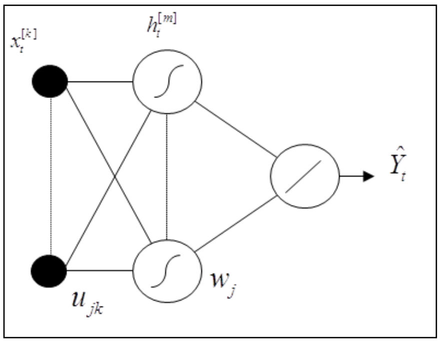
Online Appendix

A. Neural Networks and technical characteristics.

In this section, short descriptions of the MLP, RNN and PSN are presented, along with their input selection and parametrization for each ETF return series under study. Firstly, the typical MLP model is shown in the following figure.

Figure A.1: A single output, fully connected MLP model



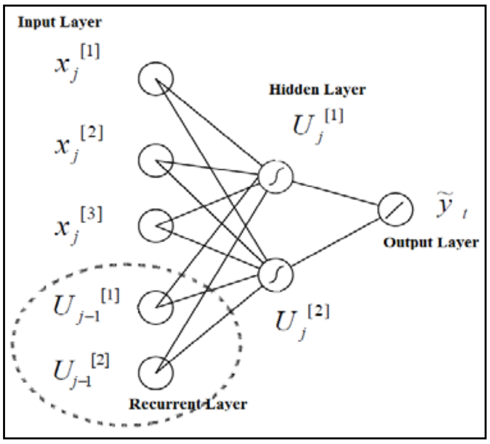
where:

- $x_t^{[n]}(n=1,2,\cdots,k+1)$ are the inputs (including the input bias node) at time t
- $h_t^{[m]}(m=1,2,...,j+1)$ are the hidden nodes outputs (including the hidden bias node) at time t
- $\hat{Y}_t$ is the MLP output
- u_{jk}, w_j are the network weights
- is the transfer sigmoid function $S(x)=1/(1+e^{-x})$
- is a linear function $F(x)=\sum_i x_i$

The Error Function to be minimized is $E(u_{jk},w_j)=\frac{1}{T}\sum_{t=1}^{T}\left(Y_t-\hat{Y}_t(u_{jk},w_j)\right)^2$, where Y_t is the target value.

The second NN applied in this study is the RNN. A simple illustration its architecture is presented below.

Figure A.2: RNN with two nodes in the hidden layer



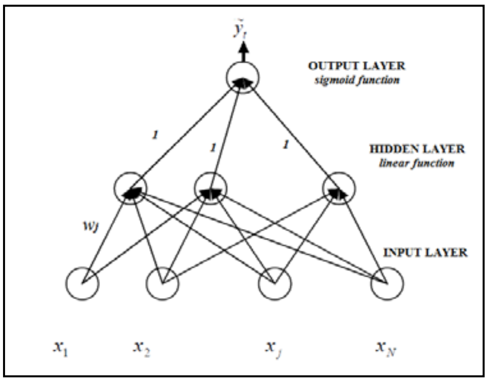
where:

- $x_t^{[n]}(n=1,2,\ldots,k+1), u_t^{[1]}, u_t^{[2]}$ are the RNN inputs at time *t* (including bias node)
- $\tilde{y}_t$ is the output of the RNN
- $d_t^{[f]}(f=1,2)$ and $w_t^{[n]}(n=1,2,\ldots,k+1)$ are the weights of the network
- $U_t^{[f]}, f=(1,2)$ is the output of the hidden nodes at time *t*
- is the transfer sigmoid function : $S(x)=1/(1+e^{-x})$
- is a linear function: $F(x)=\sum_i x_i$

The Error Function to be minimized is $E(d_t,w_t)=\frac{1}{T}\sum_{t=1}^{T}(y_t-\tilde{y}_t(d_t,w_t))^2$, where y_t is the target value. In short, the RNN architecture can provide more accurate outputs because the inputs are (potentially) taken from all previous values (see $U_{j-1}^{[1]}$ and $U_{j-1}^{[2]}$). For an exact specification of recurrent networks, see Elman (1990).

The third model is the PSN architecture, as shown in the figure below.

Figure A.3: A PSN with one output layer



where:

- x_t *(n=1,2,…,k+1)* are the model inputs (including the input bias node)
- $\tilde{y}_t$ are the PSN input and output respectively
- w_j *(j=1,2..,k)* are the adjustable weights (*k* is the desired order of the network)
- The hidden layer activation function $h(x)=\sum_i x_i$
- The output sigmoid activation function (*c* the adjustable term): $\sigma(x)=1/(1+e^{-xc})$

The Error Function minimized in this case $E(c,w_j)=\frac{1}{T}\sum_{t=1}^{T}(y_t-\tilde{y}_t(w_k,c))^2$, where y_t is the target value.

More details on the PSN model are given by Ghosh and Shin (1991).

Regarding the selection of inputs, there is no formal theory behind the selection of the NN inputs and their characteristics, such as number of hidden neurons, learning rate, momentum and iterations. For that reason, we conduct NN experiments and a sensitivity analysis on a pool of autoregressive terms of the return series in the in-sample dataset. In terms of our iterations, our experimentation starts from 5.000

iterations and stops at the 100.000 iterations. In each experiment the number of iterations is increased by 5.000, following cornerstone studies on NN training such as Tenti (1996) and Zhang *et al.* (1998). Based on the above, we select the inputs that provide the higher trading performance for each network in the in-sample period. The final sets of inputs of the three NNs for the three forecasting exercises are presented in table A.1 below:

Table A.1: Neural network inputs

SPY			DIA			QQQ		
MLP	**RNN**	**PSN**	**MLP**	**RNN**	**PSN**	**MLP**	**RNN**	**PSN**
SPY (1)*	SPY (1)	SPY (1)	DIA (2)	DIA (1)	DIA (1)	QQQ (1)	QQQ (1)	QQQ (2)
SPY (3)	SPY (2)	SPY (4)	DIA (4)	DIA (3)	DIA (2)	QQQ (2)	QQQ (4)	QQQ (4)
SPY (5)	SPY (3)	SPY (5)	DIA (5)	DIA (4)	DIA (5)	QQQ (3)	QQQ (5)	QQQ (5)
SPY (7)	SPY (5)	SPY (6)	DIA (7)	DIA (6)	DIA (6)	QQQ (5)	QQQ (6)	QQQ (6)
SPY (8)	SPY (7)	SPY (7)	DIA (9)	DIA (7)	DIA (8)	QQQ (6)	QQQ (7)	QQQ (7)
SPY (9)	SPY (8)	SPY (9)	DIA (10)	DIA (8)	DIA (9)	QQQ (8)	QQQ (9)	QQQ (8)
SPY (12)	SPY (9)	SPY (10)	DIA (11)	DIA (9)	DIA (10)	QQQ (10)	QQQ (10)	QQQ (9)
-	SPY (10)	SPY (11)	-	DIA (10)	-	QQQ (11)	QQQ (12)	QQQ (10)
-	SPY (12)	SPY (12)	-	-	-	QQQ (12)	-	QQQ (11)

Note: SPY(1) means that as input is used the SPY return series lagged by one day. Thus, today's return is used to forecast the tomorrow's one. The pool of potential inputs includes lags of daily returns running back to a month.

Table A.2 shows the training characteristics of all the above NN architectures for each forecasting exercise.

Table A.2: Neural network design and training characteristics

	Parameters	MLP	RNN	PSN
SPY	Learning algorithm	Gradient descent	Gradient descent	Gradient descent
	Learning rate	0.003	0.003	0.4
	Momentum	0.004	0.005	0.5
	Iteration steps	30000	40000	40000
	Initialisation of weights	N(0,1)	N(0,1)	N(0,1)
	Input nodes	7	9	7
	Hidden nodes	6	6	5
	Output node	1	1	1
DIA	Learning algorithm	Gradient descent	Gradient descent	Gradient descent
	Learning rate	0.002	0.005	0.3
	Momentum	0.005	0.006	0.5
	Iteration steps	45000	35000	40000
	Initialisation of weights	N(0,1)	N(0,1)	N(0,1)
	Input nodes	7	8	7
	Hidden nodes	9	7	6
	Output node	1	1	1
QQQ	Learning algorithm	Gradient descent	Gradient descent	Gradient descent
	Learning rate	0.003	0.002	0.3
	Momentum	0.005	0.005	0.4
	Iteration steps	30000	35000	25000
	Initialisation of weights	N(0,1)	N(0,1)	N(0,1)
	Input nodes	9	8	9
	Hidden nodes	8	10	8
	Output node	1	1	1

B. Statistical and trading performance measures.

The statistical and trading performance measures of the forecasting models are calculated as shown in table B.1 and B.2 respectively.

Table B.1: Statistical performance measures

Statistical performance measures	Description
Mean absolute error	$MAE = (\frac{1}{n})\sum_{\tau=t+1}^{t+n}\left\|\hat{Y}_\tau - Y_\tau\right\|$ *with* $\boldsymbol{Y}_\tau$ *being the actual value and* $\hat{\boldsymbol{Y}}_\tau$ *the forecasted value*
Mean absolute percentage error	$MAPE = \frac{1}{n}\sum_{\tau=t+1}^{t+n}\left\|\frac{Y_\tau - \hat{Y}_\tau}{Y_\tau}\right\|$
Root mean squared error	$RMSE = \sqrt{\frac{1}{n}\sum_{\tau=t+1}^{t+n}(\hat{Y}_\tau - Y_\tau)^2}$
Theil-U	$Theil-U = \frac{\sqrt{(\frac{1}{n}\sum_{\tau=t+1}^{t+n}(\hat{Y}_\tau - Y_\tau)^2}}{\sqrt{\frac{1}{n}\sum_{\tau=t+1}^{t+n}\hat{Y}_\tau^2} + \sqrt{\frac{1}{n}\sum_{\tau=t+1}^{t+n}Y_\tau^2}}$

Table B.2: Trading performance measures

Trading performance measures	Description
Annualised return after transaction costs	$R^A = 252 * \frac{1}{N} * (\sum_{t=1}^{N} R_t) - TC^A$ where R_τ the daily return and TC^A the annualized transaction cost
Annualised volatility	$\sigma^A = \sqrt{252} * \sqrt{\frac{1}{N-1} * \sum_{t=1}^{N}(R_t - \overline{R})^2}$
Sharpe ratio	$SR = \frac{R^A - R^f}{\sigma^A}$
Maximum drawdown	Maximum negative value of $\sum(R_t)$ over the period $MD = \underset{i=1,\cdots,t;t=1,\cdots,N}{Min}\left(\sum_{j=i}^{t} R_j\right)$

C. Two-asset portfolio optimization

This section summarizes the equivalent results obtained for two-asset portfolios formed by the respective ETFs. Table F.1 presents the performance of the three two-asset portfolios (equally weighted).

Table E.1: Equally weighted two-asset portfolios

	Realized return	Sharpe ratio	Sortino ratio	Max drawdown
SPY-DIA	8.765%	0.7602	1.0479	7.391%
SPY-QQQ	13.565%	1.0788	1.6217	8.281%
DIA-QQQ	11.899%	0.9733	1.4239	7.760%

The following tables present the optimization results for these portfolios. The results follow the same trend as in the case of the 1/N portfolio presented in the main text.

Table F.2: Performances of different trading strategies (Traditional M-V, two asset portfolios)

		Realized return	Sharpe ratio	Sortino ratio	Max drawdown
Panel A: Mean-Variance optimization without short-selling					
SPY-DIA	ARMA-DCC	9.94%	0.8989	1.3181	7.04%
	ARMA-ADCC	10.11%	0.9171	1.3391	7.04%
	ARMA-GAS	10.13%	0.9183	1.3338	7.04%
	RNN-DCC	14.09%	1.2222	1.8107	7.83%
	RNN-ADCC	14.11%	1.2254	1.8241	7.83%
	RNN-GAS	14.22%	1.2313	1.8274	7.83%
	PSN-DCC	14.24%	1.2325	1.8419	7.81%
	PSN-ADCC	14.25%	1.2338	1.8363	7.81%
	PSN-GAS	14.23%	1.2356	1.8542	7.81%
SPY-QQQ	ARMA-DCC	12.66%	1.0403	1.6595	7.60%
	ARMA-ADCC	12.64%	1.0406	1.6585	7.59%
	ARMA-GAS	12.82%	1.0686	1.698	7.59%
	RNN-DCC	22.77%	1.9217	3.0146	7.18%
	RNN-ADCC	22.79%	1.9218	3.0148	7.18%
	RNN-GAS	23.36%	1.959	3.0933	7.18%
	PSN-DCC	23.15%	1.9554	3.1586	7.86%
	PSN-ADCC	24.60%	2.0118	3.3553	7.74%
	PSN-GAS	24.75%	2.0088	3.3764	7.81%
DIA-QQQ	ARMA-DCC	12.33%	1.0441	1.7549	6.65%
	ARMA-ADCC	12.35%	1.0481	1.7215	6.65%
	ARMA-GAS	12.42%	1.052	1.7699	6.65%
	RNN-DCC	26.50%	2.1949	3.5555	6.75%
	RNN-ADCC	26.52%	2.196	3.5578	6.75%
	RNN-GAS	28.29%	2.3181	3.8093	6.75%
	PSN-DCC	28.30%	2.2244	3.8734	8.82%
	PSN-ADCC	28.43%	2.212	3.9033	8.82%
	PSN-GAS	28.48%	2.2145	3.9106	8.82%
Panel B: Mean-Variance optimization with short-selling					
		Realized return	Sharpe ratio	Sortino ratio	Max drawdown
SPY-DIA	ARMA-DCC	11.46%	1.0218	1.5845	7.52%
	ARMA-ADCC	11.63%	1.0373	1.5939	7.52%
	ARMA-GAS	11.64%	1.0387	1.5955	7.52%
	RNN-DCC	19.42%	1.6116	2.5722	8.72%
	RNN-ADCC	19.83%	1.6436	2.6235	8.72%
	RNN-GAS	19.97%	1.6541	2.6414	8.72%
	PSN-DCC	19.68%	1.624	2.6008	8.70%
	PSN-ADCC	20.05%	1.6501	2.6499	8.70%
	PSN-GAS	20.08%	1.652	2.6535	8.70%
SPY-QQQ	ARMA-DCC	13.51%	1.0436	1.6472	9.77%
	ARMA-ADCC	13.53%	1.0358	1.5882	9.83%
	ARMA-GAS	14.31%	1.0896	1.744	9.71%
	RNN-DCC	32.38%	2.6126	4.0333	7.41%
	RNN-ADCC	32.82%	2.6422	4.0887	7.41%
	RNN-GAS	33.25%	2.6648	4.1425	7.41%
	PSN-DCC	32.93%	2.6713	4.3888	7.40%
	PSN-ADCC	34.95%	2.7242	4.6567	7.48%

	PSN-GAS	35.05%	2.7119	4.6709	7.45%
	ARMA-DCC	12.82%	1.0803	1.8199	7.93%
	ARMA-ADCC	12.93%	1.1357	1.8279	7.93%
	ARMA-GAS	12.92%	1.1432	1.8456	7.93%
	RNN-DCC	41.76%	3.0091	5.505	8.30%
DIA-QQQ	RNN-ADCC	42.35%	2.9617	5.6068	8.52%
	RNN-GAS	42.37%	2.9627	5.6095	8.52%
	PSN-DCC	41.34%	3.069	5.6375	7.78%
	PSN-ADCC	42.36%	3.0892	5.6402	7.78%
	PSN-GAS	42.85%	3.1891	5.6378	7.78%

Note: The table presents the out-of-sample performances over the period January 2014 to March 2015 (68 weekly observations). Panel A reports performances of different M-V portfolios without short-selling. All the portfolios are weekly rebalanced tangency portfolios obtained by the M-V optimization based on various model combinations. For example, ARMA-DCC refers to the performance of the tangency portfolio of the efficient frontier of the three ETF assets, where the expected returns are obtained through ARMA forecasts, while the variance-covariance matrix is predicted by DCC. Panel B reports performances of different M-V portfolios with short-selling. '-S' denotes optimizations allowing short-selling.

Table F.3: Performances of different trading strategies (Mean-95% CVaR, two asset portfolios)

Panel A: Mean-CVaR optimization without short-selling

		Realized return	Return/CVaR	Sortino ratio	Max drawdown
	ARMA-DCC	11.43%	3.7593	1.5361	7.04%
	ARMA-ADCC	10.95%	3.6442	1.4451	7.04%
	ARMA-GAS	12.07%	3.8522	1.5516	7.04%
	RNN-DCC	15.52%	5.2265	1.936	7.59%
SPY-DIA	RNN-ADCC	15.61%	5.2449	2.1182	7.59%
	RNN-GAS	15.93%	5.4406	2.0632	7.59%
	PSN-DCC	16.19%	5.4931	2.1039	7.74%
	PSN-ADCC	16.25%	5.4176	2.1042	7.74%
	PSN-GAS	16.78%	5.6334	2.1945	7.79%
	ARMA-DCC	14.56%	4.3505	1.9341	7.60%
	ARMA-ADCC	14.53%	4.3512	1.8877	7.60%
	ARMA-GAS	14.56%	4.4526	1.9316	7.67%
	RNN-DCC	25.58%	8.3827	3.2877	7.10%
SPY-QQQ	RNN-ADCC	25.61%	8.4723	3.6059	7.11%
	RNN-GAS	25.68%	8.4913	3.4259	7.09%
	PSN-DCC	26.31%	8.5562	3.6079	8.09%
	PSN-ADCC	28.04%	8.8336	3.8447	8.06%
	PSN-GAS	28.35%	8.8972	3.8817	8.32%
	ARMA-DCC	13.51%	4.1587	1.8478	6.63%
	ARMA-ADCC	13.51%	4.1736	1.8661	6.65%
	ARMA-GAS	13.53%	4.1749	1.9175	6.65%
	RNN-DCC	29.19%	9.3864	3.8015	6.54%
DIA-QQQ	RNN-ADCC	29.16%	9.587	3.9141	6.54%
	RNN-GAS	30.19%	9.7552	4.096	6.44%
	PSN-DCC	32.16%	9.7333	4.4244	7.21%
	PSN-ADCC	32.41%	9.7125	4.4726	7.21%
	PSN-GAS	32.57%	9.7971	4.4875	7.23%

Panel B: Mean-CVaR optimization with short-selling					
		Realized return	Return/CVaR	Sortino ratio	Max drawdown
SPY-DIA	ARMA-DCC	13.64%	4.4741	1.8801	7.67%
	ARMA-ADCC	13.94%	4.5896	1.8746	7.67%
	ARMA-GAS	14.41%	4.7466	1.9429	7.67%
	RNN-DCC	20.93%	7.2314	2.738	8.30%
	RNN-ADCC	21.33%	7.446	2.6547	8.30%
	RNN-GAS	21.27%	7.3368	2.8372	8.30%
	PSN-DCC	21.52%	7.2268	2.8935	8.14%
	PSN-ADCC	21.93%	7.4467	2.9224	8.16%
	PSN-GAS	22.37%	7.4565	3.045	8.16%
SPY-QQQ	ARMA-DCC	15.77%	4.4802	1.9161	9.77%
	ARMA-ADCC	16.22%	4.583	1.8679	9.83%
	ARMA-GAS	17.72%	4.9791	2.1237	9.92%
	RNN-DCC	34.89%	11.7233	4.2933	7.05%
	RNN-ADCC	35.31%	11.9701	4.1373	7.05%
	RNN-GAS	35.42%	11.82	4.4496	7.05%
	PSN-DCC	35.29%	11.6545	4.787	7.46%
	PSN-ADCC	37.47%	12.0527	5.0349	7.47%
	PSN-GAS	38.28%	12.1199	5.2549	7.44%
DIA-QQQ	ARMA-DCC	14.96%	4.6376	2.117	7.93%
	ARMA-ADCC	15.51%	4.8249	2.1498	7.93%
	ARMA-GAS	15.69%	4.924	2.2474	7.93%
	RNN-DCC	45.00%	13.5023	5.8598	7.90%
	RNN-ADCC	45.56%	13.418	5.6734	7.20%
	RNN-GAS	45.13%	13.1415	6.0254	7.52%
	PSN-DCC	45.19%	13.6571	6.2721	7.97%
	PSN-ADCC	46.33%	13.941	6.2202	7.94%
	PSN-GAS	47.74%	13.9975	6.4695	7.91%

Note: The table presents the out-of-sample performances over the period January 2014 to March 2015 (68 weekly observations). Panel A reports performances of different mean-CVaR portfolios without short-selling. All the portfolios are weekly rebalanced tangency portfolios obtained by the different mean-CVaR optimization based on various model combinations. For example, ARMA-DCC refers to the performance of the tangency portfolio of the efficient frontier of the two ETF assets, where the expected returns are obtained through ARMA forecasts, while the variance-covariance matrix is predicted by DCC. Panel B reports performances of different mean-CVaR portfolios with short-selling. 'SKT' represents that the 95% CVaR is predicted using a Monte-Carlo simulation with the skewed t copulas to allow for asymmetric tail dependence '-S' denotes optimizations allowing short-selling.

Table F.4: Performances of different trading strategies (Mean-99% CVaR, two asset portfolios)

		Realized return	Return/CVaR	Sortino ratio	Max drawdown
Panel A: Mean-CVaR optimization without short-selling					
SPY-DIA	ARMA-DCC	11.78%	3.4047	1.5734	7.04%
	ARMA-ADCC	11.70%	3.4157	1.5434	7.04%
	ARMA-GAS	12.45%	3.4658	1.5837	7.04%
	RNN-DCC	15.29%	4.0588	1.9069	7.46%
	RNN-ADCC	15.93%	4.2188	2.1139	7.60%
	RNN-GAS	16.24%	4.3146	2.057	7.54%
	PSN-DCC	16.14%	4.3732	2.0984	6.75%
	PSN-ADCC	16.16%	4.2995	2.0921	6.70%
	PSN-GAS	16.65%	4.4627	2.1779	6.66%
SPY-QQQ	ARMA-DCC	15.01%	3.9402	1.9809	7.60%
	ARMA-ADCC	15.52%	4.0784	2.0161	7.60%
	ARMA-GAS	15.02%	4.006	1.9716	7.67%
	RNN-DCC	25.20%	6.5098	3.2382	6.98%
	RNN-ADCC	26.13%	6.8149	3.5987	7.12%
	RNN-GAS	26.17%	6.734	3.4156	7.04%
	PSN-DCC	26.48%	6.6134	3.4937	6.85%
	PSN-ADCC	27.07%	6.8062	3.7114	6.77%
	PSN-GAS	27.32%	6.843	3.7402	6.90%
DIA-QQQ	ARMA-DCC	13.92%	3.7665	1.8926	6.63%
	ARMA-ADCC	14.44%	3.9119	1.993	6.65%
	ARMA-GAS	13.96%	3.7561	1.9572	6.65%
	RNN-DCC	28.76%	7.2893	3.7443	6.43%
	RNN-ADCC	29.75%	7.7115	3.9063	6.55%
	RNN-GAS	30.77%	7.7363	4.0838	6.40%
	PSN-DCC	31.15%	7.5232	4.2842	6.10%
	PSN-ADCC	31.29%	7.4834	4.3175	6.06%
	PSN-GAS	31.38%	7.5351	4.3238	6.09%
Panel B: Mean-CVaR optimization with short-selling					
		Realized return	Return/CVaR	Sortino ratio	Max drawdown
SPY-DIA	ARMA-DCC	13.76%	3.7428	1.8935	7.67%
	ARMA-ADCC	14.08%	3.8401	1.8995	7.67%
	ARMA-GAS	14.27%	3.9107	1.9195	7.67%
	RNN-DCC	20.67%	4.9938	2.6985	8.15%
	RNN-ADCC	20.62%	5.0647	2.564	8.16%
	RNN-GAS	20.83%	4.9179	2.7181	8.14%
	PSN-DCC	21.03%	4.9702	2.8289	7.05%
	PSN-ADCC	21.55%	5.1152	2.8587	7.11%
	PSN-GAS	21.58%	5.1323	2.9847	7.03%
SPY-QQQ	ARMA-DCC	15.90%	3.7479	1.9297	9.77%
	ARMA-ADCC	16.39%	3.8346	1.8927	9.83%
	ARMA-GAS	17.55%	4.1023	2.0981	9.92%
	RNN-DCC	34.47%	8.0958	4.2313	6.92%
	RNN-ADCC	34.12%	8.1419	3.9961	6.95%
	RNN-GAS	34.68%	8.123	4.2628	6.91%
	PSN-DCC	34.49%	8.0153	4.68	6.37%
	PSN-ADCC	36.82%	8.2791	4.9251	6.42%
	PSN-GAS	36.93%	8.2796	5.151	6.42%

	ARMA-DCC	15.08%	3.8795	2.132	7.93%
	ARMA-ADCC	15.67%	4.037	2.1784	7.93%
	ARMA-GAS	15.54%	4.0569	2.2204	7.93%
	RNN-DCC	43.46%	9.3243	5.7753	7.75%
DIA-QQQ	RNN-ADCC	44.03%	9.1267	5.4797	6.28%
	RNN-GAS	44.19%	9.2088	5.7724	6.37%
	PSN-DCC	44.16%	9.3926	6.1319	6.81%
	PSN-ADCC	45.53%	9.5763	6.0846	6.90%
	PSN-GAS	46.05%	9.4468	6.3416	6.82%

Note: The table presents the out-of-sample performances over the period January 2014 to March 2015 (68 weekly observations). Panel A reports performances of different mean-CVaR portfolios without short-selling. All the portfolios are weekly rebalanced tangency portfolios obtained by the different mean-CVaR optimization based on various model combinations. For example, ARMA-DCC refers to the performance of the tangency portfolio of the efficient frontier of the two ETF assets, where the expected returns are obtained through ARMA forecasts, while the variance-covariance matrix is predicted by DCC. Panel B reports performances of different mean-CVaR portfolios with short-selling. 'SKT' represents that the 95% CVaR is predicted using a Monte-Carlo simulation with the skewed t copulas to allow for asymmetric tail dependence '-S' denotes optimizations allowing short-selling.