopolists wants to separate types ex-post serving both of them, she sets offers $x_i^S = (q_i^S, t_i^S)$, $i = H, L$, where:*

$$q_H^* = q_H^S > q_L^* \geq q_L^S$$

and q_i^* is the quantity of the first best offer designed for type i.

Lemma 4.5 shows that the quantity sold to the high type is unchanged from first best — efficiency at the top — while the quantity offered to the low type is lower — inefficiency at the bottom. On top of this, notice, from (4.21), that the tariff the high type pays is lower than the one paid in first best. This ensures the high type a positive (ex-post) surplus whilst the low type gets zero surplus. Hence, the second period separation outcome satisfies the classical properties of second degree price discrimination models.

Given M^S, ex-post profits from separation are:

$$\begin{aligned}\pi^S &= \pi(M^S)\\ &= \pi(x_L^S)\beta + \pi(x_H^S)\big(1-\beta\big)\\ &= \big[\tfrac{1}{2}\big[u(q_L^S) + v_L(q_L^S)\big] - c(q_L^S)\big]\beta\\ &\quad + \big[\tfrac{1}{2}\big[u(q_H^S) + v_H(q_H^S) - v_H(q_L^S) + v_L(q_L^S)\big] - c(q_H^S)\big]\big(1-\beta\big). \end{aligned} \tag{4.26}$$

In order to understand what is best for the monopolist ex-post when both types are willing to enter, I compare (4.17), (4.18) and (4.26) and look for the conditions that solve $\max\{\pi^S, \pi^P, \pi^{EL}\}$. First, I show in the following Lemma that pooling is never an equilibrium.

Lemma 4.6. *There exist an offer $\tilde{x} = (\tilde{q}, \tilde{t})$ that is chosen by the high type, but not by the low type, in the menu $\{0, x^P, \tilde{x}\}$ and such that $\pi(x^P)\beta + \pi(\tilde{x})(1-\beta) > \pi^P$. Therefore, there exists no equilibrium where the monopolist sets a pooling menu ex-post.*

I prove this in the appendix using a similar argument to Rothschild and Stiglitz (1976). What Lemma 4.6 says is that there always exist a, non optimal, separation menu $\{0, x^P, \tilde{x}\}$ that makes consumers self-select and grants the monopolist higher ex-post profits than the pooling one. Given this result, when both types enter the store, in equilibrium, the monopolist either excludes the low type or she separates types selling offers x_L^S and x_H^S according to condition (4.25).

To see this consider Figure 4.3 below.[14]

In the Figure, I plot the three different ex-post profits as a function of β. It is easy to see that as long as separation is possible, the ex-post profits it grants are the highest that the monopolist can obtain. When $\beta = \underline{\beta}$ the profits from separation equal those from the exclusion of the low type, and when $\beta = 1$ they equal those from pooling. Hence, when condition (4.25) holds, ex-post the monopolist sells different positive quantities of the good to different types. When it fails, she excludes the low type from the market and only sells the first best offer to the high type.

[14] In Appendix C.8, I present Claims C.1-C.4 to describe the shape of the profits depicted in the Figure.

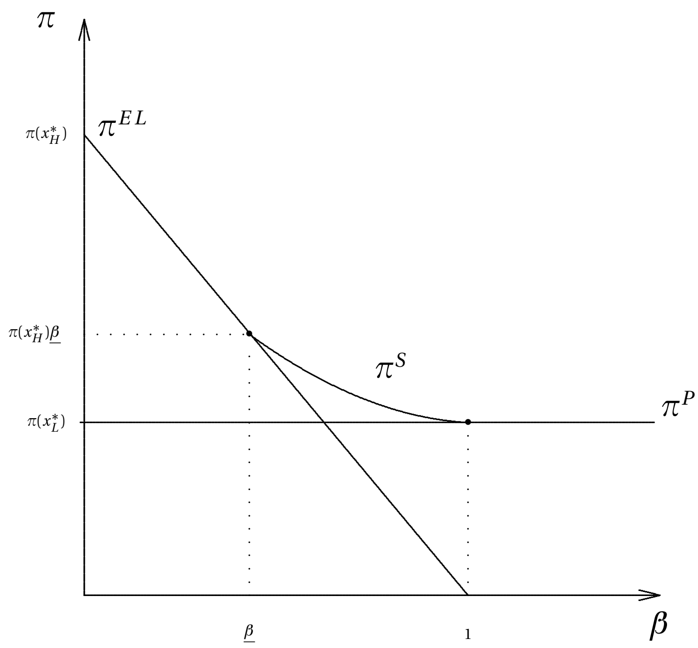

Figure 4.3: Ex-post Equilibrium

Cost and Gains of serving the low type. On the x-axis is the probability the consumer is a low type, β, while ex-post profits are measured on the y-axis. Following of Claims C.1-C.4 in Appendix C.8, π^S takes the form shown. The curve is not plotted for values of $\beta < \underline{\beta}$ since separation is not feasible when (4.25) fails. The other profits are linear in β.

Given all the above, now I move to the identification of the optimal ex-ante choice. First of all, I derive the optimal entry fees for every possible ex-post menu. Recall that the monopolist ex-ante cannot distinguish between types and she is forced to charge a single entry fee. Also, the case in which the monopolist excludes high types ex-ante is already described above.

I start by considering the case in which the monopolist excludes low types ex-post. In order to attract high types to the store, the monopolist has to set an entry fee as in (4.10). Given this, both types will enter ex-ante. The high type consumer buys x_H^* in the ex-post stage while the low type simply walks out of the store, i.e., chooses *0*. Notice that the monopolist is paying the an entry bonus ($F_H^* < 0$) to *both* types only to have the high type inside the store buying the first best offer. Hence, when, and if, this case is an equilibrium, the low type gets a positive ex-ante surplus and a zero ex-post utility while the high type obtains zero surplus both ex-ante and ex-post. The following result shows ex-ante profits in this case.

Lemma 4.7. *When the monopolist excludes the low type ex-post, by setting menu* $\{0, x_H^*\}$*, she earns ex-ante profits:*

$$\begin{aligned}\Pi^{EL} &= \pi^{EL} + F_H^* = \pi(x_H^*)(1-\beta) + F_H^* \\ &= u(q_H^*) - c(q_H^*) - \left[\tfrac{1}{2}\left[u(q_H^*) + v_H(q_H^*)\right] - c(q_H^*)\right]\beta \end{aligned} \tag{4.27}$$

where EL stands for "exclude low type".

Before moving to the case of no exclusion of types, ex-post or ex-ante, it is important to stress the connection between this equilibrium and the one where the monopolist excludes the high type ex-ante. Generally, in classical price discrimination problems (e.g. Spence, 1977), the monopolist may find it optimal to exclude from the market the type of consumer with the lowest valuation of the good (i.e. the low type). As already argued in this paper, however, because of the self-control problem of the consumer, the role of types is inverted from one stage to the other. Ex-ante, the high type anticipates a stronger self-control problem which decreases his ex-ante willingness to pay. Ex-post, the low type is less tempted than the high type, and willing to pay less than the latter for the same quantity. Hence, ex-ante is the low type to be the *most valuable* consumer for the monopolist, while ex-post this role belongs to the high type. This is the reason why the monopolist may find it optimal to exclude the high type ex-ante or the low type ex-post. Finally, notice that while I highlight that this role "swap" is a property of models with self-control preferences, the case of exclusion of the high type is original to this paper, and it follows from the inability of the monopolist to commit ex-ante to the ex-post menu. If he could commit, as shown by Esteban and Miyagawa (2005), he would decorate the menu for the low type and offer a second, separate, menu for the high type, replicating perfectly the first-best equilibrium.

Returning to the equilibrium selection, when the monopolist wants to separate types optimally ex-post, it is easy to see that $W_L(M^S, 0) = 0$ and $W_H(M^S, 0) = U(x_H^S)$.

The ex-ante utility of the high type in this case is not always positive:

$$U(x_H^S) \geq 0 \Longleftrightarrow v_H(q_L^S) - v_L(q_L^S) \geq v_H(q_H^S) - u(q_H^S) \tag{4.28}$$

which is not generally satisfied. However, when $v_H \longrightarrow u$ the RHS of (4.28) goes to 0 while the LHS remains positive. Hence as the temptation of the high type disappears, the LHS of the equation becomes relatively larger than the RHS. In other words, the quantity the high type buys ex-post gets closer to what is optimal according to his ex-ante preferences; hence the decreasing need to compensate him ex-ante.

Given this, when the monopolist separates types ex-post, she sets an entry fee $F^S = min\{U(x_H^S), 0\}$. Recall that if $F \leq 0$ both types accept it when entering the store. Therefore, the ex-ante profits of ex-post separation are described in the following result:

Lemma 4.8. *When the monopolist separates types ex-post, by setting menu M^S, she earns ex-ante profits:*

$$\begin{aligned}\Pi^S =& \pi^S + F^S \\ =& \pi(x_L^S)\beta + \pi(x_H^S)\big(1-\beta\big) + min\{U(x_H^S), 0\} \\ =& \big(1-\beta\big)\big[\tfrac{1}{2}\big(u(q_H^S) + v_H(q_H^S) - v_H(q_L^S)\big) - c(q_H^S)\big] \\ &+ \tfrac{1}{2}v_L(q_L^S) + \big[\tfrac{1}{2}u(q_L^S) - c(q_L^S)\big]\beta + min\{U(x_H^S), 0\}.\end{aligned} \tag{4.29}$$

where S stands for "separation".

The perfect Bayesian equilibrium of the game depends on whether (4.16) is larger than the profits of letting both types in the store. These latter depend, as described above, on condition (4.25). Proposition 4.9 describes the equilibrium of the game.

Proposition 4.9. *In the perfect Bayesian equilibrium of the game with asymmetric information, the monopolist charges an ex-ante entry fee F and an ex-post menu of offers M, where:*

(i) if optimal separation is not possible, there exists a β_{EH}^{EL} such that if:

$$\beta \leq \beta_{EH}^{EL} \tag{4.30}$$

$F = F_H^*$ and $M = \{0, x_H^*\}$, both types enter the store accepting the entry bonus ex-ante, the high type buys x_H^* ex-post while the low type chooses 0. If $\beta > \beta_{EH}^{EL}$, then $F = 0$ and $M = M' = \{0, x_L^*, z\}$, the monopolist charges a zero entry fee ex-ante, only the low type enters the store ex-ante and buys x_L^* ex-post.

(ii) if optimal separation is possible, there exists a β^S_{EH} *such that if:*

$$\beta \leq \beta^S_{EH} \tag{4.31}$$

$F = F^S$ *and* $M = M^S = \{0, x^S_H, x^S_L\}$*, both types enter the store accepting the entry bonus ex-ante, the high type consumer buys* x^S_H *ex-post while the low type buys* x^S_L*. If* $\beta \leq \beta^S_{EH}$*, then* $F = 0$ *and* $M = M' = \{0, x^*_L, z\}$ *as in* (i).

The perfect Bayesian Equilibrium is derived by considering the optimal ex-post menu set by the monopolist when both types enter the store (according to condition (4.25)), and then comparing the ex-ante profit of that with the one of excluding the high type ex-ante.

The Proposition shows that an equilibrium exists for all values of $\beta \in [0, 1]$, and it is summarised by Figure 4.4 below.

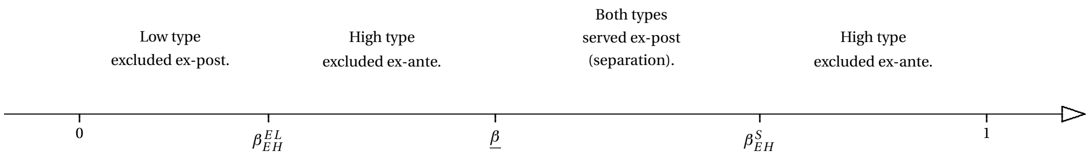

Figure 4.4: Ex-ante Equilibrium

The equilibrium of the game for all values of β. The low type is excluded ex-post (but enters ex-ante) only when β is low. The high type is excluded ex-ante either because separation ex-post is not feasible or because β is too high.

When the ex-post optimal behaviour is to exclude the low type, then, if (4.30) holds, the game has a separating equilibrium where the low type is excluded ex-post. If (4.30) does not hold then the ex-ante profit of excluding the high type ex-ante is higher than Π^{EL} and therefore the equilibrium of the game is a separating one where the high type is excluded ex-ante. Similarly, when the sub-game equilibrium is to separate types selling positive quantities to both of them, if condition (4.31) holds then both types are attracted in ex-ante. If it does not hold, the separating equilibrium where the high type is excluded ex-ante takes place.

One of the main messages of the paper is present in Proposition 4.9 and it is worth further analysis. There exists a parameter space where the monopolist is optimally paying a positive amount to both types ex-ante but selling only to the high type ex-post. This implies two things. First, the presence of consumers who suffer from self-control problems generates an extra cost on the monopolist when she decided to exclude low types from the market. In other words, she has to "pay a fee" to low types consumers in order to sell at first best to high types. Second, it implies that under some conditions the low type obtains an information rent when information becomes asymmetric. This second point generates the following Corollary.

Corollary 4.10. *There exists values for parameter β and temptation levels v_H and v_L such that the low type consumer is better off when information is asymmetric than when types are common knowledge.*

This result holds under two circumstances: first, and more obviously, when both types are induced in the store but the low type is excluded ex-post; second when ex-post separation with a negative entry fee ex-ante takes place. Notice that, here, the low type consumer obtains a zero ex-post surplus and a positive ex-ante surplus, precisely the opposite of what the high type obtains. This generates a discussion conserning the "role" that high and low types play in temptation models with self-control preferences.

In classical problems, where consumers do not suffer from self-control problems, the high type is usually considered to be the *best* type. He usually has a higher willingness-to-pay/ability, or induces the seller to face less risk. In temptation models with self-control preferences, this is only partially true. Ex-ante, in fact, the high type is no longer the best type in the market. His high willingness to pay for the good becomes a burden for him. Having a high valuation of the good (high temptation) now means he suffers from stronger self-control problems, and a lower ability to control his actions ex-post. Hence, while ex-post the roles are clear and the high type is the consumer with the highest valuation of the good, ex-ante these roles are reversed. The high type now becomes the type with the strongest self-control problem while the low type can control himself and bears a lower self-control cost. Hence, the "swap" in welfare results: the low type obtains positive rent ex-ante, while the high type obtains the ex-post.[15]

[15] Or, indeed, none at all, as described above.

4.5 Comparative Statics

The main purpose of this paper is to show the effect of asymmetric information when the level of temptation of the consumer is private information and the monopolist cannot commit herself to the menus she sets ex-post. Proposition 2 characterises the equilibrium of this game. The qualitative features of the equilibrium depend on the probability the consumer is a low type, β, and on the temptation level of both types of consumer, i.e, the relative slope of v_L and v_H. Given the level of β, v'_H and v'_L, there exists a unique equilibrium in pure strategies. In this section, I focus on the effect of these variables on conditions (4.30) and (4.31), which, ultimately, identify the equilibrium of the game. Notice that both conditions are implicit in β since the RHS depends on it also.

Condition (4.30) compares the monopolist's ex-ante profits of excluding the high type ex-ante with those of excluding the low type ex-post. β must be lower than β^{EL}_{EH} in order for the monopolist to be willing to exclude low types ex-post. When β increases, the probability of paying the consumer to enter the store only to have him choose *0* ex-post increases. Therefore, excluding the low type ex-post becomes less appealing to the monopolist.

The effect of the temptation level is, unexpectedly, symmetric in (4.30). If v'_L rises, β^{EL}_{EH} decreases, hence the monopolist is less willing to exclude the low type ex-post. The intuition is straightforward: since the low type is now more tempted, the monopolist can exploit his temptation more and excluding him ex-post becomes less attractive.

To understand why a rise in the temptation level of the high type has the same effect, consider the following. As v'_H rises, q^*_H rises and β^{EL}_{EH} decreases. The condition becomes tighter and will, eventually, fail. There is a clear explanation for this. Notice that the profits compared here are *ex-ante* profits and, therefore, the entry fee plays an important role in the condition. The higher is the level of temptation of the high type, the higher is the entry bonus that the monopolist has to offer the consumer to attract him into the store. As condition (4.30) shows, if the temptation level is high, this effect is stronger than the incentive to attract the high type into the store and extract all his surplus. When v'_H is "too high", it becomes too costly to attract the high type and the monopolist finds it optimal to exclude him ex-ante.

Condition (4.31) compares the ex-ante profits of excluding high types ex-ante with the ones of separation. An examination of this conditions yields somewhat more ambiguous comparative statics. The reason for this is that changes in all the crucial variables of the model have "ex-ante effects" and "ex-post effects" which are in sharp opposition.

First, consider a rise in the parameter β. This has two opposite effects. On the one hand, excluding the high type becomes more attractive since the consumer is more likely to be of a low type. Moreover, ex-post separation becomes less attractive, since the surplus granted to the high type ex-post is higher when β is higher.[16] However, a third effect arises if $U(x_H^S) < 0$. Notice that in this case $F^S = U(x_H^S)$ and that $\left|U(x_H^S)\right|$ decreases in β. Therefore, the "ex-ante cost of separation", i.e. the entry bonus that the monopolist has to offer consumers, decreases in β, making separation more attractive. None of these effects dominates the other for all possible levels of temptations. Hence, a change in β can, eventually, make condition (4.31) hold or fail.

Consider, now, the case of a rise in v'_H. On the one hand, the monopolist is less willing to exclude the high type ex-ante in order to exploit his higher temptation ex-post. On the other, a higher temptation level implies a higher entry bonus ex-ante paid to all types that enter. Hence, a rise in v'_H has both positive and negative effects on the profits of ex-post separation. Similar is the intuition behind a rise in v'_L. On the one hand, the difference in temptation levels between types is lower and excluding the high type becomes more attractive. On the other, the cost of separation (intended as the surplus granted to the high type ex-post) decreases making separation more attractive. Hence, a rise in v'_L has positive effects on both the profits of excluding the high type ex-ante and those of separating ex-post.

4.6 Conclusions

I construct a two period model where the consumer suffers from self-control problems (Gul and Pesendorfer, 2001). I show the effect of asymmetric information when the consumer level of temptation is heterogenous. Differently from the existing literature, I relax the general assumption that the monopolist can commit ex-ante to the menus she sets ex-post. In accordance with recent trends in online temptation

[16] To see this, notice that the ex post surplus is given by $t_H^* - t_H^S = \frac{1}{2}\left[v_H(q_L^S) - v_L(q_L^S)\right] > 0$, which is increasing in β.

markets sellers' behaviour, I show how the monopolist solves the problem of commitment by charging negative entry fees ex-ante and attracting consumers into the store in order to exploit their ex-post temptation in period 2.

The unique (pure) equilibrium of the game can take three different forms. First, the monopolist may exclude the high type ex-ante, via "decoration" of the menu, and sell only the first-best offer to the low type ex-post. Second, she may attract both consumers in the store ex-ante and then separate them ex-post. Last, and most interesting, she can attract both consumers in the store ex-ante but exclude the low type ex-post, selling the first best quantity to the high type. This case arises if the temptation of the high type is large enough to make the exclusion of the low type attractive, but not so high that the cost of compensating him ex-ante becomes overwhelming.

The paper has two main messages. First, I show how, under some conditions, in equilibrium, consumers with low level of temptation are better off when information is asymmetric than when the consumer's type is common knowledge. This gives rise to a "role swap" between high and low types discussed in the paper. Second, I show how the presence of tempted consumers can motivate the existence of entry bonuses in temptation markets even when not all consumers buy a product once in the store. The monopolist wants to exploit the high type's temptation ex-post and she is, therefore, forced to attract both types of consumer into the store with an entry bonus.

Chapter 5

Conclusion

This thesis has studied how the results of classical contract theory are affected by the presence of individuals that suffer from different behavioural biases. The Chapters presented three self-contained papers that studied different frameworks of the classical principal/agent model.

In Chapter 2, agents played the role of workers, who were unaware of their innate abilities and formed biased expectations about them, i.e. they were naïve. This bias was assumed to depend on the same abilities they tried to estimate. The principal, i.e. the employer, instead, had unbiased beliefs about workers' abilities. When designing optimal contracts for heterogonously capable agents, she faced a trade-off between exploiting strongly naïve workers and designing efficient contracts for the most widespread type of worker, according to her posteriors.

In Chapter 3 and Chapter 4, agents played the role of heterogeneously tempted consumers who suffer from self-control problems. In Chapter 3, an uninformed seller offered a loyalty scheme to consumers. If a consumer accepted it, the seller offered him personalised discounts in return for a certain degree of observability over the consumer's temptation level. Loyalty schemes became, therefore, a means of exchange used by the seller to perform personalised pricing. Results showed how this practice, under some conditions, may actually increase consumers' surplus.

Chapter 4 presented a two period game where, in period 1, the principal was not able to commit to the prices and quantities offered to consumers in period 2. In order to compensate them for her taking advantage of their temptation in period 2, the seller offered consumers a fixed payment in period 1. Under some conditions, in this framework, the consumer with a lower valuation of the good obtained an information rent. Similarly, under other conditions, the consumer with the higher valuation of the good was excluded from the market by the seller. Hence, the presence of consumers that suffer from self-control problems was shown to "reverse" the roles of the agent's types with respect to classical theory.

Of course, the papers presented are only the starting point of a deeper analysis of the behavioural biases studied. How would the results of Chapter 2 change if the employer could not perfectly monitor workers, i.e. observe the effort they exert on the job? What kind of loyalty scheme technology would the seller in Chapter 3 set up if he had control over the degree of observability over consumers' preferences it granted? How does the "role reversal" result of consumer's types of Chapter 4 extends to more general mechanism design problems with no commitment? These are all questions that I am eager to address in the near future.

Appendices

Appendix A

Appendix to Chapter 2

A.1 ϕ-Informative Beliefs — Optimal Contracts

In this appendix, I present the solutions of problem (2.2) for every value of p_P and p_U, that is, for all possible combinations generated by conditions (2.12) and (2.13). I also derive workers' utility in each equilibrium. Notice that the ranking and sign of workers' utility levels is proven in Lemma 2.1.

In what follows I define U_i^j as the utility a j-productivity and i-belief type worker obtains at the end of the game: $U_i^j \equiv U_j(w_i^j, e_i^j) = w_i^j - \theta_j e_i^j$.

If conditions (2.12) and (2.13) hold together, the optimal contracts designed for area A are obtained by solving (2.2) with $(IC_{U,\phi})$ and $(IC_{P,\delta})$ binding results in:

$$w_\delta^U = E_\delta(\theta)e_\delta^U \tag{A.1}$$
$$w_\delta^P = E_\delta(\theta)e_\delta^U + \theta_P(e_\delta^P - e_\delta^U) \tag{A.2}$$
$$w_\phi^U = (E_\delta(\theta) - E_\phi(\theta))e_\delta^U + E_\phi(\theta)e_\phi^P + \theta_U(e_\phi^U - e_\phi^P) \tag{A.3}$$
$$w_\phi^P = (E_\delta(\theta) - E_\phi(\theta))e_\delta^U + E_\phi(\theta)e_\phi^P \tag{A.4}$$

and

$$e_\delta^U : y'(e) = \tfrac{E_\delta(\theta) - (1-E(p))E_\phi(\theta) - p_P\lambda\theta_P}{(1-\lambda)p_U} \tag{A.5}$$
$$e_\delta^P : y'(e) = \theta_P \tag{A.6}$$
$$e_\phi^U : y'(e) = \theta_U \tag{A.7}$$
$$e_\phi^P : y'(e) = \tfrac{(1-E(p))E_\phi(\theta) - (1-\lambda)(1-p_U)\theta_U}{(1-p_P)\lambda}. \tag{A.8}$$

This results in:

$$U_\delta^U = E_\delta(\theta)e_\delta^U - \theta_U e_\delta^U < 0 \tag{A.9}$$
$$U_\delta^P = E_\delta(\theta)e_\delta^U - \theta_P e_\delta^U > 0 \tag{A.10}$$
$$U_\phi^U = (E_\delta(\theta) - E_\phi(\theta))e_\delta^U - (\theta_U - E_\phi(\theta))e_\phi^P \leq 0 \tag{A.11}$$
$$U_\phi^P = (E_\delta(\theta) - E_\phi(\theta))e_\delta^U - (\theta_P - E_\phi(\theta))e_\phi^P > 0. \tag{A.12}$$

If condition (2.12) holds while (2.13) fails, the optimal contracts designed for area B are obtained by solving (2.2) with $(IC_{U,\phi})$ and $(IC_{U,\delta})$ binding. This results in:

$$w_\delta^U = E_\delta(\theta)e_\delta^P - \theta_U(e_\delta^P - e_\delta^U) \tag{A.13}$$
$$w_\delta^P = E_\delta(\theta)e_\delta^P \tag{A.14}$$
$$w_\phi^U = (E_\delta(\theta) - E_\phi(\theta))e_\delta^P + E_\phi(\theta)e_\phi^P + \theta_U(e_\phi^U - e_\phi^P) \tag{A.15}$$
$$w_\phi^P = (E_\delta(\theta) - E_\phi(\theta))e_\delta^P + E_\phi(\theta)e_\phi^P \tag{A.16}$$

and

$$e_\delta^U : y'(e) = \theta_U \tag{A.17}$$
$$e_\delta^P : y'(e) = \frac{E_\delta(\theta) - (1 - E(p))E_\phi(\theta) - p_P(1-\lambda)\theta_U}{\lambda p_U} \tag{A.18}$$
$$e_\phi^U : y'(e) = \theta_U \tag{A.19}$$
$$e_\phi^P : y'(e) = \frac{(1 - E(p))E_\phi(\theta) - (1-\lambda)(1-p_P)\theta_U}{(1-p_U)\lambda}. \tag{A.20}$$

this results in:

$$U_\delta^U = E_\delta(\theta)e_\delta^P - \theta_U e_\delta^P < 0 \tag{A.21}$$
$$U_\delta^P = E_\delta(\theta)e_\delta^P - \theta_P e_\delta^P > 0 \tag{A.22}$$
$$U_\phi^U = (E_\delta(\theta) - E_\phi(\theta))e_\delta^P - (\theta_U - E_\phi(\theta))e_\phi^P \leq 0 \tag{A.23}$$
$$U_\phi^P = (E_\delta(\theta) - E_\phi(\theta))e_\delta^P - (\theta_U - E_\phi(\theta))e_\phi^P > 0. \tag{A.24}$$

If conditions (2.12) and (2.13) fail together, the optimal contracts designed for area C are obtained by solving (2.2) with $(IC_{P,\phi})$ and $(IC_{U,\delta})$ binding. This results in:

$$w_\delta^U = E_\delta(\theta)e_\delta^P - \theta_U(e_\delta^P - e_\delta^U) \tag{A.25}$$
$$w_\delta^P = E_\delta(\theta)e_\delta^P \tag{A.26}$$
$$w_\phi^U = (E_\delta(\theta) - E_\phi(\theta))e_\delta^P + E_\phi(\theta)e_\phi^U \tag{A.27}$$
$$w_\phi^P = (E_\delta(\theta) - E_\phi(\theta))e_\delta^P + E_\phi(\theta)e_\phi^U + \theta_P(e_\phi^P - e_\phi^U). \tag{A.28}$$

and

$$e_\delta^U : y'(e) = \theta_U \tag{A.29}$$
$$e_\delta^P : y'(e) = \frac{E_\delta(\theta) - (1 - E(p))E_\phi(\theta) - p_P(1-\lambda)\theta_U}{\lambda p_U} \tag{A.30}$$
$$e_\phi^U : y'(e) = \frac{(1 - E(p))E_\phi(\theta) - \lambda(1 - p_U)\theta_P}{(1 - p_P)(1-\lambda)} \tag{A.31}$$
$$e_\phi^P : y'(e) = \theta_P. \tag{A.32}$$

this results in:

$$U_\delta^U = E_\delta(\theta)e_\delta^P - \theta_U e_\delta^P < 0 \tag{A.33}$$
$$U_\delta^P = E_\delta(\theta)e_\delta^P - \theta_P e_\delta^P > 0 \tag{A.34}$$
$$U_\phi^U = (E_\delta(\theta) - E_\phi(\theta))e_\delta^P - (\theta_U - E_\phi(\theta))e_\phi^U \leq 0 \tag{A.35}$$
$$U_\phi^P = (E_\delta(\theta) - E_\phi(\theta))e_\delta^P - (\theta_U - E_\phi(\theta))e_\phi^U > 0. \tag{A.36}$$

Finally, if condition (2.12) fails while (2.13) holds, the optimal contracts designed for area D are obtained by solving (2.2) with $(IC_{P,\phi})$ and $(IC_{P,\delta})$ binding results in:

$$w_\delta^U = E_\delta(\theta)e_\delta^U \tag{A.37}$$
$$w_\delta^P = E_\delta(\theta)e_\delta^U + \theta_P(e_\delta^P - e_\delta^U) \tag{A.38}$$
$$w_\phi^U = (E_\delta(\theta) - E_\phi(\theta))e_\delta^P + E_\phi(\theta)e_\phi^U \tag{A.39}$$
$$w_\phi^P = (E_\delta(\theta) - E_\phi(\theta))e_\delta^U + E_\phi(\theta)e_\phi^U + \theta_P(e_\phi^P - e_\phi^U). \tag{A.40}$$

and

$$e_\delta^U : y'(e) = \frac{E_\delta(\theta) - (1-E(p))E_\phi(\theta) - p_U \lambda \theta_P}{(1-\lambda)p_P} \tag{A.41}$$

$$e_\delta^P : y'(e) = \theta_P \tag{A.42}$$

$$e_\phi^U : y'(e) = \frac{(1-E(p))E_\phi(\theta) - \lambda(1-p_U)\theta_P}{(1-\lambda)(1-p_P)} \tag{A.43}$$

$$e_\phi^P : y'(e) = \theta_P. \tag{A.44}$$

this results in:

$$U_\delta^U = E_\delta(\theta)e_\delta^U - \theta_U e_\delta^U < 0 \tag{A.45}$$

$$U_\delta^P = E_\delta(\theta)e_\delta^U - \theta_P e_\delta^U > 0 \tag{A.46}$$

$$U_\phi^U = (E_\delta(\theta) - E_\phi(\theta))e_\delta^P - (\theta_U - E_\phi(\theta))e_\phi^U \leq 0 \tag{A.47}$$

$$U_\phi^P = (E_\delta(\theta) - E_\phi(\theta))e_\delta^U - (\theta_P - E_\phi(\theta))e_\phi^U > 0. \tag{A.48}$$

A.2 Proof of Lemma 2.1

(1) First of all, $U_i^P \geq \max\{0, U_i^U\}$. To see this notice that from $IC_{P,i}$: $U_i^P \geq w_i^U - \theta_P e_i^U \geq U_i^U$ for all e, with strict inequality for all $e > 0$. Moreover, suppose $U_i^U < 0$, $U_i^P > 0$ is implied by the (IR_J). Using the above, the (IR_δ), the (IC_ϕ) and the fact that $\phi > \delta$ I can write the following sequence of inequalities:

$$\phi U_\phi^P + (1-\phi)U_\phi^U \geq \phi U_\delta^P + (1-\phi)U_\delta^U \geq \delta U_\delta^P + (1-\delta)U_\delta^U \geq 0$$

which proves that (IR_ϕ) holds.

Suppose now that (IR_δ) was not binding. Then the principal can decrease all wages in the contract by $\epsilon > 0$ without affecting any of the other constraints while raising profits.

(2) Given the above, $U_\delta^P \geq 0 \geq U_\delta^U$. Rearrange the ICs in the following way:

$$\delta(U_\delta^P - U_\phi^P) + (1-\delta)(U_\delta^U - U_\phi^U) \geq 0 \tag{IC_δ}$$

$$\phi(U_\delta^P - U_\phi^P) + (1-\phi)(U_\delta^U - U_\phi^U) \leq 0. \tag{IC_ϕ}$$

The above shows that the sign of the convex combination between $(U_\delta^P - U_\phi^P)$ and

$(U_\delta^U - U_\phi^U)$ changes from non-negative to non-positive when the combination gets closer to $(U_\delta^P - U_\phi^P)$ instead of $(U_\delta^U - U_\phi^U)$. This implies that $(U_\delta^P - U_\phi^P) \leq 0$ and that $(U_\delta^U - U_\phi^U) \geq 0$ which implies $U_\phi^P \geq U_\delta^P \geq 0 \geq U_\delta^U \geq U_\phi^U$.

Suppose now (IC_ϕ) was not binding, then the principal can decrease both e_ϕ^U and e_ϕ^P keeping period 2 incentive compatibility unchanged. In this way, profits would rise, (IC_δ) would be relaxed and (IR_ϕ) would still hold by the Lemma above. To see that (IC_δ) is slack rearrange the ICs in the following way:

$$\delta(U_\delta^P - U_\delta^U) + U_\delta^U \geq \delta(U_\phi^P - U_\phi^U) + U_\phi^U \qquad (IC_\delta)$$

$$\phi(U_\phi^P - U_\phi^U) + U_\phi^U = \phi(U_\delta^P - U_\delta^U) + U_\delta^U \qquad (IC_\phi)$$

From (IC_ϕ), $U_\phi^U = \phi(U_\delta^P - U_\delta^U) + U_\delta^U - \phi(U_\phi^P - U_\phi^U)$. Substitute it back into the (IC_δ) to get: $(U_\phi^P - U_\phi^U) \geq (U_\delta^P - U_\delta^U)$ which always holds given Lemma 2.

A.3 Proof of Result 1

Consider the principal objective function as in (2.4). Notice that the effect of e_δ^P is given by $\lambda E_\phi(\theta) + (1-\lambda)\theta_U - E_\delta(\theta)$ which is positive if and only if condition (2.5) holds. Hence, if that is the case, $(w_\delta^P, e_\delta^P) = (E_\delta(\theta), 1)$.

If, instead, (2.5), then the employer wants to set e_δ^P as low as possible. However, ex-post incentive compatibility implies that $e_\delta^P \geq e_\delta^U$. Hence, $(w_\delta^P, e_\delta^P) = (E_\delta(\theta) e_\delta^U, e_\delta^U)$ and the contract for δ induces (imaginary) pooling.

A.4 Proof of Result 2

To prove the statement simply work out the wage levels and notice that: if (2.5) holds:

$$w_\delta^U - \theta_U e_\delta^U = (E_\delta(\theta) - \theta_U) < 0$$
$$w_\phi^P - \theta_P e_\phi^P = (E_\delta(\theta) - E_\phi(\theta)) > 0.$$

If it does not hold:

$$w_\delta^U - \theta_U e_\delta^U = (E_\delta(\theta) - \theta_U) e_\delta^U \in [(E_\delta(\theta) - \theta_U), 0]$$
$$w_\phi^P - \theta_P e_\phi^P = (E_\delta(\theta) - E_\phi(\theta)) e_\delta^U \in [0, (E_\delta(\theta) - E_\phi(\theta))].$$

A.5 Proof of Result 3

To prove the statement simply work out the wage levels and notice that:

$$w_\phi^U - \theta_U e_\phi^U = (E_\phi(\theta) - \theta_U) < 0$$
$$w_\delta^P - \theta_P e_\delta^P = \theta_P e_\delta^P - \theta_P e_\delta^P = 0.$$

A.6 Proof of Result 4

The employer wants to design incentive compatible contracts that maximise profits. From Lemma 2.1 I know that (IR_δ) and (IC_ϕ) have to bind in period 1. The first is irrelevant for the optimistic workers' contract.

I can represent incentive compatibility in a (w_ϕ^U, w_ϕ^P) space as in Figure 2.1 in the paper. Incentive compatible contracts lie above the (IC_ϕ) between $(IC_{U,\phi})$ and $(IC_{P,\phi})$. Expected utility increases towards the top right, profits towards the bottom left. Hence, an optimal contract always lies on the (IC_ϕ) binding line. In order to select the optimal contract, I study the slope of the isoprofits, in a (w_ϕ^U, w_ϕ^P) space, and compare it to the one of the (IC_ϕ). The former is given by $-\frac{(1-\lambda)(1-p_U)}{\lambda(1-p_P)}$ while the latter is $-\frac{1-\phi}{\phi}$. Hence, isoprofits are steeper than the (IC_ϕ) if $(1-p_P) \leq (1-p_U)\frac{\phi}{1-\phi}\frac{1-\lambda}{\lambda}$ that can be rearranged to obtain (2.12). If the latter holds, the right hand side graph of Figure 2.1 shows that the optimal contract has $(IC_{U,\phi})$ binding and induces efficient effort in optimistic unproductive workers.

A.7 Proof of Result 5

The proof follows the one for Result 4. Simply substitute the (IC_ϕ) with the (IR_δ) constraint. Notice that here I use a partial equilibrium argument. that is, I assume that given the optimal contract for the pessimistic worker, the contract designed for the optimistic worker adjusts in equilibrium in order for the (IC_ϕ) to bind.

A.8 Proof of Corollary 2.5

To prove the Corollary simply notice that in area A both (2.12) and (2.13) must hold and that $p_P \geq p_U$. From the proofs of Result 4 and Result 5 the conditions are respectively equivalent to:

$$\frac{1-\phi}{\phi} \leq \frac{(1-\lambda)}{\lambda}\frac{(1-p_U)}{(1-p_P)} \qquad \text{and} \tag{A.49}$$

$$\frac{(1-\delta)\lambda}{(1-\lambda)\delta} \geq \frac{p_U}{p_P}. \tag{A.50}$$

Start from noticing that $\frac{1-\lambda}{\lambda} > \frac{1-\phi}{\phi}$. When $p_P \geq p_U$, $\frac{1-p_U}{1-p_P} \geq 1$. Hence, (A.49) always holds. For (A.50), notice that $\frac{(1-\delta)\lambda}{(1-\lambda)\delta} > 1$. Hence the condition always holds for $\frac{p_U}{p_P} \leq 1$. This proves that aera A always takes up the entire space above the 45 degree line in Figure 2.3.

A.9 Proof of Result 6

Checking for $(e_i^P - e_i^U) > 0$, it is easy to see that this is true for all contracts and all types when the $(IC_{P,\delta})$ binds. As for the rest of the contracts, it is also easy to check that the offers for productive types are always incentive compatible when they induce separation. As for the ones for the pessimistic type:

$$e_\delta^P - e_\delta^U \geq 0 \text{ if and only if } \phi < \frac{\delta}{1-E(p)} \tag{A.51}$$

which generates (2.14).

Appendix B

Appendix to Chapter 3

B.1 Incentives to enter one of the Small Stores

Assumption 1 implies that $p_x \leq \frac{u(x)}{2}$. Here I show how this is also a sufficient condition for the consumers to have the (at least weak) incentive to enter one of the small stores that offer M_s. Notice that, given Assumption 1, the condition for a consumer to enter in one of the small stores can be written as:

$$
\begin{aligned}
& W_i(M_s) \geq 0 \\
& u(x) + \phi_i v(x) - 2p_x - \max\left\{\phi_i v(x) - p_x, 0\right\} \geq 0 \qquad \text{(B.1)}
\end{aligned}
$$

Consider consumers that do not suffer from temptation when facing x, i.e., the ones for which $0 \succ_T (x, p_x)$. The above expression than boils down to $u(x)+\phi_i v(x)-2p_x \geq 0$ that is solved by any $p_x \leq \frac{u(x)+v(x)}{2} > \frac{u(x)}{2}$.

Consider now consumers that are tempted by the offer for good x,i.e., the ones for which $(x, p_x) \succ_T 0$. For them, the constraint above becomes $u(x) - p_x \geq 0$, which is also solved by any $p_x \leq \frac{u(x)}{2}$.

B.2 Proof of Result 7

Given Assumption 1, (3.8) can be rewritten as:

$$
\begin{aligned}
& \max_{p_z}[p_z - c_z] \qquad \text{s.t.} \\
& \max_{a \in M_m}[u(a) + \phi_i v(a) - 2p_a] - \max_{a \in M_m}[\phi_i v(a) - p_a] \\
& \qquad \geq u(x) + \phi_i v(x) - 2p_x - \max_{a \in M_s}[\phi_i v(a) - p_a] \\
& u(z) + \phi_i v(z) - 2p_z \geq u(x) + \phi_i v(x) - 2p_x.
\end{aligned}
$$

The IC holds if and only if

$$p_z \leq \tfrac{1}{2}\phi_i(v(z) - v(x)) + p_x \tag{B.2}$$

where I used $u(x) = u(z)$.

Consider now the most tempting offer as defined in Section 3.2. Notice that, for i:

$$(x, p_x) \succ_T 0 \text{ if } \quad \phi_i > \frac{p_x}{v(x)} \tag{B.3}$$

$$(z, p_z) \succ_T (x, p_x) \text{ if } \quad p_z \leq \phi_i(v(z) - v(x)) + p_x \tag{B.4}$$

$$(z, p_z) \succ_T 0 \text{ if } \quad p_z \leq \phi_i v(z). \tag{B.5}$$

It is easy to see then that every price that satisfies the IC, also makes $(z, p_z) \succ_T (x, p_x)$ for i. To understand why this is true, remember that the tempting features of good z are not hurting the consumer ($u(z) = u(x)$). Hence, if he values the good enough to find it optimal to fall to temptation and buy it, the temptation utility he obtains from offer (z, p_z) is also high enough to make (B.4) hold.

Start by assuming that **(B.3) holds**. Notice that conditions (B.2) and (B.4) are identical. Hence, a solution to (3.8) cannot fail to satisfy (B.4). I now solve (3.8) assuming that the solution satisfies (B.4) and then check that this is true afterwards. Notice that in this case (B.4) implies (B.5) and $W_i(M_s) = u(x) - p_x$. Hence the PC becomes:

$$\begin{aligned} & u(z) + \phi_i v(z) - 2p_z - \phi_i v(z) + p_z \geq u(x) - p_x \\ & \Rightarrow p_z^* = p_x \end{aligned}$$

where p_z^* clearly satisfies (B.2) and, therefore, (B.4).

Assume now that **(B.3) does not hold**. Given this, then (B.5) implies (B.4) and $W_i(M_s) = u(x) + \phi_i v(x) - 2p_x$, since (x, p_x) does not tempt 0. As above I solve (3.8) assuming (B.5) holds and then check for it with the solution found. When (B.5) holds the PC becomes:

$$\begin{aligned} & u(z) + \phi_i v(z) - 2p_z - \phi_i v(z) + p_z \geq u(x) + \phi_i v(x) - 2p_x \\ & \Rightarrow p_z^*(\phi_i) = 2p_x - \phi_i v(x). \end{aligned}$$

Notice, however, that this $p_z^*(\phi_i)$ is compatible with (B.5) only for $\phi_i \geq \frac{2p_x}{v(z)+v(x)} < \frac{p_x}{v(x)}$. For $\phi_i < \frac{2p_x}{v(z)+v(x)}$, (B.5) fails at the $p_z^*(\phi_i)$ derived which brings to a contradiction.

When (B.5) and (B.3) both fail, IC and PC coincide. Hence,

$$p_z^*(\phi_i) = \tfrac{1}{2}\phi_i(v(z) - v(x)) + p_x$$

from (B.2). This solution violates (B.5) for all $\phi_i < \frac{2p_x}{v(z)+v(x)}$. Hence, the latter is a saddle point in the graph of $p_z^*(\phi)$ (see Figure 3.1 in the text).

B.3 Proof of Result 8

First of all notice that, since loyalty cards are ϕ-uninformative, in equilibrium, there cannot be two different consumers buying good z at different prices. The reason for this is that the monopolist has no information to discriminate between two different consumers. Hence, since she cannot change the quantity/quality of an offer but only the price, she cannot set up an incentive compatible pricing scheme that makes consumers self-select.[1] Every consumer that buys z, purchases one unit of the good at the lowest possible price.

Given this, the monopolist sets either p_A or p_R equal to a price that induces the optimal interval of consumers to enter the store. Differently from the case of asymmetric information with no loyalty scheme however, now strongly consumers are not driven away from the superstore. They can be attracted back in to buy x by setting either p_A or p_R equal to p_{max}. Hence, the optimal p_z set for consumers that buy good z will be $p_z^\dagger$ as describe din the paper.

Later in the proof I argue that it makes no difference which of the two prices is set equal to $p_z\dagger$. Hence, I let it be p_A. Given this, the only use left for p_R is to be set high enough to create a commitment device for all those consumers that are tempted by $p_z^\dagger$, but do not find it optimal to choose $(z, p_z^\dagger)$ form the menu. I define in the paper the price that has this function: p_{max}. Notice that p_{max} can take any value greater or equal to the one I defined. Its value, however, does not affect the outcome of the game.

As for consumers' equilibrium strategy, notice that types in $\left[0, \underline{\phi}(p_z^\dagger)\right]$ have no

[1]Unless of course, one where some consumers "self-exclude" from the market, as the one described here.

interest is buying good z at $p_z^\dagger$ nor are they tempted by it. Hence, they obtain the same ex-ante utility no matter what they do.

Consumers in $\Phi(p_z^\dagger)$ obtain a strictly larger $W_i(\cdot)$ if they accept the card, enter the superstore and buy good z, with respect to any other action at their disposal. To see why this is true, notice that every consumer with consumer in $\Phi(p_z^\dagger)$ has a strictly larger willingness to pay than consumers $\underline{\phi}(p_z^\dagger)$ and $\overline{\phi}(p_z^\dagger)$. Also, notice that $p_z^\dagger = p_z^*(\underline{\phi}(p_z^\dagger)) = p_z^*(\overline{\phi}(p_z^\dagger))$.
Hence, consumers inside the interval obtain $W_i\left(\left\{(x,p_x),(z,p_z^\dagger),0\right\}\right) > W_i\left(\left\{(x,p_x),0\right\}\right) = W_i\left(\left\{(x,p_x),(z,p_R),0\right\}\right)$.

Finally, consider consumers in $\left[\overline{\phi}(p_z^\dagger),1\right]$. They are tempted by $(z,p_z^\dagger)$ so much that they would choose it, were they to accept the loyalty card and enter the superstore. However, it is easy to check that their ex-ante utility obtained by doing so is $W_i\left(\left\{(x,p_x),(z,p_z^\dagger),0\right\}\right) < W_i\left(\left\{(x,p_x),0\right\}\right) = W_i\left(\left\{(x,p_x),(z,p_R),0\right\}\right)$. Hence, they strictly prefer to reject the loyalty card. Once rejected it, the utility from entering any store is equal to $W_i\left(\left\{(x,p_x),0\right\}\right)$. Hence, they have a weak incentive to enter the monopolist's superstore. This concludes the proof.

Finally, notice that, since p_A is constant over ϕ_i, there exist another perfectly equivalent equilibrium of the following type:

Corollary B.1. *When the partition is a trivial one, i.e., of fineness* 0*, and loyalty cards are* ϕ*-uninformative, there exist a second equilibrium where the monopolist sets:*

$$p_z = \begin{cases} p_R = p_z^\dagger & \text{if } i \text{ accepts} \\ p_A \geq p_{max} & \text{if } i \text{ rejects} \end{cases} \tag{B.6}$$

Consumers with type in:

- $\left[0,\underline{\phi}(p_z^\dagger)\right]$ *are indifferent between accepting or rejecting the loyalty card and have a (weak) incentive to enter the superstore to buy* x*,*

- $\Phi(p_z^\dagger)$ *are strictly better off by rejecting the loyalty card and buy* z *in the superstore,*

- $\left[\overline{\phi}(p_z^\dagger),1\right]$ *strictly prefer to accept the loyalty card. They have (weak) incentive to buy* x *from the superstore.*

B.4 Proof of Lemmas

B.4.1 Lemma 3.1

I first prove that $\underline{p}_z = p_x$. Consider consumers for which (B.3) holds, (B.4) and (B.5) also hold from the proof of Result 7. For these consumers, the PC becomes $u(z) - p_x \geq u(x) - p_x$ at $p_z = p_x$, and it holds with equality. Therefore, there is no higher price that attracts consumers with $\phi_i \geq \frac{p_x}{v(x)}$ into the monopolist's superstore. It remains to check if all other types enter at $p_z = p_x$. Consider now consumers for which (B.3) fails. It is easy to see that (B.4) holds for all of them at $p_z = p_x$. For (B.5) to hold instead, ϕ_i has to be larger than the ratio $\frac{p_x}{v(z)}$. If this is the case, the PC becomes $u(z)+p_x \geq u(x)+\phi_i v(x)$ which holds with inequality for all $\phi_i < \frac{p_x}{v(x)}$. Finally, for consumers with $\phi_i < \frac{p_x}{v(z)}$, 0 is the most tempting offer in menu $M_m(p_x)$. Hence the PC simplifies to $u(z)+\phi_i v(z) \geq u(x)+\phi_i v(x)$, which holds with inequality. This proves that p_x is the largest possible value of p_z that attracts every consumer into the monopolist's superstore.

Since all consumers are entering her superstore, the expected profits she obtains at $\underline{p}_z$ are given by:

$$\Pi(z, p_x) = \int_0^1 (p_x - c_z) d\phi_i = (p_x - c_z) \tag{B.7}$$

I now prove that $\bar{p}_z = \phi^* v(z)$. Consider first-best p_z in (3.9) and consider all three intervals. Notice that the highest value of $p_z^*(\phi_i)$ is reached when $\phi_i = \phi^*$, $p_z^*\left(\phi^*\right) = \phi^* v(z) \equiv \bar{p}_z$. At this price, only consumer ϕ^* is willing to enter the monopolist's superstore. The expected profits she obtains at $\bar{p}_z$ are given by:

$$\Pi(z, \bar{p}_z) = \left(\bar{p}_z - c_z\right) \Pr\left\{\phi_i = \phi^*\right\} = 0 \tag{B.8}$$

since ϕ_i is a continuous random variable.

For the second part of the Lemma, it is obvious that no consumer for which $(x, p_x) \succ_T 0$ will enter the superstore for $p_z > p_x$. Hence the only relevant interval of ϕ_i is $\left[0, \frac{p_x}{v(x)}\right)$. Notice, from the derivation of Result 7, that consumers $\phi_i \in \left[0, \phi^*\right)$ enter the monopolist's superstore if and only if $p_z \leq \frac{1}{2}\phi_i\left(v(z) - v(x)\right) + p_x$. Solving this inequality for ϕ_i, I find the subset of $\left[0, \phi^*\right)$ of consumers that enter the superstore for a given p_z. This is given by $\left[\frac{2(p_z - p_x)}{v(z)-v(x)}, \phi^*\right)$. A similar reasoning can be applied to consumers in $\left[\phi^*, \frac{p_x}{v(x)}\right)$. They enter the monopolist's superstore if and only if $p_z \leq 2p_x - \phi_i v(x)$. Solving this for ϕ_i I find the subset of $\left[\phi^*, \frac{p_x}{v(x)}\right)$ consumers the enter at a given p_z. This is given by $\left[\phi^*, \frac{2p_x - p_z}{v(x)}\right]$. To conclude the proof notice that the union of the two subset is the continuous interval $\Phi_1(p_z)$.[2]

B.4.2 Lemma 3.3

The proof of this Lemma is straightforward. Simply notice that:

$$
\begin{aligned}
& arg \max_{p_z} \int_{\Phi(p_z)} \pi_z d\phi + \int_{\phi^l}^{\underline{\phi}(p_z)} \pi_x d\phi + \int_{\overline{\phi}(p_z)}^{\phi^r} \pi_x d\phi \\
&= arg \max_{p_z} \left(\underline{\phi}(p_z) - \phi^l\right)(p_x - c_x) + \left(\overline{\phi}(p_z) - \underline{\phi}(p_z)\right)(p_z - c_z) + \left(\phi^r - \overline{\phi}(p_z)\right)(p_x - c_x) \\
&= arg \max_{p_z} \left(\overline{\phi}(p_z) - \underline{\phi}(p_z)\right)(p_z - c_z - p_x + c_x) \quad \text{if} \quad [\phi^l, \phi^r] \supseteq \Phi(p_z^\dagger)
\end{aligned}
$$

Hence, as long as $[\phi^l, \phi^r] \supseteq \Phi(p_z^\dagger)$, the equilibrium price for p_z is $p_z^\dagger$.

B.4.3 Lemma 3.4

The proof of this Lemma follows the discussion in the paper. Consider the maximization problem in (3.20). Element Φ_1 is composed only by weakly tempted consumers. Hence, the seller knows that by setting a price level p_1 she excludes from the market for z only consumers that do not value the good enough to pay p_1, i.e., $[0, \underline{\phi}(p_1))$. The latter, however, are still willing to enter the superstore to buy x. Of course this maximization ignores the cutoff $\underline{\phi}(p_1) \geq 0$. Hence, if the solution is smaller than p_x, the seller will set $p_1 = p_x$.

When Φ_1 is composed also of some strongly tempted consumers, the maximisation problem is ignoring the portion of negatively sloped willingness to pay that goes from ϕ^* to ϕ^1. This is the reason for the extra cutoff at $p_z^*(\phi^1)$, as explained in the text.

B.4.4 Lemma 3.5

The proof follows the one for Lemma (3.5).

[2] Of course, $\Phi_1(p_x)$ coincides with $\left[0, \frac{p_x}{v(x)}\right)$ and $\Phi_1(\bar{p}_z)$ collapses to the single value ϕ^*.

B.4.5 Lemma 3.6

The proof follows the one for the previous Lemmas. Notice that p_x plays no role in the maximisation unless $\phi^r > \frac{p_x}{v(x)}$. While $p_n = p_x$ always when $\frac{p_x}{v(x)}$.

B.4.6 Lemma 3.7

The proof follows from the one for the previous Lemmas and the discussion in the text.

B.5 Proof of Result 9

When deciding whether to accept or reject the loyalty card, consumers have the (weak) incentive to accept. They also, have the (weak) incentive to enter the superstore to buy z. This is because:

$$W_i\left(\left\{(x,p_x),(z,p_z^*(\phi),0\right\}\right) = W_i\left(\left\{(x,p_x),0\right\}\right) = W_i\left(\left\{(x,p_x),(z,p_R),0\right\}\right)$$

B.6 Consumers without self-control problems

In the paper I have analysed the market equilibrium for the case of consumers that suffer from self-control problems. In this Appendix, I describe the results of a “classical” model that fails to account for this aspect of consumer’s preferences. I show how such a model would miss important qualitative characteristics captured by this paper.

First of all, I need to clarify a controversial point. The fact that consumers are free of self-control problems does not mean that they are not tempted by good z. Their temptation utility function $V_i(z,p_z)$, and so their valuation of the good, is unchanged. What does change, is their ex-ante utility, since now consumers are assumed not to care for the self-control cost of resisting temptation. A good example for this case is to consider again the consumer described in Section 3.2 in the paper in front of the menu with the healthy salad and the tasty hamburger. This consumer is still tempted by the burger but he is now *not* considering the cost of sticking to his diet. He now does not care about resisting to the temptation of ordering the burger and the self-control cost that this creates. He takes his decision accounting for how tempting the burger is. If temptation is overwhelming he is "happy" to fall to it and order the burger.

Given this, the model becomes one with classical consumers that, considering their preferences, buy good $a = z, x$ solving:

$$\max_{a \in M_j} [U(a, p_a) + V_i(a, p_a)]. \tag{B.9}$$

This new, simpler, model has no ex-ante nor ex-post stage. Consumer i simply evaluates $U(a, p_a) + V_i(a, p_a)$ at $a = x, z, 0$ and enters the store that sells the offer that maximises his utility. I now find the equilibrium of first-best, second-best (asymmetric information) and of the loyalty card schemes system case.

B.6.1 First-best

If the monopolist can perfectly observe consumers' type without the need of any technology, she solves:

$$\begin{gathered} \max_{p_z} [p_z - c_z] \qquad \text{s.t.} \\ u(z) + \phi_i v(z) - 2p_z \geq u(x) + \phi_i v(x) - 2p_x \\ \Longrightarrow p_z^*(\phi_i) = \tfrac{1}{2}\phi_i(v(z) - v(x)) + p_x \\ \text{for all } \phi_i \in [0, 1]. \end{gathered}$$

B.6.2 Asymmetric Information

If the monopolist cannot observe consumers' consumer and has no access to any loyalty card schemes technology, she compares the expected profits of attracting all consumers into the store ($p_z = p_x$) with the ones of excluding consumers with a low valuation of the good.[3] Following a similar reasoning as the one in Section 3.5.2 in the paper, she understands that by setting price p_z only consumers with $\phi_i \geq \frac{2(p_z - p_x)}{v(z) - v(x)}$ enter her superstore. That is: $\Phi(p_z) \equiv \left[\underline{\phi}(p_z), 1\right]$. Hence, the optimal price to set if she

[3] Price p_x attracts all consumers in since it is the highest price that type $\phi = 0$ is willing to pay for z. All other consumers have higher willingness to pay.

wants to exclude less tempted consumers from the market is given by:

$$p'_z = \arg\max_{p_z} \int_{\Phi(p_z)} (p_z - c_z) d\phi + \int_0^{\underline{\phi}(p_z)} (p_x - c_x) d\phi \tag{B.10}$$

$$= \tfrac{1}{4}((v(z) - v(x))) + \tfrac{1}{2}(p_x + c_z - c_x) > p_x \tag{B.11}$$

which only attracts consumers with $\phi_i \geq \frac{1}{2} + \frac{c_z - c_x}{v(z) - v(x)} = \underline{\phi}(p'_z)$. This case is represented, together with first best, in Figure B.1.

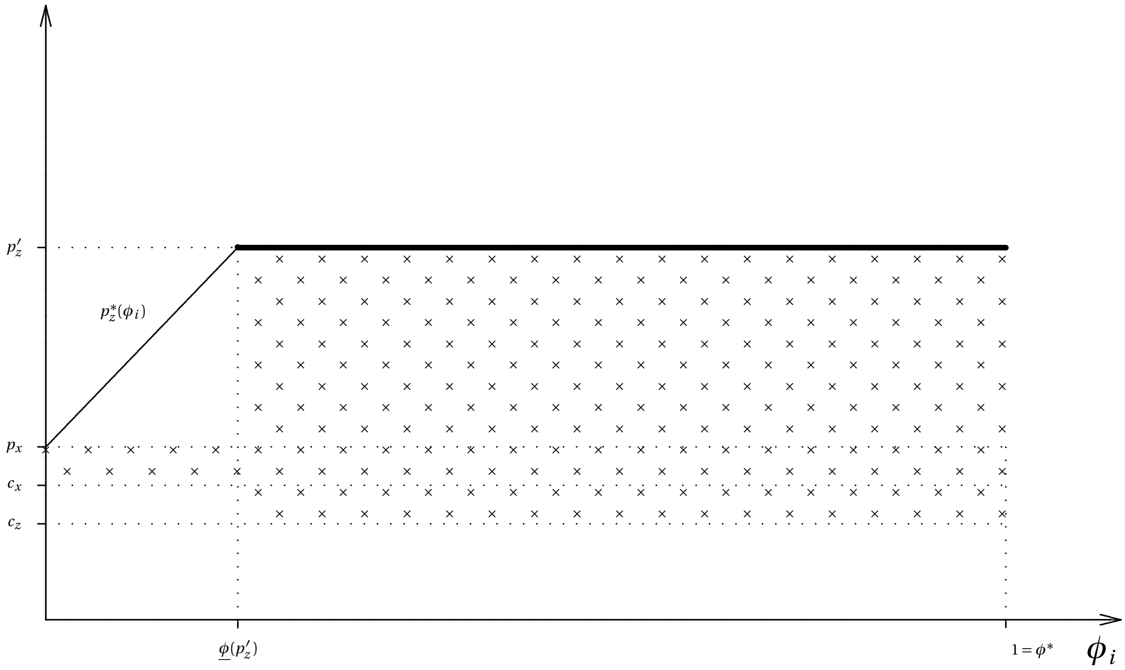

Figure B.1: Loyalty Schemes Without Self-Control

In the Figure, the case of a model with consumer free from self-control problem is described. The first-best price is represented by the $p^*_z(\phi_i)$ curve. The asymmetric information profits are depicted by the "×" shaded area and the asymmetric information equilibrium price, p'_z, by the bold line.

B.6.3 Loyalty Cards Scheme.

Suppose now that the monopolist has access to a technology as the one described in Section 3.6 in the paper. By looking at Figure 3.5 it is easy to see how the algorithm for p_A described in the paper works for this case as well with some exemptions. First of all, notice that if consumers do not suffer from self-control problems, none of them finds an offer "too tempting not to buy it but too costly to buy it optimally". That is to say, there is no downward sloping portion of the willingness to pay curve. Hence, to find the equilibrium prices, one should use the algorithm described in the paper considering elements composed of only weakly tempted consumers.[4] This implies that, unlike a model with individuals that suffer from self control problems, the equilibrium price of asymmetric information and the one for ϕ-uninformative loyalty schemes coincides.

A model without consumers with self-control problems would misreport completely the situation of highly tempted consumers. Here, they are never excluded from the market. No consumer in the market is strictly willing to reject the loyalty card for "commitment" reasons. There is no reason for them to reject a loyalty card since when they are *too* tempted they are "happy" to fall to temptation. Hence, the asymmetric information and ϕ-uninformative loyalty scheme equilibria coincide. This is a crucial difference that deserves discussion and analysis. It is, however, not a necessary condition for the non-monotonicity of acceptance of the loyalty card.

B.7 The Case of $c_z > c_x$

In this appendix I analyse the case of $c_z > c_x$. An example for this are new innovative technological products that tempt consumers in the market and are, usually, more costly to produce than the old products. With this new assumption, I carry on my analysis following the structure of the paper. I present the modified results for this section as Corollaries to the Propositions and Results of the paper. The intuitions and mechanics of the equilibrium do not change. However, more consumers are now excluded from the market for good z. Furthermore, there are conditions over c_z for the monopolist to find it profitable to sell good z in the first place. These are (B.14) and (B.15) described below.

[4] Notice that in the case of consumers without self-control problems, $\phi^* = 1$, there is no $\overline{\phi}(p_z')$ and the fraction $\frac{p_x}{v(x)}$ looses importance.

B.7.1 First-Best

In this new setting the monopolist is willing to sell good z to consumer i only when she obtains higher profits from doing so with respect to selling x. That is if and only if:

$$p_z \geq p_x + c_z - c_x. \tag{B.12}$$

Consider equation (3.9). It is easy to see that condition (B.12) is not satisfied for all ϕ_i. The following Corollary and its Proof show how Result 7 changes when $c_z > c_x$.

Corollary B.2. *Let $c_z > c_x$. Given the optimal pricing scheme for good z of Result 7, if the monopolist is able to perfectly observe consumers' level of temptation, she is willing to sell good z only to consumers with $\phi_i \in \Phi_0 \equiv \left[\frac{2(c_z-c_x)}{v(z)-v(x)}, \frac{(p_x+c_x)-c_z}{v(x)}\right]$. Hence, in equilibrium she offers $(z, p_z^*(\phi_i))$ to consumers in Φ_0 and (x, p_x) to consumers outside Φ_0.*

From selling good z she obtains expected profits:

$$\Pi(z, p_z^*(\phi_i)) = \int_{\Phi_0} (p_z^*(\phi_i) - c_z) d\phi_i \tag{B.13}$$

Consumers inside (outside) Φ_0 enter the superstore and buy good z (good x).

Proof. As described above, the seller is willing to sell good z to consumer i if and only if condition (B.12) holds. It is easy to show how this condition holds only for specific values of ϕ_i. First consider the case of $\phi_i < \phi^*$. In this case $\frac{1}{2}\phi_i(v(z) - v(x)) + p_x \geq p_x + c_z - c_x$ must hold. That happens if and only if $\phi_i \geq \frac{2(c_z-c_x)}{v(z)-v(x)}$. If, instead, $\phi_i \in \left[\phi^*, \frac{p_x}{v(x)}\right]$ then $2p_x - \phi_i v(x) \geq (p_x + c_z) - c_x$ must hold. That happens if and only if $\phi_i \leq \frac{(p_x+c_x)-c_z}{v(x)}$. for the case of $\phi_i > \frac{p_x}{v(x)}$, it is obvious to see that $p_z = p_x$ cannot satisfy (B.12). □

The intuition of the Result is quite simple. Since good z is now more costly to produce than good x, the monopolist will sell it only to consumers with a willingness to pay high enough. The consumers at the boundaries of Φ_0 are the ones with an ex-ante willingness to pay high enough to grant the firm a profit equal to the one of selling good x. The left boundary one does not value z enough to pay a high price. The right boundary one does not want to fall to temptation when the price is too high.

Finally, notice that the existence of interval Φ_0 depends on the relation between c_z, c_x and p_x. Comparing the left and the right boundary of Φ_0, I obtain:

$$\begin{aligned} c_z &\leq \delta p_x + c_x \\ \delta &\geq \tfrac{c_z - c_x}{p_x} \end{aligned} \tag{B.14}$$

where $\delta = \frac{v(z)-v(x)}{v(z)+v(x)} \in [0,1]$ is a relative measure of the difference in the temptation value of the two goods. Since $c_z > c_x$, in order for the monopolist to be willing to sell good z at all, the extra perks of good z have to increase the enjoyment that consumers obtain with respect to good x enough. If $v(z)-v(x)$ is too low, then there are no consumers in the market willing to pay a price large enough to cover the higher cost of production of z.

B.7.2 Asymmetric Information

The Results of Section 3.5.2 and Proposition 3.2 change in the following way. First of all, from B.7.1, I know that it is not optimal to sell good z to consumers outside Φ_0. Hence, if the monopolist wants to sell good z she has three choices similar to the case in 3.5.2. She can (i) charge the highest possible price that attracts all consumer's in Φ_0 into the superstore, $\underline{p}_z$, (ii) sell the good to the highest possible price (satisfying PC and IC) that attracts at least one consumer into the store, $\overline{p}_z$, and (iii) sell good z at an intermediate price, $p'_z \in \left(\underline{p}_z, \overline{p}_z\right)$ and attract only a subset of consumers, $\Phi(p'_z) \subset \Phi_0$, into the store. Using a similar argument to the one for Lemma 3.1, it is easy to show that $\left[\underline{p}_z, \overline{p}_z\right] = \left[p_x + c_z - c_x, \phi^* v(z)\right]$. Notice that the assumption that $c_z > c_x$ does not affect consumers' willingness to pay. Hence, $\overline{p}_z$ is not affected by the change. As for $\underline{p}_z$, notice that its value comes directly from condition (B.12). To see that all consumers in Φ_0 would enter at this price simply notice that the optimal price in Corollary B.2 for consumers at the boundary of Φ_0 is exactly $p_x + c_z - c_x$. Also similar to Lemma 3.1 is the discussion about profit levels from selling z for case (i) and (ii). They are, respectively $\Pi(z, \underline{p}_z) = \int_{\Phi_0} (p_x - c_x) d\phi_i$ and $\Pi(z, \overline{p}_z) = 0$.

Finally, notice that the structure of interval $\Phi(p'_z)$ is unchanged. Hence, in order to find p'_z, the monopolist solves (3.14). When $c_z > c_x$, there is no flat portion of the willingness to pay curve in the set of "valuable" consumers — i.e. the ones to which it is profitable to sell z. To see this, simply notice that $\frac{p_x}{v(x)} \notin \Phi_0$, since $\frac{p_x}{v(x)} > \frac{(p_x + c_x) - c_z}{v(x)}$

when $c_z > c_x$.

This does not imply, however, that the maximisation in (3.14) will always identify the correct maximum. The problem ignores that beyond the boundaries of Φ_0 it is unprofitable to sell good z. Hence, if the solution of the problem is lower than $p_x + c_z - c_x$, the retailer sets $p_z = p_x + c_z + c_x$ to sell z to all consumers in Φ_0 and sell x to all other consumers. It can be shown that:

$$\begin{aligned} &p'_z \geq p_x + c_z - c_x \quad \text{is equivalent to} \\ &c_z \leq \frac{2c_x(2v(x) - v(z))}{v(z) - v(x)} \end{aligned} \tag{B.15}$$

The next Corollary shows how Proposition 3.2 changes with $c_z > c_x$.

Corollary B.3. *Let $c_z > c_x$. If the monopolist cannot observe consumers' level of temptation in any way, she sets p'_z as in (3.14) if the cost of producing z is low enough, i.e. condition (B.15) holds. She sets $p_z = p_x + c_z - c_x$ otherwise. Consumers behave as in Proposition 3.2.*

B.7.3 Loyalty Scheme Technology

The modifications to the equilibrium of the model with loyalty schemes under the case of $c_z > c_x$ are quite simple to identify. First of all, recall that the monopolist is not willing to sell good z to consumers that are outside interval Φ_0. It is easy to see then that the algorithm described in Section 3.6 applies entirely to the case of $c_z > c_x$. Given partition $\mathscr{P}_n$, however, the monopolist will use the algorithm, and charge the according price p_A, only to consumers that belong to elements of the partition that are inside, or at least have a non empty intersection with, Φ_0. Elements of the partition that are outside Φ_0 are irrelevant for the equilibrium definition. The result is presented in the following Corollary of Proposition 3.9.

Corollary B.4. *Let $c_z > c_x$. If the monopolist can offer a loyalty card scheme that partitions the consumers's space with $\mathscr{P}_n$, in equilibrium she charges p_z equal to:*

(i) $p_R = p_{max}$ *if i rejects the loyalty card.*

(ii) p_A *as described by (3.19)-(3.28) if i accepts the loyalty card and proves to belong to an element of $\mathscr{P}_n$ that is a subset of Φ_0.*

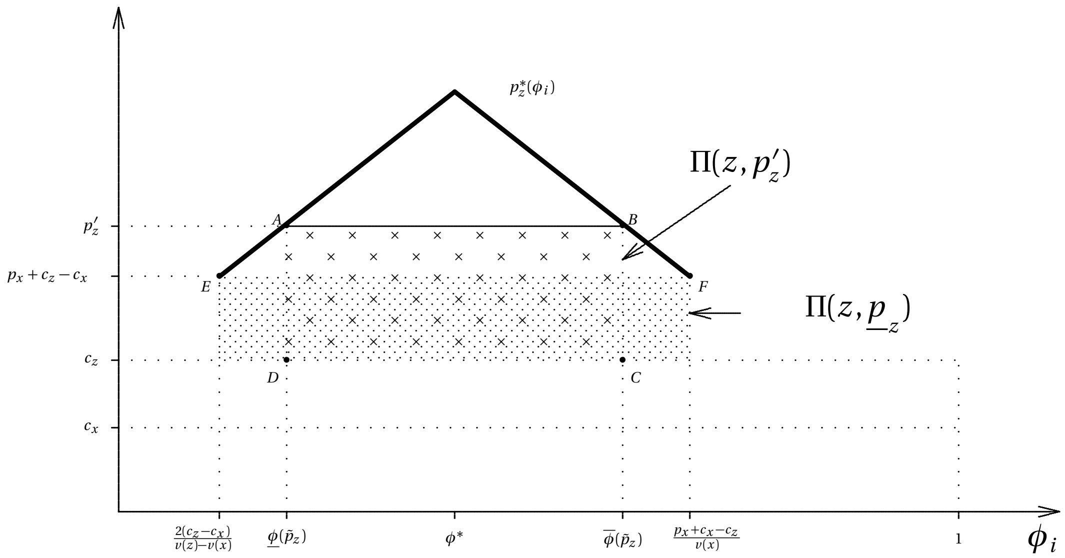

Figure B.2: The case of $c_z > c_x$

The Figure shows how results in Figure 3.1 and 3.2 change when $c_z > c_x$ is assumed. It shows the value of first-best price $p_z^*(\phi_i)$ (the bold curve from E to F). Under first-best, the monopolist obtains profits (from selling z) equal to the area between the bold curve and the c_z line. When information becomes asymmetric, she charges p_z' if (B.15) holds and $\underline{p}_z = p_x + c_z - c_x$ if not.

(iii) $p_A = p_x + c_z + c_x$ *if i accepts the loyalty card and proves to belong to an element of $\mathscr{P}_n$ that is outside of Φ_0.*

(iv) p_A *as described by (3.19)-(3.28) with the left and right cutoffs replaced by:*

$$\max\left\{p_z^*(\phi_l), p_x + c_z - c_x\right\} \quad \textit{and} \tag{B.16}$$

$$\max\left\{p_x + c_z - c_x, p_z^*(\phi_r)\right\}, \tag{B.17}$$

if i accepts the loyalty card and proves to belong to an element of $\mathscr{P}_n$ that intersects with Φ_0.

Consumers to the left of Φ_0 have the (weak) incentive to enter the superstore to buy good x. Consumers to the right of Φ_0 have the strict incentive to reject the loyalty card in order to enter the superstore to buy good x.

Appendix C

Appendix to Chapter 4

C.1 Generality of $v_H(0) = v_L(0) = u(0)$

Given that the high type is more tempted than the low type, an interesting alternative assumption would be $v_H(0) > v_L(0) = u(0)$ — or, equivalently, $v_H(0) > u(0) > v_L(0)$; I study the former case for simplicity. In this case, function v_H lies strictly above v_L for every q. This has no qualitative implications on the result. Suppose $v_H(0) = \delta > 0$. In the first best, the optimal offer for the low type does not vary. The one for the high type changes in the following sense: the ex-post tariff and the ex-ante entry bonus decrease by $\delta/2$. Notice, this new assumption is equivalent to an ex-post outside option for the high type, since now offer *0* grants him an ex-post utility of δ. Qualitatively nothing changes: both types obtain ex-ante and ex-post utility equal to the one they would obtain from their outside option. Quantitatively, the ex-post surplus of the high type is now positive and equal to $v_H(0)$.
If information is, instead, asymmetric then problem (4.7) becomes:

$$\max_M \pi(M) = \pi(x_L)\beta + \pi(x_H)(1-\beta)$$

$$\text{s.t.} \quad U(x_H) + V_H(x_H) \geq \delta \qquad (\overline{PC}_H)$$

$$U(x_L) + V_L(x_L) \geq 0 \qquad (\overline{PC}_L)$$

$$U(x_L) + V_L(x_L) \geq U(x_H) + V_L(x_H) \qquad (\overline{IC}_L)$$

$$U(x_H) + V_H(x_H) \geq U(x_L) + V_H(x_L). \qquad (\overline{IC}_H)$$

However, the maximisation can be solved in the exact same way as in Section 4.4. It remains to check whether the participation constraint of the high type holds with this new outside option. Plugging-in the solution from (4.20), I get: $U(x_H^S) + V_H(x_H^S) =$

$v_H(q_L^S) - v_L(q_L^S)$. Since $v_H \succ v_L$ and $v_H(0) - v_L(0) = \delta$, then $U(x_H^S) + V_H(x_H^S) > \delta$.

C.2 Proof of Proposition 4.1

By backward induction, I start from the ex-post problem (4.8). Let the ($\overline{PC}_i$) bind, then $t_i^* = \frac{1}{2}\left[u(q_i) + v_i(q_i)\right]$. Substitute this back into (4.8) and solve for q_i^* and t_i^*. Moving, then, to the ex-ante problem:

$$\max_{F_i} \Pi^S = \max_{F_i} \left[\pi(x_i^*) + F_i\right] \tag{C.1}$$

$$\text{s.t. } W_i(\{0, x_i^*\}, F_i) \geq 0 \tag{$\underline{PC}_i$}$$

$$\text{for } i = L, H.$$

Since F enters with a negative sign in the constraint, and with a positive sing in the profit, let the participation constraint bind to obtain F_H^* and F_L^*.

C.3 Proof of Lemma 4.3

I will show that only PC_L and IC_H bind while PC_H and IC_L are redundant — I omit the upper bar on constraints, but all constraints are ex-post ones.

First of all, since $V_H \succ U \succ V_L$, PC_L implies that $U(x_L) + V_H(x_L) > 0$ which, along with IC_H, implies that PC_H is slack.

Constraint PC_L, instead, has to bind at the solution. Suppose this is not true then the monopolist can increase t_L and t_H of an amount $\epsilon > 0$ such that PC_L binds, not affecting the incentive compatibility constraints, raising her profits.

Similarly for IC_H: if it is slack, the monopolist can increase t_H by $\epsilon > 0$ such that IC_H binds, not affecting PC_L, relaxing IC_L, and raising profits.

Finally, consider IC_L. Suppose it is not redundant and suppose the solution to the reduced problem, subject only to PC_L and IC_H, is given by two different offers x'_H and x'_L. If IC_L is not redundant then the low type would be at least as happy to buy x'_H as to buy x'_L. Since IC_H binds, however, also the high type is as happy to buy x'_L as to buy x'_H. This means that the monopolist would be better off by simply offering the offer x'_i such that $\pi(x'_i) > \pi(x'_j)$. This contradicts two distinct offers, as x'_H to x'_L, to be the solution to profit maximisation. Since problem (4.19) in the paper yields two distinct solutions, IC_L can be considered slack and checked afterwards.

C.4 Proof of Lemma 4.4

Since c is increasing in q and such that $c(0)=0$, in order for q_L^S to be positive, the slope $c'(q_L^S)$ has to be positive. This happens when the LHS in the definition of q_L^S from (4.24) is positive. Hence:

$$\left[(u'(q)+v_H'(q))\right] \geq \tfrac{1}{\beta}\left[(v_H'(q)-v_L'(q))\right] \Rightarrow \beta \geq \frac{v_H'(q_L^S)-v_L'(q_L^S)}{u'(q_L^S)+v_H'(q_L^S)} \equiv \underline{\beta}.$$

C.5 Proof of Lemma 4.5

$q_H^* = q_H^S$ is obvious. Recall that $c(q)$ is increasing and convex in q, all utility functions are increasing and concave in q and that q_H^S, q_L^*, q_L^S are described by the equations in (4.9), (4.20) and (4.24).

To see that $q_H^S > q_L^*$ suppose the contrary, $q_H^S \leq q_L^*$. Then $c'(q_H^S) \leq c'(q_L^*)$, $u'(q_H^S) \geq u'(q_L^*)$ and $v_H'(q_H^S) \geq v_H'(q_L^*)$. Also $v_H'(q_L^*) > v_L'(q_L^*)$ since $v_H \succ v_L$. This makes:

$$u'(q_H^S)+v_H'(q_H^S) > u'(q_L^*)+v_L'(q_L^*) \quad \text{and} \quad c'(q_H^S) \leq c'(q_L^*)$$

which contradicts (4.9) and (4.20). To see that $q_L^* = q_L^S$ is possible, notice that, for $\beta = 1$, (4.24) is identical to (4.9).

A similar proof for $q_L^* \geq q_L^S$ is possible. Suppose this is not true, i.e. $q_L^* < q_L^S$. Then $c'(q_L^S) > c'(q_L^*)$, $u'(q_L^S) < u'(q_L^*)$ and $v_i'(q_L^S) < v_i'(q_L^*)$ for $i = H, L$. I will show that his brings to a contradiction since when $q_L^* < q_L^S$, $c'(q_L^S) > c'(q_L^*)$ cannot happen. The latter is true if:

$$\tfrac{1}{2\beta}\left[(u'(q_L^S)+v_H'(q_L^S))\beta - (v_H'(q_L^S)-v_L'(q_L^S))\right] > \tfrac{1}{2}(u'(q_L^*)+v_L'(q_L^*))\beta$$

which can be rearranged as:

$$(u'(q_L^S)-u'(q_L^*))\beta + \left[v_L'(q_L^S)-v_H'(q_L^S)(1-\beta)-v_L'(q_L^*)\beta\right] > 0$$

which never holds. To see this, notice that the first element is negative, since $u'(q_L^S) <$

$u'(q_L^*)$, and the second element is also negative, since $v_L'(q_L^S) < v_L'(q_L^*)$ and $v_L'(q_L^S) < v_H'(q_L^S)$.

C.6 Proof of Lemma 4.6

Define $\tilde{x} = (\tilde{q}, \tilde{t})$ where $\tilde{q} = q_L^* + \epsilon$ and $\tilde{t} = \frac{1}{2}\left[u(\tilde{q}) + v_L(\tilde{q}) + \delta\right]$. From Lemma 2, $q_H^* > q_L^*$, hence there exist a small $\epsilon > 0$ such that $q_H^* > \tilde{q} > q_L^*$. Then it is easy to see that:

$$U(\tilde{x}) + V_L(\tilde{x}) = -\delta \tag{C.2}$$

$$U(\tilde{x}) + V_H(\tilde{x}) = v_H(\tilde{q}) - v_L(\tilde{q}) - \delta \tag{C.3}$$

$$\pi(\tilde{x}) = \tfrac{1}{2}\left[u(\tilde{q}) + v_L(\tilde{q})\right] - c(\tilde{q}) + \tfrac{1}{2}\delta. \tag{C.4}$$

To show that pooling is never an equilibrium, I prove that $\tilde{x}$ is not chosen by the low type in $\{0, x^P, \tilde{x}\}$, but it is chosen by the high type, and yields strictly higher profits for the monopolist—recall that $x^P = x_L^*$. Formally, there exist a $\delta > 0$ such that:

$$U(\tilde{x}) + V_L(\tilde{x}) \leq 0 \tag{C.5}$$

$$U(\tilde{x}) + V_H(\tilde{x}) \geq U(x^P) + V_H(x^P) \tag{C.6}$$

$$\pi(\tilde{x}) > \pi(x^P). \tag{C.7}$$

Equation (C.5) holds by (C.2) and the positivity of δ. Let (C.6) bind, then:

$$\delta = v_H(\tilde{q}) - v_L(\tilde{q}) - \left[v_H(q_L^*) - v_L(q_L^*)\right] > 0 \tag{C.8}$$

by definition of $\tilde{q}$. Substituting δ into (C.7) I get:

$$\left[\tfrac{1}{2}\left[u(\tilde{q}) + v_L(\tilde{q})\right] - c(\tilde{q})\right] - \left[\tfrac{1}{2}\left[u(q_L^*) + v_L(q_L^*)\right] - c(q_L^*)\right] > 0 \tag{C.9}$$

which is always true by definition of $\tilde{q}$. This concludes the proof of the Lemma.

C.7 Proof of Proposition 4.9

The monopolist has two decisions to make. Ex-ante, she decides whether to exclude the high type or not. If she does not, then ex-post she decides whether to separate or exclude the low type. Solving the game backwards, I consider this latter choice first. This is described by (4.25). If (4.25) fails, then, in period 1, the monopolist compares the ex-ante profits of excluding the high type ex-ante with the ones of excluding the low type ex-post. Comparing (4.16) with (4.27) I obtain (4.30) where

$$\beta_{EH}^{EL} \equiv \frac{2\left[u(q_H^*)-c(q_H^*)\right]}{u(q_L^*)+v_L(q_L^*)-2c(q_L^*)+u(q_H^*)+v_H(q_H^*)-2c(q_H^*)}.$$

If (4.25) holds, instead, then he compares the ex-ante profits of excluding the high type in period 1 and the ones of serving both types ex-post. Comparing (4.16) with (4.29) I obtain (4.31) where

$$\beta_{EH}^{S} \equiv \frac{u(q_H^S)-2c(q_H^S)+\left(v_H(q_H^S)-v_H(q_L^S)+v_L(q_L^S)\right)\mathbb{I}\left[U(x_H^S)>0\right]}{2\pi(x_H^S)+u(q_L^*)+v_L(q_L^*)-2c(q_L^*)-\left[u(q_L^S)+v_L(q_L^S)-2c(q_L^S)\right]}.$$

C.8 Figure 4.3 and Ex-post Equilibrium Analysis

The next four Claims explain the shape of the profits in Figure 4.3.

I first show that $\pi^S(\beta)$ is decreasing and convex in β. Then I show what happens to x_H^S, x_L^S and $\pi^S(\beta)$ at the extreme values of $\beta \in \left[\underline{\beta},1\right]$. Claim C.4 contains a technical requirement for the result.

Claim C.1. *Profit $\pi^{EL}(\beta)$ is a linearly decreasing function of β while $\pi^S(\beta)$ is decreasing and convex in β.*

Proof. While the first is trivial, to see that the latter is true notice that:

$$\frac{\partial \pi^S}{\partial \beta} = \left\{\tfrac{1}{2}\left[u(q_L^S)+v_H(q_L^S)\right]-c(q_L^S)\right\}-\left\{\tfrac{1}{2}\left[u(q_H^*)+v_H(q_H^*)\right]-c(q_H^*)\right\} \tag{C.10}$$

where I used the fact that $\frac{1}{2}\left[v_L'(q_L^S)+u'(q_L^S)\beta-v_H'(q_L^S)(1-\beta)\right]-c'(q_L^S)=0$, by definition of q_L^S. Hence:

$$\frac{\partial \pi^S}{\partial \beta}<0 \quad \text{for all}\, \beta.$$

Moreover, since q_H^S is independent of β:

$$\frac{\partial^2 \pi^S}{\partial \beta^2} > 0 \quad \text{for all}\, \beta \tag{C.11}$$

since q_L^S is increasing in β. This proves the claim. □

Claim C.2. *At $\beta = \underline{\beta}$, $x_H^S = x_H^*$, $x_L^S = 0$ and $\pi^S\left(\underline{\beta}\right) = \pi^{EL}\left(\underline{\beta}\right)$.*
At $\beta = 1$, $x_L^S = x_L^*$ and $\pi^S(1) = \pi^P$ (x_H^S is irrelevant).

Proof. Notice that $q_H^S = q_H^*$ and $t_L^S = t_L^*$ for all β. Moreover, at $\beta = \underline{\beta}$, $q_L^S = 0$ which makes $t_H^S = t_H^*$ and $t_L^S = 0$. Therefore the profit from the low type is zero and $\pi^S\left(\underline{\beta}\right) = \pi(x_H^*)\left(1 - \underline{\beta}\right) = \pi^{EL}\left(\underline{\beta}\right)$. This proves the first part.
At $\beta = 1$, instead, $q_L^S = q_L^*$ proving $x_L^S = x_L^*$. Moreover, since the probability of the consumer to be high type is zero, $\pi^S(1) = \pi(x_L^*) = \pi^P$. This concludes the proof of the claim. □

Claim C.3. *When $\beta = \underline{\beta}$, separation yields higher profits than pooling.*

Proof. There is a simple way to prove this. Notice that, from claim C.2, at $\beta = 1$, $\pi^S(1) = \pi^P$. Also, from claim C.1, π^S is decreasing in β. Since $\underline{\beta} < 1$ it must then be that $\pi^S\left(\underline{\beta}\right) > \pi^P$. □

Claim C.4. *At $\beta = \underline{\beta}$:*

$$\frac{\partial \pi^S}{\partial \beta}\left(\underline{\beta}\right) = \frac{\partial \pi^{EL}}{\partial \beta}\left(\underline{\beta}\right).$$

Hence $\pi^{EL}(\beta)$ is tangent to $\pi^S(\beta)$ in $\beta \in \left[\underline{\beta}, 1\right]$. The point of tangency is $\underline{\beta}$.

Proof. The proof is straightforward. π^{EL} is linear so its slope does not depend on β and it is equal to $-\pi(x_H^*)$. It follows from the fact that $q_L^S = 0$ at $\beta = \underline{\beta}$ that (C.10), evaluated at $\underline{\beta}$, is, also equal to $-\pi(x_H^*)$. To prove that the two functions intersect only at $\underline{\beta}$, where π^{EL} is tangent to π^S, simply notice that, from claim C.1, the slope of π^S is strictly larger — less negative — than the slope of π^{EL} for every $\beta \in \left(\underline{\beta}, 1\right]$. □

These four Claims prove the representation of π^{EL}, π^S and π^P in Figure 4.3.

Because of Claims C.1-C.4, π^{EL} is tangent to π^S in the interval $\left[\underline{\beta}, 1\right]$. Hence, π^{EL} lies below π^S for all $\beta \in \left[\underline{\beta}, 1\right]$.

This implies that, when $\beta \in \left[\underline{\beta}, 1\right]$, a further increase in the parameter value makes optimal separation more attractive. When $\beta \in \left[0, \underline{\beta}\right)$, of course, an increase in the parameter will, eventually, make separation possible.

Notice that this also implies that when pooling yields higher profits that excluding the low type, separation yields the highest possible profits. Therefore, $\pi^P < \max\{\pi^S, \pi^{EL}\}$. This is an alternate argument to Lemma 4.6 to show that pooling is never an equilibrium.

C.9 Temptation in a Multi-self Model

As mentioned in Section 4.3, the temptation aspect of decision-making can be also modelled as in a multi-self model (Strotz, 1956) with self one's preferences $U(x)$ and self two's preferences $U(x) + V_i(x)$. A temptation model like the one considered in this paper, however, endogenises the change from a classical model with consumers's preferences $U(x) + V_i(x)$ to a multi-self model. In this sense, the multi-self model is a special case of a model à la Gul and Pesendorfer (2001) (hereafter called *temptation model*) where consumers are always tempted by the offer they choose in period 2. Temptation models, instead, have the ability to account for the self-control cost the consumer bears. In other words, they account for the possibility of an offer non chosen ex-post to tempt the consumer and affect his choice ex-ante. This cannot happen in a multi-self model. Below I show how the equilibrium of a multi-self model like the one described is qualitatively different from the one derived in this paper. For the following analysis, notice that the ex-post problem of the seller does not change since period 2 preferences are still given by $U(x) + V_i(x)$.

Let information about types be private. First, suppose it is optimal for the monopolist to exclude the low type ex-post. The equilibrium does not change with respect to a temptation model. This is because the low type is always happy to enter when he chooses offer *0* ex-post. The high type, instead is tempted by x_H^* and, therefore, his period 1 utility is the same regardless of which model I consider. The entry fee is set to $F_H^* = U(x_H^*) < 0$.

Notice now, that an offer z that tempts the consumer, but is not chosen by him, ex-post (i.e. of the type described in Section 4.4.2) would have no effect on his behaviour in a multi-self model. Hence in equilibrium, in a multi-self model, the high type always enters ex-ante since there is no way for the monopolist to "decorate" the menu designed for the low type.

Finally, the separation equilibrium strongly differs from the one described in the paper. The downward tempted consumer in a temptation model is *not* tempted by the offer he chooses ex-post, x_L^S. Hence, he behaves as a classical consumer not affected by self-control problems and decides whether or not to enter the store according to $U(x) + V_i(x)$. In the multi-self model presented in this appendix, instead, he evaluates entrance according to $U(x)$. Hence, his ex-ante utility from the separation menu is given by $W_L(M_S, 0) = U(x_L^S) \geq 0$. Recall from Section 4.4.2 that the sign of the ex-ante utility of the high type depends on condition (4.28). Let the condition hold, then the ex-ante utility of both types is positive and the monopolist can extract this surplus with a positive entry fee: $F_S^* = min\{U(x_L^S), U(x_H^S)\} \geq 0$.

To conclude, a multi-self model provides qualitatively different results from the ones derived in the paper. These results are, inevitably, less general, since a multi-self model would fail to account for the self-control cost of decision making and assume exogenously the difference in ex-ante and ex-post preferences. The temptation model, instead, endogenises this difference and provides more general results.

Bibliography

ACKERMANN, P. (2010): "Loyalty Rewards and Monopoly Pricing," *Department of Economics, Universität Bern.*

AMADOR, M., I. WERNING, AND G. M. ANGELETOS (2006): "Commitment Vs. Flexibility," *Econometrica*, 2, 365–396.

ANDERSON, S. P., AND R. RENAULT (1999): "Pricing, Product Diversity and Search Cost: A Bertrand-Chamberlin-Diamond Model," *The RAND Journal of Economics*, 30(4), 719–735.

——— (2006): "Avertising Content," *The American Economic Review*, 96(1), 93–113.

ASHEIM, G. B. (2008): "Procrastination, Partial Naïvete, and Behavioral Welfare Analysis," *Department of Economics, University of Oslo.*

BAGUES, M., AND M. J. PEREZ-VILLADONIGA (2012): "Do Recruiters Prefer Applicants with Similar Skills? Evidence from a Randomized Natural Experiment," *Journal of Economic Behavior & Organization*, 82, 12–20.

BANKS, D. L., A. E. LAWSON, AND M. LOGVIN (2007): "Self-efficacy, Reasoning Ability and Achieving in College Biology," *Journal of Research in Science Teaching*, 44(5), 706–724.

BANNER, M., D. DUNNING, J. EHRLINGER, K. JOHNSON, AND J. KRUGER (2008): "Why the Unskilled are Unaware: Further Exploration of (Absent) Self-insight among the Incompetent," *Organizational Behavior and Human Decision Processes*, 105(1), 98–121.

BASSO, L. J., M. T. CLEMENTS, AND T. W. ROSS (2009): "Moral Hazard and Customer Loyalty Programs," *American Economic Journal: Microeconomics*, 1(1), 101–123.

BURDETT, K., AND K. L. JUDD (1983): "Equilibrium Price Dispersion," *Econometrica*, 51(4), 955–969.

BYROM, J. (2001): "The Role of Loyalty Card Data within Local Marketing Initiatives," *International Journal of Retail & Distribution Management*, 29(7), 333–342.

CAMINAL, R. (2012): "The Design and Efficiency of Loyalty Rewards," *Journal of Economics & Management Strategy*, 21(2), 339–371.

CAMINAL, R., AND A. CLAICI (2007): "Are Loyalty-Rewarding Pricing Schemes Anti-Competitive?," *International Journal of Industrial Organization*, 25(4), 657–674.

CHEN, Y., AND J. PEARCY (2010): "Dynamic Pricing: When to Entice Brand Switching and When to Reward Consumer Loyalty," *The RAND Journal of Economics*, 41(4), 674–685.

CHI, M., T. GLASER, AND R. REES (1982): "Expertise in Problem Solving," *Pittsburgh University*.

CHRISTENSEN, E. G. B., AND J. NAFZIGER (2016): "Packaging of sin goods–Commitment or exploitation?," *Journal of Economic Behavior & Organization*, 122, 62–74.

CONLIN, M., T. O'DONOGHUE, AND T. J. VOGELSANG (2007): "Projection Bias in Catalog Orders," *The American Economic Review*, 97, 1217–49.

COURTY, P., AND H. LI (2000): "Sequential Screening," *Review of Economic Studies*, 67, 697–717.

DEB, R., AND M. SAID (2015): "Dynamic Screening with Limited Commitment," *Journal of Economic Theory*, 159, 891–928.

DIAMOND, A. P. (1971): "A Model of Price Adjustment," *Journal of Economic Theory*, 3(2), 156–168.

DITTRICH, D. A., W. GÜTH, AND B. MARCIEJOVSKY (2005): "Overconfidence in Investment Decisions: an Experimental Approach," *The European Journal of Finance*, 11(6), 471–491.

DUNNING, D., J. EHRLINGER, K. JOHNSON, AND J. KRUGER (2003): "Why People Fail to Recognize Their Own Incompetence," *Current Directions in Psychological Science*, 12(3), 83–87.

DUNNING, D., AND J. KRUGER (1999): "Unskilled and Unaware of It: How Difficulties in Recognising One's Own Incompetence Lead to Inflated Self-assessments," *Journal of Personality and Social Psychology*, 77, 1121–1134.

ELIAZ, K., AND R. SPIEGLER (2006): "Contracting with Diversely Naïve Agents," *Review of Economic Studies*, 73, 689–714.

——— (2008): "Consumer Optimism and Price Discrimination," *Theoretical Economics*, 3, 459–497.

ESTEBAN, S., AND E. MIYAGAWA (2005): "Optimal Menu of Menus with Self-control Preferences," *NAJ Economics*.

——— (2006): "Temptation, Self-Control, and Competitive Nonlinear Pricing," *Economic Letters*, 90(3), 348–355.

ESTEBAN, S., E. MIYAGAWA, AND M. SCHUM (2007): "Nonlinear Pricing with Self-Control Preferences," *Journal of Economic Theory*, 135(1), 306–338.

EVANS, R., AND S. REICHE (2015): "Contract Design and Non-Cooperative Renegotiation," *Journal of Economic Theory*, 157, 1159–1187.

FERRARO, P. J. (2010): "Know Thyself: Competence and Self-awareness," *Atlantic Economic Journal*, 38, 183–196.

FOSCHI, M. (2014): "Asymmetric Information, Commitment and Self-Control," *Department of Economics, University of Leicester*.

——— (2015): "Price Discrimination in the Retailing Industry," .

GALPERTI, S. (2015): "Commitment, flexibility, and optimal screening of time inconsistency," *Econometrica*, 83(4), 1425–1465.

GILPATRIC, M. S. (2008): "Present-bias Preferences, Self-awareness and Shirking," *Journal of Economic Behaviour and Organisation*, 67, 735–754.

GOMEZ, B. G., A. G. ARRANZ, AND J. G. CILLÀN (2006): "The Role of Loyalty Programs in Behavioral and Affective Loyalty," *Journal of Consumer Marketing*, 23(7), 387–396.

GREENLEE, P., D. REITMAN, AND D. S. SIBLEY (2008): "An Antitrust Analysis of Bundled Loyalty Discounts," *Internartional Journal of Industrial Organization*, 26(5), 1132–1152.

GRUBB, M. D. (2009): "Selling to Overconfident Consumers," *The American Economic Review*, 99, 1770–1807.

GUL, F., AND W. PESENDORFER (2001): "Temptation and Self-Control," *Econometrica*, 69(6), 1403–1435.

HANNAK, A., G. SOELLER, D. LAZER, A. MISLOVE, AND C. WILSON (2014): "Measuring Price Discrimination and Steering on E-Commerce Web Sites," *Proceedings of the 2014 Conference on internet measurement conference*, pp. 305–318.

HARRIS, M., AND A. RAVIV (1979): "Optimal Incentive Contracts with Imperfect Information," *Journal of Economic Theory*, 20, 231–259.

HEIDHUES, P., AND B. KÖSZEGI (2010): "Exploiting Naïvete About Self-control in the Credit Market," *The American Economic Review*, 5, 2279 – 2303.

KOVÁC, E., AND D. KRÄHMER (2013): "Optimal Sequential Delegation," *SFB/TR 15 Discussion Paper Series*, 427.

KUMRU, C. S., AND A. C. THANOPOULOS (2008): "Social Security and Self-control Preferences," *Journal of Economic Dynamics and Control*, 32(3), 757–778.

LAFFONT, J. J., AND D. MARTIMORT (2002): *The Theory of Incentives*. Princeton University Press, New Jersey, United States of America.

LEWIS, M. (2004): "The Influence of Loyalty Programs and Short-Term Promotions on Customer Retention," *Journal of Marketing Research*, 41(3), 281–292.

LICHTENSTEIN, S., B. FISCHOFF, AND L. D. PHILLIPS (1982): *Calibration of Probabilities: The State of the Art to 1980*. Cambridge University Press, Cambridge, UK.

LOEWENSTEIN, G., T. O'DONOGHUE, AND M. RABIN (2003): "Projection Bias in Predicting Future Utility," *Quarterly Journal of Economics*, 118, 1209–48.

MEYER-WAARDEN, L. (2007): "The Effects of Loyalty Programs on Customer Lifetime Duration and Share of Wallet," *Journal of Retailing*, 83(2), 223–236.

MIKIANS, J., L. GYARMATI, V. ERRAMILLI, AND N. LAOUTARIS (2012): "Detecting Price and Search Discrimination on the Internet," *Proceedings of the 11th ACM Workshop on Hot Topics in Networks*.

——— (2013): "Crowd-Assisted Search for Price Discrimination in E-Commerce: First Results," *Proceedings of the 9th ACM conference on Emerging networking experiments and technologies.*

MOORE, D. A., AND P. J. HEALY (2008): "The Trouble with Overconfidence," *Psychological Review*, 115(2), 502.

NOWER, L., AND A. BLASZCZYNSKY (2004): "Impulsive and Pathological Gambling: A Descriptive Model," *International Gambling Studies*, 6(1), 61–75.

O'DONOGHUE, T., AND M. RABIN (2001): "Choice and Procrastination," *Quarterly Journal of Economics*, 1, 121–160.

RHODES, A. (2014): "Multi-product Retailing," *The Review of Economic Studies*, 2, 333–359.

ROTHSCHILD, M., AND J. STIGLITZ (1976): "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect information," *The Quarterly Journal of Economics*, 90(4), 629–649.

SHILLER, B., AND J. WALDFOGEL (2011): "Music for a Song: an Empirical Look at Uniform Pricing and its Alternatives," *The Journal of Industrial Economics*, 59(4), 630–660.

SHILLER, B. R. (2015): "Approximating Reservation Prices From Broad Consumer Tracking," *Department of Economics, Brandeis University.*

SPENCE, M. (1977): "Nonlinear Prices and Welfare," *Journal of Public Economics*, 8(1), 1–18.

STAHL, D. O. (1989): "Oligopolistic Pricing with Sequential Consumer Search," *The American Economic Review*, 79(4), 700–712.

STROTZ, R. (1956): "Myopia and Inconsistency in Dynamic Utility Maximization," *The Review of Economics Studies*, 23(3), 165–180.

SVENSON, O. (1981): "Are We Less Risky and More Skilful than Our Fellow Drivers?," *Acta Psychologica*, 47, 143–146.

TOWNSEND, R. M. (1979): "Optimal Contracts and Competitive Markets with Costly State Verification," *Journal of Economic Theory*, 21(2), 265–293.

VARIAN, H. R. (1980): "A Model of Sales," *The American Economic Review*, 70(4), 651–659.

VON THADDEN, E.-L., AND X. ZHAO (2012): "Incentives for Unaware Agents," *Review of Economic Studies*, 79, 1151–1174.

WALDFOGEL, J. (2015): "First Degree Price Discrimination Goes to School," *The Journal of Industrial Economics*, 63(4), 569–597.

WERNERFELT, B. (1994): "Selling Formats for Search Goods," *Marketing Science*, 13(3), 298–309.