Cover sheet

Catania, L., Grassi, S., & Ravazzolo, F. (2019). Forecasting Cryptocurrencies Under Model and Parameter Instability. International Journal of Forecasting, 35(2), 485-501.
https://doi.org/10.1016/j.ijforecast.2018.09.005

Publication metadata

Title:	Forecasting Cryptocurrencies Under Model and Parameter Instability.
Author(s):	Catania, Leopoldo ; Grassi, Stefano ; Ravazzolo, Francesco.
Journal:	International Journal of Forecasting.
DOI/Link:	10.1016/j.ijforecast.2018.09.005
Document version:	Accepted manuscript (post-print)

Forecasting Cryptocurrencies Under Model and Parameter Instability

Leopoldo Catania[a], Stefano Grassi[b], Francesco Ravazzolo[c]

[a]*Department of Economics and Business Economics, Aarhus University and CREATES*
[b]*Department of Economics and Finance, University of Rome, Tor Vergata and CREATES*
[c]*Faculty of Economics and Management, Free University of Bozen-Bolzano, Italy, and CAMP, BI Norwegian Business School, Norway*

Abstract

This paper studies the predictability of cryptocurrencies time series. We compare several alternative univariate and multivariate models in point and density forecasting of four of the most capitalized series: Bitcoin, Litecoin, Ripple and Ethereum. We apply a set of crypto–predictors and rely on Dynamic Model Averaging to combine a large set of univariate Dynamic Linear Models and several multivariate Vector Autoregressive models with different forms of time variation. We find statistically significant improvements in point forecasting when using combinations of univariate models and in density forecasting when relying on selection of multivariate models. Both schemes deliver sizable directional predictability.

Keywords: Cryptocurrency; Bitcoin; Forecasting; Density Forecasting; VAR; Dynamic Model Averaging

1. Introduction

Bitcoin was the first decentralized cryptocurrency created in 2009 and documented in Nakamoto (2009). Since its introduction, it has gained a growing attention from the media, academics, finance industry and in recent months the global interest in Bitcoin and cryptocurrencies has spiked dramatically. There are numerous reasons for this intensified interest, and we just mention a few of them: Japan and South Korea have recognised Bitcoin as a legal method of payment (Bloomberg, 2017a; Cointelegraph, 2017); some central banks are exploring the use of the cryptocurrencies (Bloomberg, 2017c); a large number of companies and banks created the Enterprise Ethereum Alliance[1] to make use of the cryptocurrencies and the related technology called blockchain, (Forbes, 2017). Finally the Chicago Mercantile Exchange (CME) started the Bitcoin futures on 18 December 2017, see Chicago Mercantile Exchange (2017), and Nasdaq and Tokyo Financial Exchange will follow in 2018, see (Bloomberg, 2017b; Tokyo Financial Exchange, 2017).

This interest has been reflected on the market capitalization of the cryptocurrencies that exploded from around 19 billion in February 2017 to around 800 billion in December 2017 and more than 1000

[1]Source: https://entethalliance.org/members/

cryptocurrencies. Although Bitcoin is a relatively new currency, there has already been some initial analysis into the cryptocurrency. Hencic and Gourieroux (2014) applied a non–causal autoregressive model to detect the presence of bubbles in the Bitcoin/USD exchange rate. Sapuric and Kokkinaki (2014) measure the volatility of the Bitcoin exchange rate against six major currencies. Chu et al. (2015) provide a statistical analysis of the log–returns of the exchange rate of Bitcoin versus the USD. Catania and Grassi (2018) analyse the main characteristics of the cryptocurrency volatility. Hotz-Behofsits et al. (2018) apply to model cryptocurrencies a time-varying parameter VAR with t-distributed measurement errors and stochastic volatility.

Despite all this effort a detailed analysis of the forecasting performances of different models to this series has not been provided yet. This paper tries to fill this gap and compares a large set of different models for point and density forecasting of four of the most capatilised cryptocurrencies, precisely: Bitcoin, Litecoin, Ripple and Ethereum. We compare univariate autoregressive models to univariate linear regression models based on a large set of crypto–predictors. The predictors include commodity prices, other financial assets such as stock prices and bond prices, and volatility indices to proxy market sentiments, following evidence in Bianchi (2018) that returns on cryptocurrencies are moderately correlated (in-sample analysis) with commodities and few more financial assets. Moreover, we apply dynamic selection of the large set of models based on our predictor lists using Dynamic Model Selection (DMS) and dynamic averaging of the same model set using the Dynamic Model Averaging (DMA) methodology proposed by Raftery et al. (2010). DMS and DMA have been found to provide forecasting gains in macroeconomic applications, see for example Koop and Korobilis (2011) and Koop and Korobilis (2012), and have not yet been applied to cryptocurrencies. Then, we generalise the exercise to multivariate models where we predict jointly the four series using Vector Autoregressive (VAR) models, Bayesian VAR, timevarying parameters VAR with stochastic volatility as in Koop and Korobilis (2013), selecting and averaging of these models with different degrees of smoothness and different sets of predictors. See, among other, Stambaugh (1999), Pastor (2000), Pastor and Stambaugh (2000) and Barberis (2000) for the use of multivariate modelling and Bayesian inference in asset predictions and allocation, Dangl and Halling (2012) for application of model averaging to stock price prediction and Johannes et al. (2014) for time-varying parameters and stochastic volatility VAR models for stock price prediction. We extend this methodology to cryptocurrencies and enlarge the model set by allowing for different sources of time variation and model uncertainty.

In total, we have 24 classes of models and combine in the univariate analysis up to 2,621,440 models and in the multivariate case up to 4 time-varying VAR models. We separate univariate analysis from multivariate analysis and in the former one we predict and report results separately for each cryptocurrency; in the latter one we predict and evaluate forecasts jointly for the four cryptocurrencies giving information for building portfolios of cryptocurrencies. We consider prediction from one day ahead to seven days (one week) ahead.

Our results show that DMA and DMS of a large set of models provide forecasting gains in terms of point forecasting relative to the autoregressive benchmark with time–varying volatility. Gains are economically and statistically significant for Bitcoin up to 11% at one day ahead horizon and for Ethereum up to 14% at one day ahead horizon. Evidence is weaker for Litecoin and Ripple. When focusing on density forecasting, gains in predicting Bitcoin and Ethereum disappear, even if moderate improvements emerge for Litecoin and Ripple. Therefore, the combinations of a large set of predictors increase point forecast accuracy for the major currencies, but do not improve density forecasting.

When we focus on multivariate models, we get contrasting evidence as only a few models provide marginally more accurate point predictions and then for longer horizons. Generally, point predictability does not emerge. However, when the complete distribution is predicted, most of the multivariate schemes offer statically significant gains at all horizons. In particular, the selection of time–varying VARs with different sets of predictors and different levels of smoothness provide the largest gains. Our finding corroborates and extends the evidence presented by Hotz-Behofsits et al. (2018), who found sizable improvements in density forecasting but not in point forecasting by the use of their time-varying parameter VAR, and by Catania et al. (2018), who showed that (more) sophisticated volatility models can improve volatility predictions at different forecast horizons, to a large set of multivariate models. Directional predictability indicates that combinations of both univariate and multivariate models can be used to create profitable investment strategies.

The remaining of the paper proceeds as follows. Section 2 provides details on cryptocurrencies and crypto–predictors. Section 3 presents our univariate and multivariate models. Section 4 presents the metrics used to assess our results and explains the major findings in detail. Section 5 provides results for different hyper-parameters of the combination schemes and finally Section 6 concludes.

2. Dataset description

2.1. Cryptocurrencies

The data used in this study are closing log returns of the cryptocurrencies. The crypto–market is open 24 hours a day, seven days a week; hence for computing returns we use the closing price at midnight (UTC). Those data are freely available from CoinMarketcap.[2]

Since the introduction of Bitcoin in 2009, hundreds of other cryptocurrencies have been created and, as of January 2018, 1440 cryptocurrencies exist. The analysis and forecast of such a big dataset are outside the scope of this paper. Here we focus on four major cryptocurrencies: i) Bitcoin, ii) Ethereum, iii) Ripple and iv) Litecoin.

[2]https://coinmarketcap.com/

Bitcoin is the most popular and prominent cryptocurrency based on the decentralization and cryptography. The decentralization means that the Bitcoin network is controlled and owned by all of its users, who must adhere to the same set of rules. The cryptography controls the money creation (fixed to a maximum of 21 million coins) and transactions, and no central bank is needed, see Nakamoto (2009). This decentralised nature offers many advantages, such as being free from government control and regulation, but critics often argue that apart from its users, there is nobody overlooking the whole system and that the value of Bitcoin is unfounded. In spite of this criticism, it touched 20000 USD in December 2017 and at the time of writing it is above 10000 USD, a steep increase since its starting point of a few cents in 2010.

Ethereum is a decentralised platform featuring smart contract functionality that facilitates online contractual agreement applications that elinminate any possibility of downtime, censorship, fraud, or third party interference. The Ethereum also provides a cryptocurrency token called Ether, which can be transferred between accounts and used to compensate participant nodes for computations performed, see Ethereum (2014).

Ripple, developed by the banking industry in 2012, is a blockchain network that incorporates a payment system and a currency system known as XRP. It enables banks to send real–time international payments and for this reason is currently used by many banks such as UBS, Santander and Standard Chartered among many others, see Ripple (2012).

Litecoin was created in 2011 and is based on the same protocol as used by Bitcoin. For this reason it is often considered Bitcoin's leading rival. It has one main feature that distinguishes it from Bitcoin: it is significantly faster regarding transactions and it is particularly attractive in time–critical situations, see Litecoin (2014).

We collect data in the sample between August 8, 2015 to December 28, 2017, giving a total of 874 daily observations and compute percentage daily log returns. Table 1 reports the descriptive statistics for the cryptocurrencies and Table A.1 in Appendix A describes the data transformation. Series are far from being normally distributed as documented in Chu et al. (2015) and they display high volatility, non-zero skewness, very high kurtosis and several spikes, see Figure 1.

2.2. Crypto-predictors

Currently, cryptocurrencies are mainly considered as an alternative investment since their use as payment is still limited. This can create correlations with other assets for at least two main reasons. The first regards investors who usually allocate wealth in a global portfolio and hedge across investments; the second relates to market sentiments that spread fast among different assets. See Bianchi (2018) for similar arguments.

Our list of crypto–predictors includes international stock index prices (the S&P 500, Nikkei 225,

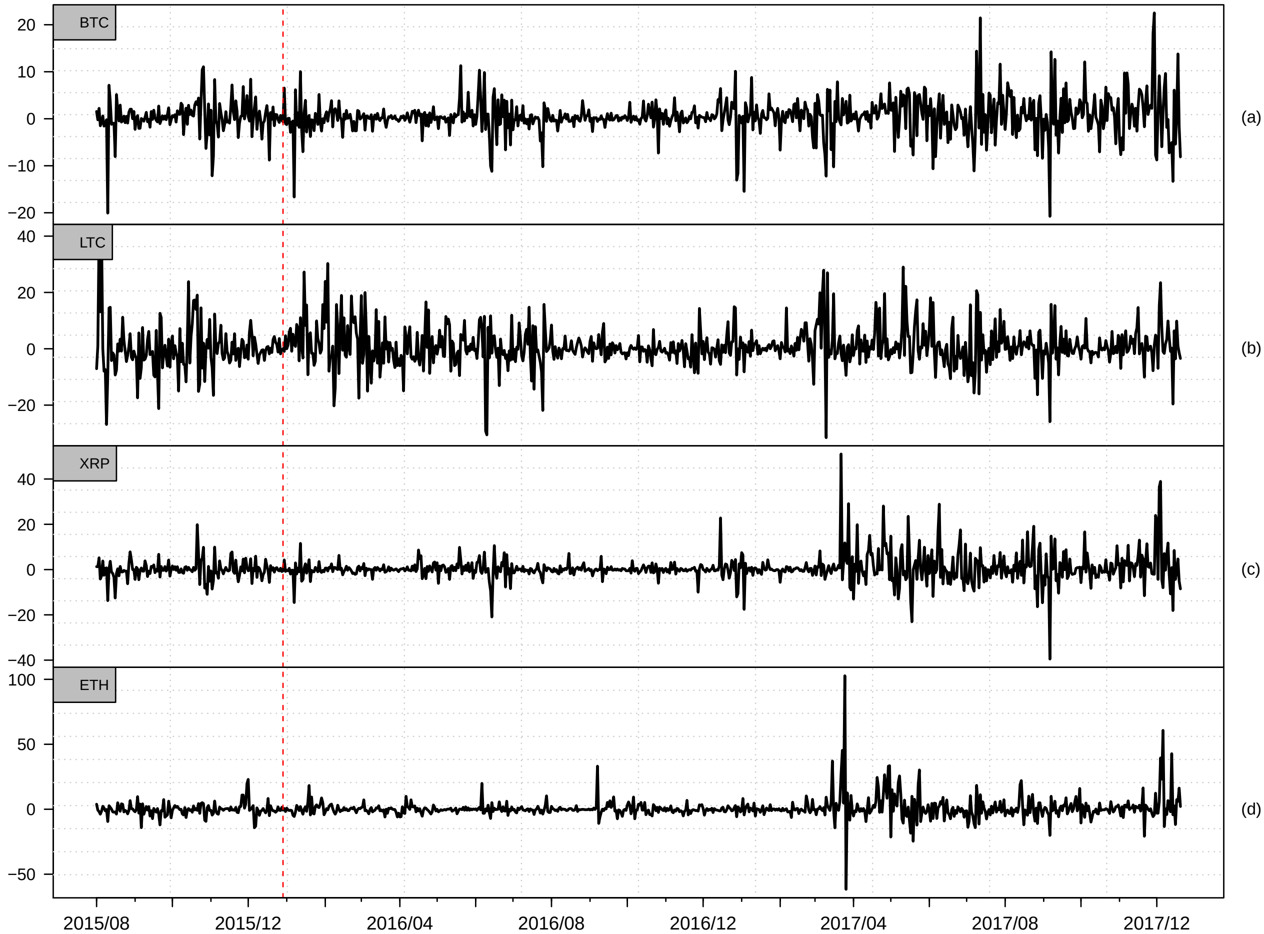


Figure 1: *The plots show the daily percentage log returns of the four cryptocurrencies considered in this study: Bitcoin (BTC) in panel (a), Litecoin (LTC) in panel (b), Ripple (XRP) in panel (c) and Ethereum (ETH) in panel (d). The dashed horizontal red line indicates the beginning of the out-of-sample period on 6th June, 2016. The full sample spans from 9th August, 2015 to 28th December, 2017 for a total of 873 observations.*

Stoxx Europe 600 to be precise); commodity prices (gold and silver prices to be precise); interest rates and CDS (the 5-year Europe credit default swap and the 1-month and 10 year US Treasury rates to be precise); volatility index and VIX closing price. See Table A.1 in Appendix A for the data transformation.

We also apply lags of each of the cryptocurrencies (that would be each of the four currencies discussed in this paper), and a transformation of previous day cryptocurrencies, labelled in the paper as "crypto–explicative", to account for intra-day patterns by taking the difference between the highest and the lowest price, as a proxy of the volatility of the cryptocurrencies.

3. Competing models

In this section, we introduce our different models to make daily forecasts for the cryptocurrencies. As anticipated, we consider univariate and multivariate models with and without exogenous variables and selection and combinations of such models. The full list of univariate and multivariate models

Coin	Bitcoin	Ethereum	Ripple	Litecoin
Created	03-Jan-09	01-Aug-14	01-Jul-13	01-Nov-13
Supply	21 Million Total	18 Million Yearly	100 Billion Total	84 Million Total
Market Cap	277 Billion	466 Billion	27 Billion	15.5 Billion
Maximum	22.512	41.234	51.035	102.736
Minimun	-20.753	-130.211	-39.515	-61.627
Mean	0.453	0.639	0.467	0.591
Median	0.318	-0.051	0.000	-0.338
Std Dev.	3.84	8.535	5.785	7.743
Skewness	-0.091	-3.721	1.637	3.767
Kurtosis	9.391	67.442	19.091	50.455

Table 1: *Descriptive statistics for the four large cryptocurrencies by market capitalization calculated between 08/08/2015 and 28/12/2017, for a total of 874 observations. The table reports the name of the coins, the creation date, the maximum number of coins in million and billion and the market capitalization in December 2017 as reported in https://coinmarketcap.com/. The Ethereum has a total supply of 18 million coins per year; the other three have a prefixed total amount.*

is reported in Table 2 and Table 3, respectively. They are compared with an autoregressive model of order 1 with time–varying volatility driven by an exponential weighted moving average (EWMA) scheme with a forgetting factor $\kappa = 0.96$ (AR(1)–EWMA).[3]

3.1. Univariate models

Linear regression models. Our data set includes 13 different crypto–predictors including macro and finance variables and crypto–explicatives for each series. We apply a linear regression where we include lags of the dependent variable and all predictors. This model is usually referred to as "Kitchen Sink" (KS). We also consider a restricted version where we only include lags of the cryptocurrencies and crypto–explicative, labelled KS-NR.

Model combinations. The previous linear models can suffer from massive model uncertainty. Indeed, a model with 16 predictors and up to 3 lags of the dependent variable results in more than 524,288 possible combinations.[4] To mitigate this fact, we propose to apply model combination techniques.

We rely on the methodology of Raftery et al. (2010) who introduce an estimation technique to predict the output strip thickness for a cold rolling mill, which they refer to as DMA. Recently, DMA has also shown to be useful in macroeconomic and financial applications, see Koop and Korobilis (2011) and Koop and Korobilis (2012), and housing prices, see Bork and Møller (2015).

To provide more detail on the underlying mechanism of DMA, we start by assuming that any combination of the elements on the right–hand–side of a linear regression can be expressed

[3] We also compute forecasts using a simple AR(1) without time–varying volatility as well as with different choices of κ. The benchmark used in the analysis reports best results.

[4] We also include an intercept term in all models, hence we have a total of 19 predictors resulting in $2^{19} = 524,288$ combinations.

as a Dynamic Linear Model (DLM), see West and Harrison (1999) and Raftery et al. (2010). Particularly, let $\mathbf{F}_t^{(i)}$ denote a $p \times 1$ vector based on a given combination of our total predictors, $\mathbf{F}_t = (1, y_{t-1}, y_{t-2}, y_{t-3}, x_{1,t-1}, \ldots, x_{16,t-1})'$. Then, we can express our i–th DLM as:

$$\begin{aligned} y_t &= \mathbf{F}_t^{(i)\prime} \boldsymbol{\beta}_t^{(i)} + \varepsilon_t^{(i)}, \quad \varepsilon_t^{(i)} \sim N\left(0, V_t^{(i)}\right), \\ \boldsymbol{\beta}_t^{(i)} &= \boldsymbol{\beta}_{t-1}^{(i)} + \boldsymbol{\eta}_t^{(i)}, \quad \boldsymbol{\eta}_t^{(i)} \sim N\left(0, \mathbf{Q}_t^{(i)}\right), \end{aligned} \tag{1}$$

where the $p \times 1$ vector of time–varying regression coefficients, $\boldsymbol{\beta}_t^{(i)} = \left(\beta_{1t}^{(i)}, \ldots, \beta_{pt}^{(i)}\right)'$, determines the impact of $\mathbf{F}_t^{(i)}$ on y_t. The Random Walk specification of $\boldsymbol{\beta}_t^{(i)}$ does not assume any systematic movements but considers changes in $\boldsymbol{\beta}_t^{(i)}$ as unpredictable.

The conditional variances, $V_t^{(i)}$ and $\mathbf{Q}_t^{(i)}$, are unknown quantities. We assume an Inverse–Gamma prior for $V_0^{(i)}$ and update its estimate at each point in time via its posterior distribution as in Prado and West (2010). We employ the strategy as in Prado and West (2010) and introduce an additional forgetting factor, $\kappa \in (0, 1)$, to induce time variation in V_t. Usual choices are $\kappa = 0.96$ and $\kappa = 0.99$, see Prado and West (2010). Regarding $\mathbf{Q}_t^{(i)}$ we employ the forgetting factor approach detailed in Dangl and Halling (2012). This additional forgetting factor, labelled as λ, is allowed to take one of the $d = 5$ values in the grid $\{0.91, 0.93, 0.95, 0.97, 0.99\}$ and to augment the number of possible models up to $k = 2^{19}d = 2'621'440$.[5] Notably, when $\mathbf{Q}_t^{(i)} = \mathbf{0}$ for $t = 1, \ldots, T$, then $\boldsymbol{\beta}_t^{(i)}$ is constant over time. Thus, (1) nests the specification of the constant regression coefficients. For $\mathbf{Q}_t^{(i)} \neq \mathbf{0}$, $\boldsymbol{\beta}_t^{(i)}$ varies according to equation 1. However, this does not mean that $\boldsymbol{\beta}_t^{(i)}$ needs to change at every time period; we can easily have periods where $\mathbf{Q}_t^{(i)} = \mathbf{0}$ and thus $\boldsymbol{\beta}_t^{(i)} = \boldsymbol{\beta}_{t-1}^{(i)}$. Ultimately, the nature of the time variation in the regression coefficients is dependent on the data at hand.

DMA then averages forecasts across the k different combinations using a recursive updating scheme based on the predictive likelihood that measures the ability of a model to predict y_t, thus making it the central quantity of interest for model evaluation. This updating scheme relies on the specification of an additional forgetting factor, $\alpha \in (0, 1)$, which determines how fast the performance of past models determines the weights of the models. Common choices are $\alpha = 0.99$ and $\alpha = 1.0$, see equations (15) and (16) in Koop and Korobilis (2012). Besides averaging, we can also use the model receiving the highest probability among all model combinations to forecasts, this is the so-called DMS, see Koop and Korobilis (2012). For both DMA and DMS, we also include restricted versions with only lags of the dependent variable, see Table 2 for more details. Estimation and prediction are made exploiting the **eDMA** package for R of Catania and Nonejad (2018). Finally, we apply Bayesian Model Averaging (BMA) and combine all possible regression models based on predictive likelihood

[5] Dangl and Halling (2012) refer to this parameter as δ.

at each point.

Abbreviation	*Full Description*
AR(1)–EWMA	Autoregressive model of order one with EWMA variance and κ=0.96, benchmark model.
KS	Kitchen Sink specification, *i.e.*, a linear multiple regression including all variables.
KS–NR	Kitchen Sink specification with only the lagged values of the series as covariates.
DMA	Dynamic Model Averaging across all models and forgetting factor combinations. See Dangl and Halling (2012).
DMS	Dynamic Model Selection, selecting at each time t the best model between all models and forgetting factor combinations. See Dangl and Halling (2012).
DMA–NR	Dynamic Model Averaging with only the lagged values of the series as covariates.
DMS–NR	Dynamic Model Selection with only the lagged values of the series as covariates.
BMA	Bayesian Model Averaging: forecast combinations of all possible regression models.

Table 2: *Univariate models considered in the forecasting exercise. The first column is the abbreviation of the model. The second column provides a brief description of each individual model.*

3.2. Multivariate models

Constant parameter VARs. The first class of multivariate models we consider is the constant parameter Vector Autogressive (VAR) specification. VARs are among the most common models applied in financial and macroeconomic forecasting, see among others Lutkepohl (2007) and Koop and Korobilis (2010). Among the standard OLS estimation we also consider the Bayesian VAR (BVAR) specification as in Koop (2003). For both VAR and BVAR we select three lags based on the BIC.

Time–varying parameter specifications. Cryptocurrencies are subject to several instabilities, both in mean and at higher moments. Large parametrised constant parameter VAR models might fail to capture these instabilities and we extend the model set with 14 different time–varying specifications. The starting point for the analysis is the time-varying parameter vector autoregression model (TVP-VAR) as described in Koop and Korobilis (2013)(henceforth KK), where a new approach to estimate large–dimensional TVP-VAR is provided. Focusing on the case where the VAR has one lag and intercepts are suppressed, the TVP-VAR(1) model can be written as:

$$\mathbf{y}_t = \mathbf{T}_t \mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t, \tag{2}$$

with $\mathbf{y}_t$ being an $(M \times 1)$ vector containing observations on M time series at time t, T_t an $(M \times M)$ matrix containing M^2 parameters and $\boldsymbol{\varepsilon}_t \sim \mathrm{N}(0, \Sigma_t)$. KK close the model by specifying dynamics for the time-varying VAR parameters and rewrite the model in state space form:

$$\begin{aligned} \mathbf{y}_t &= \mathbf{F}_t \boldsymbol{\beta}_t + \boldsymbol{\varepsilon}_t, \qquad \boldsymbol{\varepsilon}_t \sim \mathrm{N}(0, \boldsymbol{\Sigma}_t), \\ \boldsymbol{\beta}_{t+1} &= \boldsymbol{\beta}_t + \boldsymbol{\eta}_t, \qquad \boldsymbol{\eta}_t \sim \mathrm{N}(\mathbf{0}, \mathbf{Q}_t), \end{aligned} \tag{3}$$

where the $(M \times M^2)$ matrix $\mathbf{F}_t$ collects lagged observations and the time-varying vector $\boldsymbol{\beta}_t$ captures the time-varying VAR parameters (with $\boldsymbol{\beta}_t = vec(\mathbf{T}_t') = vecr(\mathbf{T}_t)$), which are taking random walks with innovations $\boldsymbol{\eta}_t \sim \mathrm{N}(\mathbf{0}, \mathbf{Q}_t)$. In line with the constant parameter VARs, we consider three lags of the dependent variables.

As the model is equation 2 can be put in state space form as shown in equation 3, the usual estimation approaches based on the Kalman filter seem attractive. For example, one could use the frequentist approach of maximizing the likelihood, or the Bayesian approach of using Markov Chain Monte Carlo (MCMC) methods. Unfortunately, for large-dimensional VARs these standard approaches turn out to be unfeasible. For a VAR of dimension 7 and a model with 4 lags, the amount of time-varying parameters would be $4 \times 7^2 = 196$, making either approach computationally very demanding.

To reduce the computational burden, KK propose to make two adjustments such that the usual Kalman filter can still be used. The parameters of the model are in the variance matrices $\mathbf{Q}_t$ and $\boldsymbol{\Sigma}_t$, the idea of KK is to take $\mathbf{Q}_t$ out of the model and replace it with an approximation. In this case $\boldsymbol{\beta}_t$ can be obtained using closed-form expressions without having to maximize a likelihood first in order to get parameter estimates (or to do so using MCMC methods). Typically the latent state innovation variance $\mathbf{Q}_t$ enters the Kalman filter in the updating step, where the state variance matrix is updated through $\mathbf{P}_{t|t} = \mathbf{P}_{t-1|t-1} + \mathbf{Q}_t$ (see Durbin and Koopman, 2012). If one instead writes $\mathbf{P}_{t|t} = \frac{1}{\lambda}\mathbf{P}_{t-1|t-1}$ (for some λ), an estimate of $\mathbf{Q}_t$ is no longer necessary. This is often referred to as a forgetting factor set-up. The second adjustment is to replace the measurement error variance matrix $\boldsymbol{\Sigma}_t$ by an EWMA filter. The EWMA filtering recursion $\hat{\boldsymbol{\Sigma}}_t = \kappa\hat{\boldsymbol{\Sigma}}_{t-1} + (1-\kappa)\bar{\boldsymbol{\varepsilon}}_t\bar{\boldsymbol{\varepsilon}}_t'$ (with $\bar{\boldsymbol{\varepsilon}}_t = \mathbf{y}_t - \mathbf{F}_t\boldsymbol{\beta}_{t|t-1}$) gives the measurement error variance estimates, which we can plug into the filter. As discussed above this methodology requires the specification of the hyperparameters λ and κ (and the specification of the initial condition of the states $\boldsymbol{\beta}_0$ and $\boldsymbol{\Sigma}_0$), we refer to KK for an extensive discussion of the problem.

As explained in KK we carry out model selection using a model space involving all the variables reported in Table A.1. We cluster them in four different databases and consequently have four TVP-VARs of different size: a small TVP-VAR with only the four cryptocurrencies; a medium TVP-VAR with the four cryptocurrencies series plus the four crypto–explicatives for a total of eight dependent variables; a second medium TVP-VAR with the four cryptocurrencies series plus the seven financial and macro variables for a total of 11 dependent variables; and a large TVP-VAR with all the 17 variables. As described in KK the algorithm selects between the four TVP-VARs based on past predictive likelihoods for the set of variables the researcher is interested in forecasting, allowing for model switching accordingly to an ad hoc hyperparameter α. Moreover, as described in KK, the forgetting factor λ can be dynamically selected together with the optimal value of the shrinkage parameter at different points in time. This results in 14 different models, labelled $\mathcal{M}_4$–$\mathcal{M}_{18}$ in Table 3.

Abbreviation	*Full Description*
$\mathcal{M}_1$	AR(1)–EWMA with $\kappa = 0.96$, benchmark model.
$\mathcal{M}_2$	Vector Autoregressive with 3 lags estimated using the OLS.
$\mathcal{M}_3$	Baysian Vector Autoregressive with 3 lags as described in Koop and Korobilis (2010).
$\mathcal{M}_4$	TVP-VAR(3) with 4 cryptocurrencies series and stochastic volatility. Optimal value of the shrinkage parameter is selected using DMS as outlined in Koop and Korobilis (2013). In this model λ is dynamically selected and $\kappa = 0.96$ and $\alpha = 0.99$.
$\mathcal{M}_5$	TVP-VAR(3) with 4 cryptocurrencies series and stochastic volatility. Optimal value of the shrinkage parameter is selected using DMS as outlined in Koop and Korobilis (2013). In this model $\lambda = 0.99$, $\kappa = 0.96$ and $\alpha = 0.99$.
$\mathcal{M}_6$	TVP-VAR(3) with 4 cryptocurrencies series and stochastic volatility. Optimal value of the shrinkage parameter is selected using DMS as outlined in Koop and Korobilis (2013). In this model λ is dynamically selected and $\kappa = 0.96$ and $\alpha = 0.001$.
$\mathcal{M}_7$	TVP-VAR(3) with two models: the first uses the 4 cryptocurrencies series, the second uses cryptocurrencies plus crypto–explicative. Stochastic volatility is also present. Models and optimal value of the shrinkage parameter are selected using DMS as outlined in Koop and Korobilis (2013). In this model λ is dynamically selected and $\kappa = 0.96$ and $\alpha = 0.99$.
$\mathcal{M}_8$	TVP-VAR(3) with two models: the first uses the 4 cryptocurrencies series, the second uses cryptocurrencies plus crypto–explicative. Stochastic volatility is also present. Models and optimal value of the shrinkage parameter are selected using DMA as outlined in Koop and Korobilis (2013). In this model λ is dynamically selected and $\kappa = 0.96$ and $\alpha = 0.99$.
$\mathcal{M}_9$	TVP-VAR(3) with two models: the first uses the 4 cryptocurrencies series, the second uses cryptocurrencies plus crypto–explicative. Stochastic volatility is also present. Models and optimal value of the shrinkage parameter are selected using DMS as outlined in Koop and Korobilis (2013). In this model $\lambda = 0.99$, $\kappa = 0.96$ and $\alpha = 0.99$.
$\mathcal{M}_{10}$	TVP-VAR(3) with two models: the first uses the 4 cryptocurrencies series, the second uses cryptocurrencies plus crypto–explicative. Stochastic volatility is also present. Models and optimal value of the shrinkage parameter are selected using DMA as outlined in Koop and Korobilis (2013). In this model $\lambda = 0.99$, $\kappa = 0.96$ and $\alpha = 0.99$.
$\mathcal{M}_{11}$	TVP-VAR(3) with two models: the first uses the 4 cryptocurrencies series, the second uses cryptocurrencies plus macroeconomic variables. Stochastic volatility is also present. Models and optimal value of the shrinkage parameter are selected using DMS as outlined in Koop and Korobilis (2013). In this model λ is dynamically selected and $\kappa = 0.96$ and $\alpha = 0.99$.
$\mathcal{M}_{12}$	TVP-VAR(3) with two models: the first uses the 4 cryptocurrencies series, the second uses cryptocurrencies plus macroeconomic variables. Stochastic volatility is also present. Models and optimal value of the shrinkage parameter are selected using DMA as outlined in Koop and Korobilis (2013). In this model λ is dynamically selected and $\kappa = 0.96$ and $\alpha = 0.99$.
$\mathcal{M}_{13}$	Time varying parameters VAR(3) with two models: the first uses the 4 cryptocurrencies series, the second uses cryptocurrencies plus macroeconomic variables. Stochastic volatility is also present. Models and optimal value of the shrinkage parameter are selected using DMS as outlined in Koop and Korobilis (2013). In this model $\lambda = 0.99$, $\kappa = 0.96$ and $\alpha = 0.99$.
$\mathcal{M}_{14}$	TVP-VAR(3) with two models: the first uses the 4 cryptocurrencies series, the second uses cryptocurrencies plus macroeconomic variables. Stochastic volatility is also present. Models and optimal value of the shrinkage parameter are selected using DMA as outlined in Koop and Korobilis (2013). In this model $\lambda = 0.99$, $\kappa = 0.96$ and $\alpha = 0.99$.
$\mathcal{M}_{15}$	TVP-VAR(3) with four models: the first uses the 4 cryptoseries, the second uses cryptocurrencies plus crypto–explicative, the third uses cryptocurrencies plus macroeconomic variables and the fourth uses all the series. Stochastic volatility is also present. Models and optimal value of the shrinkage parameter are selected using DMS as outlined in Koop and Korobilis (2013). Here λ is dynamically selected, $\kappa = 0.96$ and $\alpha = 0.99$.
$\mathcal{M}_{16}$	TVP-VAR(3) with four models:the first uses the 4 cryptocurrencies series, the second uses cryptocurrencies plus crypto–explicative, the third uses cryptocurrencies plus macroeconomic variables and the fourth uses all the series. Stochastic volatility is also present. Models and optimal value of the shrinkage parameter are selected using DMA as outlined in Koop and Korobilis (2013). Here λ is dynamically selected and $\kappa = 0.96$ and $\alpha = 0.99$.
$\mathcal{M}_{17}$	TVP-VAR(3) with four models: the first uses the 4 cryptocurrencies series, the second uses cryptocurrencies plus crypto–explicative, the third uses cryptocurrencies plus macroeconomic variables and the fourth uses all the series. Stochastic volatility is also present. . Models and optimal value of the shrinkage parameter are selected using DMS as outlined in Koop and Korobilis (2013). Here λ is dynamically selected, $\kappa = 0.96$ and $\alpha = 0.99$.
$\mathcal{M}_{18}$	TVP-VAR(3) with four models: the first uses the 4 cryptocurrencies series, the second uses cryptocurrencies plus crypto–explicative, the third uses cryptocurrencies plus macroeconomic variables and the fourth uses all the series. Stochastic volatility is also present. . Models and optimal value of the shrinkage parameter are selected using DMA as outlined in Koop and Korobilis (2013). Here $\lambda = 0.99$, $\kappa = 0.96$ and $\alpha = 0.99$.

Table 3: *The table reports all the multivariate models considered in the forecasting exercise plus the benchmark model. The first column is the abbreviation of the model. The second column provides a brief description of each model.*

4. Empirical application

Our results are based on the $h = 1, \ldots, 7$ days-ahead forecasting process using an expanding window and an initial in-sample period of 146 days for Bitcoin, Litecoin, Ripple and Ethereum. Multistep-ahead predictions are obtained through direct forecasting, see Marcellino et al. (2006). Hence, our forecast evaluation period is from January 1, 2016 to December 28, 2017. In all the analyses we use the AR(1)–EWMA as a benchmark for the univariate and multivariate models described in the previous section. We discuss forecast metrics in subsection 4.1, univariate results in subsection 4.2 and multivariate results in subsection 4.3. For the univariate analysis, we present results for each cryptocurrency separately; for the multivariate analysis, we provide joint results.[6]

4.1. Forecast metrics

We assess the performance of our forecasts using different point and density metrics. Considering the accuracy of point forecasts, we use the mean squared errors (MSEs) for each of the forecast horizons, $h = 1, \cdots, 7$ we consider.[7] For the univariate exercise, the metric is computed separately for each cryptocurrency series, $i =$ Bitcoin, Litecoin, Ripple and Etherum:

$$\text{MSE}_{i,h} = \frac{1}{T-R} \sum_{t=R}^{T-h} \left(\hat{y}_{i,t+h|t} - y_{i,t+h}\right)^2, \tag{4}$$

where T is the number of observations, R is the length of the rolling window, $\hat{y}_{i,t+h|t}$ is the ith-cryptocurrency forecasts made at time t for horizon h and $y_{i,t+h}$ is the realization. For the multivariate application, we compute a MSE for each forecast and an average MSE as:

$$\text{MSE}_{h} = \frac{1}{(T-R)N} \sum_{t=R}^{T-h} \sum_{i=1}^{N} \left(\hat{y}_{i,t+h|t} - y_{i,t+h}\right)^2, \tag{5}$$

where $N = 4$ is the number of cryptocurrencies. To evaluate the density forecasts, we use predictive log score (LS). The LS is commonly viewed as the broadest measure of density accuracy, see Geweke and Amisano (2010). As for the MSE, we compute it for each horizon and series separately in the univariate application

$$s_{i,h}(y_i) = \sum_{t=R}^{T-h} \ln\left(f(y_{i,t+h}|I_{i,t})\right), \tag{6}$$

where $f(y_{i,t+h}|I_{i,t})$ is the predictive density for $y_{i,t+h}$ constructed using information up to time t. The multivariate version is

$$s_{h}(y) = \sum_{t=R}^{T-h} \ln\left(f(\mathbf{y}_{t+h}|I_{t})\right), \tag{7}$$

[6] Separate statistics for each currency are reported in Appendix B.
[7] In Appendix B we also report mean absolute deviations (MADs).

where $f(\mathbf{y}_{t+1}|I_t)$ is the joint predictive density for the 4-variate $\mathbf{y}_{t+h}$ constructed using information up to time t. For the AR(1)-EWMA, we assume a joint distribution composed by the four independent marginal predictions; therefore, we assume a diagonal variance-covariance matrix.

More specifically, we report the MSEs and the LSs for the benchmark model. For the other models, we report the ratios of each model's MSE with respect to the baseline. Entries smaller than 1 indicate that the given model yields forecasts that are more accurate than those from the baseline, and differences in score relative to the baseline, such that a positive number, indicates a model that beats the baseline. In order to statistically assess the differences among alternative models, we apply Diebold and Mariano (1995) t-tests for equality of the average loss (with loss defined as squared error and negative log score) of each model versus the AR(1)-EWMA benchmark,[8] and we also employ the Model Confidence Set procedure of Hansen et al. (2011) using the R package MCS detailed in Bernardi and Catania (2016) to jointly compare all predictions. Differences are tested separately for each horizon h.

Finally, we provide an economic evaluation of our forecasts by studying directional predictability of cryptocurrencies returns. As pointed out by, for example, Christoffersen and Diebold (2006), sign predictability may exist even in the absence of mean predictability and it complementes density forecasting by providing useful indications in terms of creating (simple) profitable investment strategies. We compute the success rate as the percentage of times a model correctly predicts the sign of future returns.[9] A success rate of one means that a model predicts the correct sign for all forecasts, a success rate of zero means the model never predicts the correct sign. Similarly to the MSE evaluation, for multivariate models we report an average success rate across the four cryptocurrencies for each forecast horizon.

4.2. Univariate forecasting results

Table 4 reports MSEs for predicting the four cryptocurrencies using univariate models.[10] The largest gains are found when using DMA for predicting Bitcoin and Litecoin at daily horizon and Ethereum at several horizons. For Bitcoin, DMA gives statistically significant reductions at one day-ahead of 11%. Even when excluding crypto–predictors and applying combinations or selections of lags, the strategies labelled DMA–NR and DMS–NR, the forecast accuracy statistically improves. Considering the large volatility of the series and that we focus on a daily forecast horizon, the gains are economically sizeable,

[8] We also compute the Amisano and Giacomini (2007) test for differences in log score performance. Notice that this test is only a rough gauge since the asymptotic validity of the Amisano and Giacomini test requires the models to be estimated with a rolling, rather than expanding, sample of data as in our case. Evidence is qualitatively similar.

[9] We do not perform sign predictability tests because as indicated in Christoffersen and Diebold (2006) similar "tests that rely only on the sign sequence omit important information about volatility dynamics, which is potentially valuable for detecting sign predictability." On this background, we leave for future research market timing tests and portfolio strategies.

[10] See Table B.2 in Appendix B for MAD scores.

in particular when compared to other highly volatile assets such as stock prices and exchange rates. When the forecast horizon increases, the forecast accuracy decreases and even if DMA provides MSE lower than the AR(1)–EWMA benchmark in four cases out of six, the difference is not statistically significant.

For the other models, straightforward linear regressions based on direct forecasting, labelled KS and KS-NR in the Table, do not seem a credible strategy. KS statistically outperforms the benchmark at horizon four, but it produces enormous losses at other horizons, see for example statistics at the second horizon. BMA also outperforms the benchmark at horizon four, but it is inferior at all other horizons. Therefore, methodology other than DMA does not seem robust across horizons.

When focusing on Ethereum, we find evidence of statistically superior predictability by the alternative models to the benchmark at several horizons. Again, DMA gives the best statistic at one day ahead with a 14% reduction in MSE; and other versions of DMA and DMS also provide economically sizeable gains at those horizons. Interestingly, a combination or selection of its own lags and lags of other currencies is a more valuable strategy for Ethereum than Bitcoin, somewhat confirming its central role as major exchanger of currencies in the crypto–system. Ethereum is highly connected to several currencies and actually more so than Bitcoin, which traditionally drives movements in the crypto–market, and DMA-NR and DMS-NR provide statistically significant gains at horizons $h = 3, 4, 5, 6, 7$.

With Litecoin, DMA gives lower MSE at one and four days ahead, but not statistically significant. KS and BMA improve the predictability at some horizons, but they also perform very poorly at other horizons, see for example at three days ahead.

For Ripple, the predictability is weaker and model averaging seems to reduce the accuracy we have found with Bitcoin, Litecoin and Ethereum, even if Table 7 indicates still large uncertainty among predictors. Only DMA-NR outperforms the benchmark at two, five, six and seven days ahead. The lower capitalization of Litecoin and Ripple relative to Bitcoin and Ethereum results in lower correlations with other assets and less accurate predictions. New predictors based on crypto–market sentiments might be considered to investigate heterogeneity across cryptocurrencies.

Table 5 provides the log score results for the univariate models. Evidence is different than for point forecasting. Gains vanish for predictions of Bitcoin and Ethereum with no models providing a higher score than the benchmark, and in the case of Ethereum the AR model is the only specification included in the model confidence set at all horizons. Weaker evidence is found for Litecoin only at longer horizons with DMA-NR and DMS-NR. However, evidence reverts for Ripple with several models outperforming the AR model at all horizons. In particular, DMA, DMS, DMA-NR and DMS-NR improve the accuracy at all horizons for both series. In general, univariate DMA, even if it allows for time–varying volatility, seems to fail to capture higher moments dynamics of cryptocurrencies.

h	1	2	3	4	5	6	7
				Bitcoin			
AR(1)–EWMA	10.69	10.63	10.62	10.61	10.81	10.81	10.98
KS	2.14	117.56	2.51	**0.91**	1.04	1.01	1.11
KS–NR	3.26	44.24	2.27	1.02	1.01	1.05	0.98
DMA–NR	**0.96**	**0.96**	1.01	1.00	0.98	0.97	**0.94**
DMS–NR	**0.98**	1.00	1.02	1.04	1.02	1.00	0.96
DMA	**0.89**	0.99	1.03	0.99	**0.96**	0.96	1.03
BMA	1.13	1.00	1.21	**0.89**	**0.98**	1.01	0.97
DMS	0.97	1.01	1.05	1.01	0.99	0.98	1.11
				Litecoin			
AR(1)–EWMA	21.88	21.76	21.76	21.76	21.80	21.83	21.95
KS	0.96	1.24	44.77	**0.98**	1.23	1.10	1.10
KS–NR	**0.98**	1.06	3.45	0.99	1.01	1.03	**0.97**
DMA–NR	1.03	1.03	1.12	1.01	0.99	1.07	1.04
DMS–NR	1.04	1.07	1.17	1.01	1.02	1.04	0.99
DMA	0.97	1.02	1.09	1.00	1.11	1.09	1.12
BMA	**0.98**	1.00	1.28	**0.98**	0.99	**0.91**	0.99
DMS	0.98	1.13	1.21	1.02	1.12	1.04	1.04
				Ripple			
AR(1)–EWMA	25.79	26.49	26.79	26.75	26.70	27.63	27.63
KS	1.27	1.45	1.46	1.22	2.58	1.20	3.07
KS–NR	1.10	1.06	1.04	0.99	1.14	**0.91**	0.98
DMA–NR	1.00	**0.94**	0.98	1.04	**0.96**	**0.92**	**0.89**
DMS–NR	1.04	**0.93**	1.01	0.98	1.03	0.99	0.98
DMA	1.04	1.03	1.14	1.17	1.03	1.00	1.10
BMA	1.09	1.00	1.06	1.06	0.99	**0.96**	**0.97**
DMS	1.14	1.08	1.25	1.28	1.16	1.16	1.20
				Ethereum			
AR(1)–EWMA	25.92	25.09	25.13	25.19	25.20	25.23	25.22
KS	1.06	153.49	1.14	0.98	1.00	2.47	1.07
KS–NR	0.99	6.90	1.02	**0.97**	**0.99**	1.00	**0.97**
DMA–NR	**0.91**	1.00	**0.91**	**0.92**	**0.94**	**0.94**	**0.94**
DMS–NR	**0.91**	1.01	**0.93**	**0.94**	**0.95**	**0.95**	**0.96**
DMA	**0.87**	1.03	1.00	0.97	1.04	1.01	1.01
BMA	**0.96**	1.35	**0.98**	**0.96**	**0.98**	0.98	**0.96**
DMS	**0.88**	1.06	1.04	1.00	1.07	1.08	0.99

Table 4: *Mean squared error (MSE), computed over the forecast horizon. Results are reported relative to the benchmark specification (AR(1)–EWMA) for which the absolute score is reported. The description of the model is reported in Table 2. Values in* ***bold****, indicate rejection of the null hypothesis of Equal Predictive Ability between each model and the benchmark according to the Diebold–Mariano test at the* 5% *confidence level. Grey cells indicate those models that belong to the Superior Set of Models delivered by the Model Confidence Set procedure at confidence level* 10%*.*

Catania et al. (2018) find that sophisticated univariate models are required to produce accurate forecasts of cryptocurrency volatility.

h	1	2	3	4	5	6	7
				Bitcoin			
AR(1)–EWMA	−799.65	−802.52	−803.13	−811.52	−818.94	−822.67	−828.14
KS	−456.19	−1805.61	−455.45	−81.66	−69.36	−102.79	−67.23
KS–NR	−651.63	−1437.81	−492.32	−72.77	−46.67	−59.03	−27.04
DMA–NR	−5.62	−6.88	−11.69	6.07	17.13	22.22	33.64
DMS–NR	−2.30	−8.40	−11.10	−11.34	12.40	17.91	36.26
DMA	−17.93	−14.53	−21.81	−4.64	4.54	6.16	24.42
BMA	−144.65	−96.42	−112.33	−59.03	−46.43	−47.05	−25.12
DMS	−40.90	−39.35	−34.03	−30.92	−6.38	0.61	25.36
				Litecoin			
AR(1)–EWMA	−730.06	−738.61	−744.18	−748.66	−752.92	−759.41	−758.45
KS	−65.48	−108.06	−1413.61	−19.55	−90.31	−105.06	−10.06
KS–NR	−31.31	−55.68	−417.78	−4.93	−14.94	−16.20	20.54
DMA–NR	64.64	70.74	43.12	**75.34**	**87.59**	**89.53**	**96.51**
DMS–NR	67.44	65.12	27.32	**79.65**	**89.49**	**89.75**	**95.48**
DMA	51.80	50.21	32.66	56.25	62.51	73.39	75.40
BMA	−29.67	−29.21	−109.76	−11.22	−3.85	−9.27	9.85
DMS	51.75	48.33	15.74	50.62	49.76	67.46	76.31
				Ripple			
AR(1)–EWMA	−655.37	−668.20	−672.74	−683.92	−687.56	−687.58	−691.48
KS	−5.19	−94.57	−74.14	27.90	−335.13	−96.59	−545.20
KS–NR	33.87	22.12	53.53	46.03	12.25	**80.50**	**89.28**
DMA–NR	**97.81**	**107.13**	**109.27**	**122.05**	**125.44**	**128.66**	**132.97**
DMS–NR	**97.20**	**107.93**	**110.39**	**123.13**	**129.60**	**132.45**	**132.93**
DMA	**79.84**	**81.71**	**80.23**	**99.66**	**116.45**	**92.73**	**122.92**
BMA	35.13	50.78	22.82	55.70	65.21	**83.63**	**88.45**
DMS	**67.55**	**94.90**	63.11	**97.54**	**123.65**	**87.38**	**117.15**
				Ethereum			
AR(1)–EWMA	−543.42	−550.18	−561.32	−563.28	−564.31	−563.93	−562.63
KS	−291.33	−2016.31	−287.32	−253.52	−253.66	−616.23	−260.69
KS–NR	−269.70	−928.56	−242.87	−229.06	−230.16	−231.28	−225.11
DMA–NR	−196.69	−200.03	−171.89	−166.81	−166.06	−165.13	−161.50
DMS–NR	−188.05	−215.10	−177.90	−171.51	−169.59	−170.47	−172.89
DMA	−200.78	−200.66	−185.76	−182.36	−175.45	−174.30	−177.80
BMA	−258.54	−316.78	−233.68	−226.49	−224.91	−237.61	−220.75
DMS	−204.48	−224.47	−201.92	−182.01	−192.63	−193.78	−204.26

Table 5: *Log Score (LS), computed over the forecast horizon. Results are reported relative to the benchmark specification (AR(1)–EWMA) for which the absolute score is reported. The description of the model is reported in Table 2. Values in* ***bold****, indicate rejection of the null hypothesis of Equal Predictive Ability between each model and the benchmark according to the Diebold–Mariano test at the 5% confidence level. Grey cells indicate those models that belong to the Superior Set of Models delivered by the Model Confidence Set procedure at confidence level 10%.*

h	1	2	3	4	5	6	7
				Bitcoin			
AR(1)–EWMA	0.53	0.53	0.53	0.53	0.53	0.53	0.52
KS	0.52	0.52	0.55	0.63	0.51	0.53	0.45
KS–NR	0.52	0.54	0.53	0.53	0.53	0.53	0.52
DMA–NR	0.62	0.63	0.61	0.61	0.62	0.61	0.61
DMS–NR	0.59	0.60	0.60	0.59	0.60	0.61	0.56
DMA	0.63	0.60	0.59	0.60	0.60	0.57	0.54
BMA	0.53	0.50	0.53	0.71	0.54	0.51	0.52
DMS	0.60	0.57	0.58	0.61	0.58	0.55	0.47
				Litecoin			
AR(1)–EWMA	0.56	0.57	0.56	0.57	0.56	0.56	0.55
KS	0.55	0.50	0.52	0.53	0.51	0.54	0.47
KS–NR	0.59	0.59	0.55	0.55	0.58	0.58	0.58
DMA–NR	0.59	0.60	0.58	0.60	0.61	0.61	0.60
DMS–NR	0.59	0.59	0.60	0.58	0.59	0.59	0.58
DMA	0.58	0.59	0.57	0.57	0.58	0.57	0.53
BMA	0.59	0.59	0.59	0.59	0.58	0.51	0.57
DMS	0.57	0.58	0.54	0.56	0.55	0.58	0.55
				Ripple			
AR(1)–EWMA	0.63	0.65	0.64	0.64	0.65	0.65	0.65
KS	0.59	0.50	0.53	0.62	0.62	0.63	0.51
KS–NR	0.62	0.63	0.63	0.63	0.61	0.62	0.64
DMA–NR	0.63	0.65	0.63	0.64	0.63	0.62	0.62
DMS–NR	0.63	0.64	0.63	0.64	0.64	0.63	0.64
DMA	0.64	0.61	0.65	0.63	0.64	0.60	0.59
BMA	0.63	0.65	0.69	0.65	0.65	0.65	0.65
DMS	0.61	0.60	0.63	0.62	0.62	0.62	0.60
				Ethereum			
AR(1)–EWMA	0.56	0.56	0.56	0.57	0.57	0.57	0.57
KS	0.55	0.46	0.53	0.55	0.58	0.55	0.45
KS–NR	0.50	0.43	0.52	0.59	0.59	0.53	0.53
DMA–NR	0.61	0.60	0.59	0.60	0.60	0.61	0.59
DMS–NR	0.56	0.57	0.56	0.57	0.57	0.58	0.59
DMA	0.61	0.58	0.51	0.58	0.56	0.57	0.55
BMA	0.51	0.43	0.52	0.52	0.55	0.52	0.55
DMS	0.58	0.55	0.53	0.57	0.55	0.55	0.54

Table 6: *Success Rate (SR), computed over the forecast horizon. Models' description is reported in Table 2.*

Table 6 reports the results for directional predictability of our models. DMA does well for all cryptocurrencies, with success rates higher than 60% for Bitcoin, Ethereum and Ripple at one day horizon and statistics above 50% in all cases. DMA-NR also provides positive results and in several cases its success rate is the highest and often close to 60%. The benchmark model is never above 60% for Bitcoin, Ethereum and Litecoin, and the poorest performance is given by the KS model with success rates below 50% in 5 cases out of 28.

h = 1	h = 2	h = 3	h = 4	h = 5	h = 6	h = 7
			Bitcoin			
VIX(37)	Lag3(13)	SV(45)	ETH(38)	ETH_HL(46)	BTC_HL(43)	CDS_5y(41)
ETH(33)	ETH_HL(12)	Lag3(42)	CDS_5y(37)	BD_1m(41)	XRP(41)	SV(39)
XRP_HL(28)	SP_500(10)	Lag4(37)	LTC(36)	Lag5(39)	NK_225(40)	ES_600(38)
SV(24)	BD_1m(9)	BD_1m(37)	ETH_HL(36)	NK_225(38)	GLD(39)	LTC_HL(38)
NK_225(24)	SV(8)	SP_500(35)	BTC_HL(35)	LTC(38)	VIX(38)	BD_1m(38)
			Litecoin			
XRP_HL(36)	SV(43)	Lag5(57)	ETH(38)	BTC(45)	SP_500(50)	SV(45)
SV(35)	Lag3(38)	XRP(33)	ETH_HL(38)	NK_225(36)	ETH(48)	SP_500(37)
ES_600(33)	Lag4(35)	SV(31)	XRP_HL(35)	BD_10y(35)	XRP(41)	GLD(36)
Lag2(33)	Lag2(34)	ETH(22)	SV(35)	BTC_HL(34)	LTC_HL(40)	LTC_HL(35)
ETH_HL(33)	LTC_HL(33)	SP_500(21)	GLD(34)	ETH_HL(33)	ES_600(39)	ES_600(34)
			Ripple			
ETH(40)	SP_500(36)	LTC(38)	BTC_HL(37)	BTC(38)	NK_225(29)	BD_1m(38)
BD_10y(38)	BTC_HL(32)	ES_600(30)	BTC(36)	SP_500(33)	ETH(28)	GLD(36)
XRP_HL(33)	ETH(32)	VIX(30)	ETH(35)	ETH(33)	LTC(28)	XRP_HL(35)
ES_600(33)	ES_600(31)	ETH(29)	NK_225(31)	LTC(31)	BD_1m(28)	NK_225(35)
BTC_HL(32)	LTC(31)	CDS_5y(28)	Lag5(31)	BTC_HL(29)	Lag8(28)	VIX(32)
			Ethereum			
LTC_HL(41)	Lag2(39)	SP_500(38)	NK_225(35)	XRP(41)	CDS_5y(38)	CDS_5y(36)
Lag3(37)	Lag3(37)	XRP(35)	ES_600(34)	BTC_HL(38)	ES_600(38)	BD_1m(35)
BD_1m(37)	Lag4(34)	Lag3(35)	VIX(34)	NK_225(38)	Lag7(35)	SP_500(34)
ES_600(35)	BTC(27)	GLD(35)	BD_1m(33)	GLD(38)	Lag6(35)	NK_225(33)
VIX(34)	ETH_HL(25)	BTC_HL(33)	SP_500(33)	BTC(34)	LTC_HL(34)	VIX(33)

Table 7: *Top 5 crypto predictors for different cryptocurrencies and forecast horizons. Number in brackets is the average (%) inclusion probability of the selected predictor over the forecasting period.*

To provide a gauge of what drives the documented predictability, Table 7 shows the inclusion probability of the most likely five cryptopredictors. For Bitcoin, one of the other cryptocurrencies or one of the crypto–explicatives is always included; but also other assets, such as VIX, silver and the Nikkei 225 index at one day-ahead horizon, both bonds and the SP&500 at two day-ahead horizon, have large positive probabilities. Probabilities for macro and finance crypto–predictors become more relevant for longer horizons, suggesting a strengthening of the relationship with other assets, and lags of the dependent variable. The results for other currencies are qualitatively similar, with a more relevant roles for lags of the dependent variables at short horizons for Litecoin and Ethereum.

Table 7 provides the average inclusion probabilities. Figure 2 shows how weights of the corresponding five variables have evolved over time for the first horizon. We document large instabilities, in particular from the second half of 2016 to the first semester of 2017, but with sizeable differences across the four currencies. For Bitcoin, the coefficients of other cryptocurrencies, see variables Ethereum and Ripple high minus Ripple low, triple over time, even if with some drops. The interconnection with other markets, see variables Nikkei 225 and Silver, also increases over our sample. But the linkage to stock market risk, see VIX, is substantially reduced from 2016Q2. Therefore, our results suggest that Bitcoin has become more interconnected with other assets, but it has also developed its own risk intrinsic to cryptocurrencies and seems idiosyncratic to other markets.

Evidence is somewhat different for other cryptocurrencies: despite large changes in the middle of our sample, most of the coefficients assume similar values at the beginning of the sample and at the end of the sample. This is particular evident for the coefficients for lags of the dependent variable, other cryptocurrencies, and crypto–predictors. For example, the coefficient of Ethereum for Ripple starts almost at 0.4, increases to a maximum above 0.8, decreases to a minimum of almost zero, and at the end of the sample is back to a value around 0.5. A slightly different pattern has coefficients with other financial variables, and the coefficient of the Stock Europe 600 increases and doubles to a final value of 0.4 for all three cryptocurrencies. Therefore, Litecoin, Ripple and Ethereum have experienced a period of large changes in the second half of 2016 and first half of 2017, and the relationships have stabilised and become more similar to Bitcoin from the second half of 2017 when all the crypto–market has experienced an impressive increase in value.

4.3. Multivariate forecasting results

Tables 8 and 9 report MSE and predictive log score for the multivariate models. The evidence is striking and the results are almost opposite to the univariate case in terms of forecast metrics: only model M_2, a standard VAR model, provides statistically superior point forecasts at long horizons; several models provide large gains when density forecasting cryptocurrencies. Focusing on MSE results, simpler constant–parameter VAR and BVAR specifications, labelled $\mathcal{M}_2$ and $\mathcal{M}_3$, respectively, are very imprecise at short horizons with losses up to 20%, but they perform more similar to the AR(1)–EWMA at longer horizons with slight improvements for $\mathcal{M}_2$ at long horizons as discussed above. Time–varying specifications provide more similar performance across horizons, but they are never superior to the benchmark, even if several of them are included in the 5% model confidence set.

If we focus on predictive log score, several of the models in Table 9 provide statistically superior forecasts relative to the benchmark at almost all horizons. Model $\mathcal{M}_9$ (selection among a model with only cryptocurrencies and a model with also crypto–explicative), model $\mathcal{M}_{13}$ (selection among a model with only cryptocurrencies and a model with also macroeconomic variables), model $\mathcal{M}_{17}$ (selection among a model with only cryptocurrencies, a model with also crypto–explicative, a model with also macro predictors and a model with all variables) give the highest gains. In particular, $\mathcal{M}_{17}$ is always statistically superior to the benchmark; it is included in the 5% model confidence set in all horizons, and it is the only model in the confidence set at two day ahead horizon. As in the univariate case, the crypto–explicatives and macro and financial predictors improve forecast accuracy. However, in contrast to the univariate case, a selection of models containing clusters of them, instead of averaging predictors, provides the largest gains. We speculate that the importance of

h	1	2	3	4	5	6	7
$\mathcal{M}_1$	21.35	21.36	21.51	21.63	21.84	22.00	22.22
$\mathcal{M}_2$	1.12	1.11	1.07	1.01	**0.99**	**0.99**	**0.99**
$\mathcal{M}_3$	1.22	1.08	1.02	1.02	1.00	1.00	1.00
$\mathcal{M}_4$	1.02	1.02	1.01	1.01	1.00	1.00	1.00
$\mathcal{M}_5$	1.02	1.02	1.02	1.01	1.01	1.00	1.00
$\mathcal{M}_6$	1.04	1.02	1.02	1.00	1.00	1.00	1.00
$\mathcal{M}_7$	1.02	1.02	1.02	1.01	1.00	1.00	0.99
$\mathcal{M}_8$	1.01	1.02	1.01	1.01	1.00	1.00	1.00
$\mathcal{M}_9$	1.03	1.02	1.03	1.02	1.01	1.01	1.00
$\mathcal{M}_{10}$	1.02	1.02	1.02	1.01	1.00	1.01	1.00
$\mathcal{M}_{11}$	1.02	1.03	1.01	1.01	1.01	1.00	1.00
$\mathcal{M}_{12}$	1.02	1.01	1.01	1.00	1.00	1.00	1.00
$\mathcal{M}_{13}$	1.02	1.02	1.01	1.02	1.01	1.01	1.01
$\mathcal{M}_{14}$	1.02	1.02	1.02	1.01	1.01	1.01	1.00
$\mathcal{M}_{15}$	1.02	1.01	1.01	1.02	1.00	1.00	1.00
$\mathcal{M}_{16}$	1.02	1.01	1.01	1.00	0.99	0.99	1.00
$\mathcal{M}_{17}$	1.03	1.03	1.03	1.03	1.02	1.01	1.01
$\mathcal{M}_{18}$	1.02	1.03	1.03	1.01	1.01	1.01	1.00

Table 8: *(Multivariate) Mean Squared Error, computed over the forecast horizon. Results are reported relative to the benchmark specification (AR(1)–EWMA) for which the absolute score is reported. Models' description is reported in Table 3. Values in* ***bold****, indicate rejection of the null hypothesis of Equal Predictive Ability between each model and the benchmark according to the Diebold–Mariano test at the* 5% *confidence level. Gray cells indicate those models that belong to the Superior Set of Models delivered by the Model Confidence Set procedure at confidence level* 10%*.*

crypto–predictors differs across currencies and a flexible multivariate combination scheme that allows for very different weights across series and clusters of predictors could improve the accuracy, see for example Casarin et al. (2018).

When focusing on the performance over time, Figure 3 reports the cumulative predictive log score over time relative to the benchmark for three different horizons, h=1, 4 and 7. At each point in time, a positive number indicates that the alternative model outperforms the benchmark. The plots show that DMS of TVP-VAR models provides constant gains relative to the AR(1)–EWMA benchmark over all the out-of-sample period, and in some cases the increase is very large, such as at the end of March 2017 when all currencies experienced a break in volatility, see Figure 1. The models based on DMA also outperform the benchmark, but gains are lower. VAR ($\mathcal{M}_2$) and BVAR ($\mathcal{M}_3$) provide some gains in the initial part of the sample, but their performance deteriorates substantially with instability in March 2017. Only the BVAR at long horizons maintains part of its predictive power, but scores are substantially lower than those of the DMS alternatives.

Tables B.4-B.7 in Appendix B report log score results for multivariate models in predicting each cryptocurrency separately. Previous evidence is confirmed for each currency, and gains are not a result of the models performing well just for a subset of asset.

Table 10 reports the directional predictability of the multivariate models. In all cases, the success

h	1	2	3	4	5	6	7
$\mathcal{M}_1$	−3032.71	−3128.53	−3194.86	−3214.03	−3245.41	−3269.08	−3320.47
$\mathcal{M}_2$	−265.77	−140.83	−69.17	−28.24	73.97	49.22	90.07
$\mathcal{M}_3$	−358.20	−154.34	17.43	16.86	56.69	67.90	108.48
$\mathcal{M}_4$	**258.51**	169.72	163.93	168.22	145.05	**174.37**	**193.94**
$\mathcal{M}_5$	**248.60**	165.01	168.85	173.97	163.13	**185.79**	**192.71**
$\mathcal{M}_6$	**236.18**	188.96	190.42	173.58	208.37	**204.29**	**213.80**
$\mathcal{M}_7$	**261.41**	212.98	187.53	197.74	168.05	**201.18**	**208.90**
$\mathcal{M}_8$	−50.69	−33.54	−53.93	−62.35	−71.33	−82.23	−66.46
$\mathcal{M}_9$	**277.70**	232.99	**243.79**	**231.61**	**216.97**	**231.20**	**241.32**
$\mathcal{M}_{10}$	−43.00	−26.04	−28.84	−39.92	−41.94	−44.39	−32.78
$\mathcal{M}_{11}$	**256.33**	186.46	189.63	173.86	206.05	**184.02**	**188.96**
$\mathcal{M}_{12}$	−398.50	−394.92	−416.52	−422.55	−438.19	−437.01	−430.37
$\mathcal{M}_{13}$	**301.44**	233.10	213.17	**253.19**	**263.83**	**237.92**	**260.27**
$\mathcal{M}_{14}$	−380.10	−348.54	−357.57	−368.02	−378.57	−365.61	−348.23
$\mathcal{M}_{15}$	**264.02**	204.87	219.89	169.88	206.16	**203.15**	**185.84**
$\mathcal{M}_{16}$	−733.43	−705.93	−734.21	−758.45	−766.37	−769.23	−770.58
$\mathcal{M}_{17}$	**279.05**	**254.17**	**238.94**	**272.90**	**267.62**	**234.99**	**263.45**
$\mathcal{M}_{18}$	−720.52	−668.18	−703.15	−700.65	−707.64	−706.95	−701.83

Table 9: *(Multivariate) Log Score (LS), computed over the forecast horizon. Results are reported relative to the benchmark specification (AR(1)–EWMA) for which the absolute score is reported. The description of the models is reported in Table 3. Values in* ***bold****, indicate rejection of the null hypothesis of Equal Predictive Ability between each model and the benchmark according to the Diebold–Mariano test at the* 5% *confidence level. Grey cells indicate those models that belong to the Superior Set of Models delivered by the Model Confidence Set procedure at confidence level* 10%*.*

rate is higher than 50%, with higher numbers almost at 60% at two and three-day ahead horizons. Results are very similar across models. When investigating directional predictability for each currency individually, see Tables B.4-B.7, the highest gains are found in predicting Ripple, with success rates well above 60%. Ethereum seems, on the contrary, the most difficult to predict in terms of sign movements. Considering the large gains we find in log score evaluation, more advanced investment strategies with more weight on higher moments and not only on return direction could be investigated.

5. Robustness: forgetting factor comparison

Our combination schemes in equations (1) and (3) depend on three parameters to choose: 1) the forgetting factor in the variance of the measurements, κ, 2) the forgetting factor in the variance of the states, λ and 3) the forgetting factor of the models, α. In both the univariate and the multivariate models we set $\alpha = 0.99$ and $\kappa = 0.96$, while λ is selected by averaging over the grid of values $\{0.91, 0.93, 0.95, 0.97, 0.99\}$ at each point in time. In the next two subsections, we report the results for different values of α and κ in the univariate and multivariate schemes.[11]

[11] See McCormick et al. (2012) and Bork and Møller (2015) for an extension to recursive estimation of the forgetting factors.

h	1	2	3	4	5	6	7
$\mathcal{M}_1$	0.57	0.58	0.58	0.58	0.58	0.57	0.57
$\mathcal{M}_2$	0.54	0.56	0.57	0.57	0.56	0.55	0.54
$\mathcal{M}_3$	0.54	0.55	0.56	0.55	0.55	0.55	0.55
$\mathcal{M}_4$	0.56	0.56	0.57	0.57	0.56	0.57	0.56
$\mathcal{M}_5$	0.56	0.57	0.57	0.57	0.57	0.57	0.57
$\mathcal{M}_6$	0.56	0.56	0.56	0.57	0.57	0.57	0.56
$\mathcal{M}_7$	0.56	0.56	0.57	0.57	0.56	0.56	0.56
$\mathcal{M}_8$	0.57	0.57	0.56	0.57	0.56	0.55	0.56
$\mathcal{M}_9$	0.56	0.57	0.58	0.57	0.57	0.56	0.57
$\mathcal{M}_{10}$	0.56	0.58	0.57	0.57	0.57	0.57	0.57
$\mathcal{M}_{11}$	0.56	0.56	0.56	0.56	0.55	0.57	0.55
$\mathcal{M}_{12}$	0.56	0.56	0.56	0.56	0.56	0.57	0.56
$\mathcal{M}_{13}$	0.56	0.58	0.58	0.57	0.57	0.57	0.57
$\mathcal{M}_{14}$	0.56	0.57	0.58	0.57	0.57	0.57	0.57
$\mathcal{M}_{15}$	0.56	0.55	0.57	0.56	0.56	0.55	0.54

Table 10: *(Multivariate) Average Success Rate (SR), computed over the forecast horizon. The description of the models is reported in Table 3.*

5.1. Robustness: univariate models

Tables C.12–C.20 in Appendix C report the MSE, log score and success rate for different choices of the α and κ forgetting factors. For MSE and log score evaluation, the metrics for the DMA with $\alpha = 0.99$ and $\kappa = 0.96$ are computed over the out-of-sample and for the other specifications relative to them. For the success rate, the full sample statistics for all models are given. We consider eight different combinations, based on three values for $\alpha = \{0.01, 0.99, 1.00\}$ and for $\kappa = \{0.96, 0.98, 1.00\}$. The results are qualitatively similar to the main case for small variations of α and κ. In contrast, the results substantially deteriorate when $\alpha = 0.01$, indicating that inducing persistence in the selection of the models improves forecasts. Interestingly, MSE and success rate improve in several cases for $\alpha = 1.00$, but the log score decreases for the same setting.

5.2. Robustness: multivariate models

Tables C.15–C.20 in Appendix C report the MSE, log score and success rate for different choices of the κ forgetting factors. For MSE and log score evaluation, the metrics for the benchmark model AR(1)-EWMA are computed over the out-of-sample and for the other specifications relative to them. For the success rate, the full sample statistics for all models are given, we consider $\kappa = 0.96$, $\kappa = 1.00$. The results are qualitatively similar to those in Tables 8, 9 and 10, even if marginally inferior in several cases, therefore, fixing $\kappa = 0.98$ seems the better strategy. We also try to vary $\alpha = \{0.01, 1.00\}$, but results deteriorates, in particular for the log score evaluation and when applying low values of α. This evidence is similar to the univariate combinations.

6. Conclusions

This paper compares several alternative univariate and multivariate models for predicting four of the most capitalised cryptocurrencies: Bitcoin, Litecoin, Ripple, and Ethereum. A set of crypto–predictors is applied and univariate and multivariate model combinations are proposed to combine these predictors. The results show large statistically significant improvements in point forecasting of Bitcoin and Ethereum when using combinations of univariate models and in density forecasting for all the cryptocurrencies when relying on a selection of time–varying multivariate models. Both schemes deliver sizeable directional predictability gains.

We believe that our analysis opens various research agendas in predicting cryptocurrencies. For example, flexible multivariate combination schemes that allow for different weights across series could improve point and density forecast accuracy. Moreover, new predictors based on crypto–market sentiments might be considered to investigate heterogeneity across cryptocurrencies and could result in (point) forecast gains across all series.

Figure 2: Inclusion Probabilities. *The plots show inclusion probabilities of the top 5 predictors over time, computed over a one day forecast horizon, for the four cryptocurrencies.*

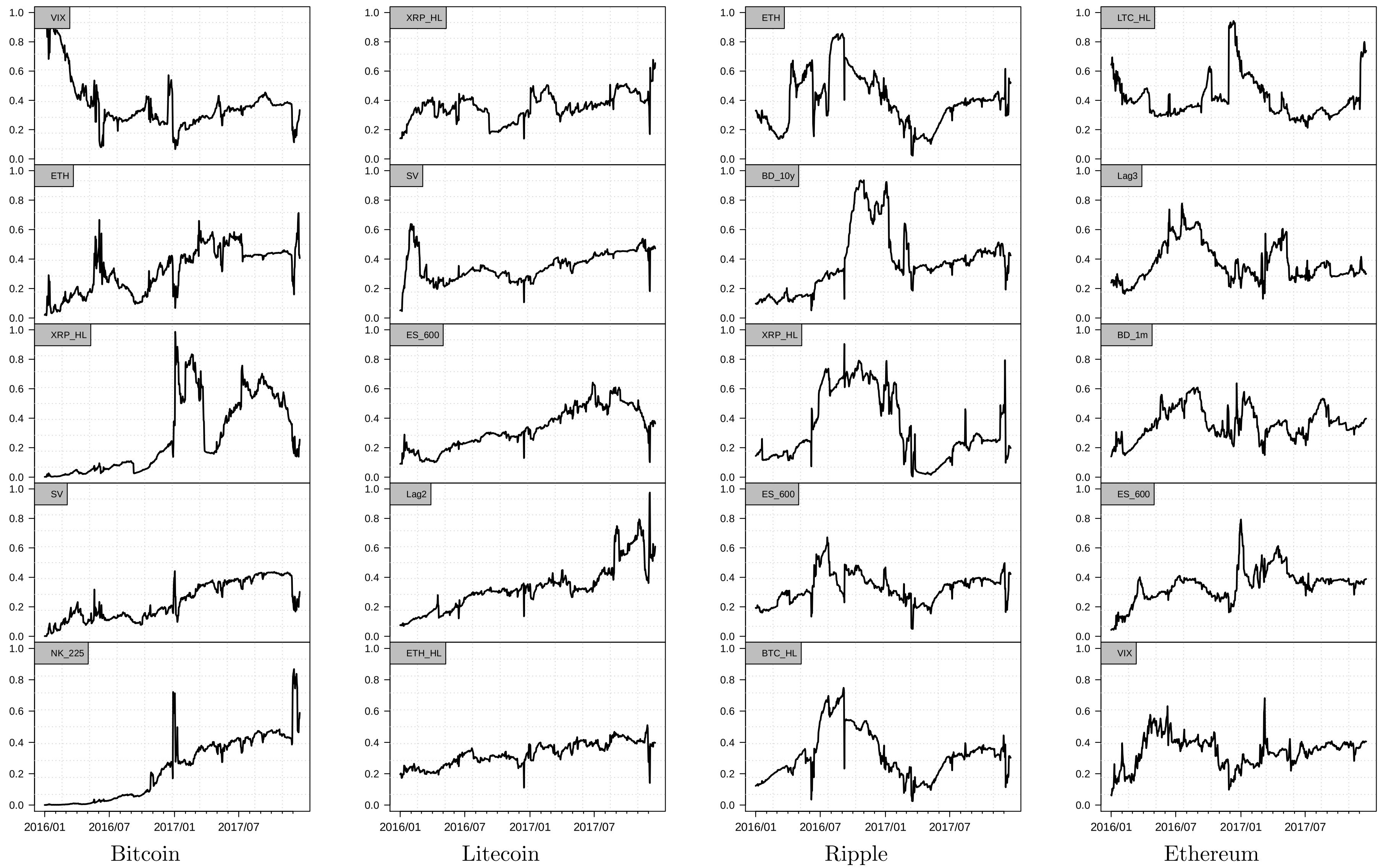
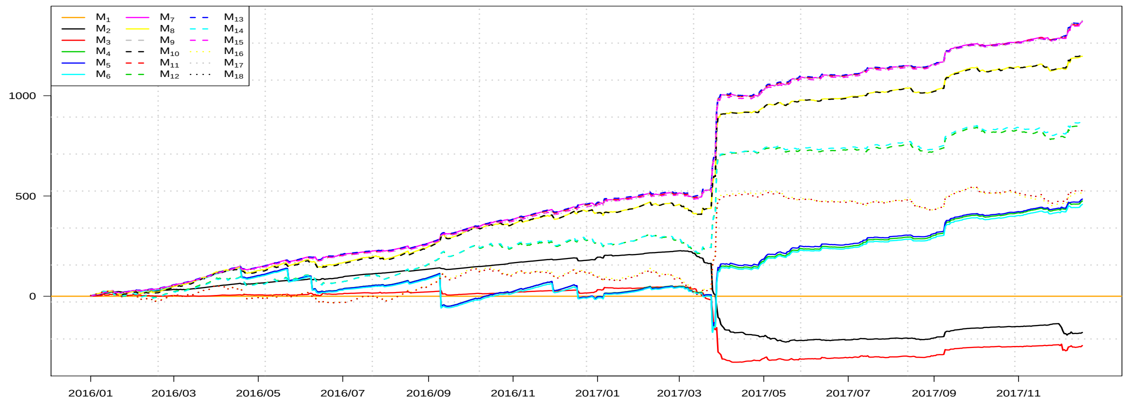
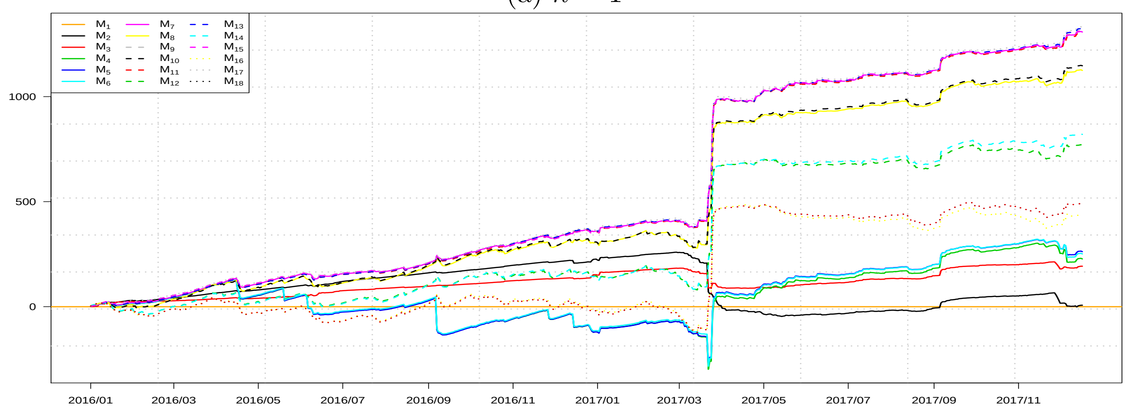


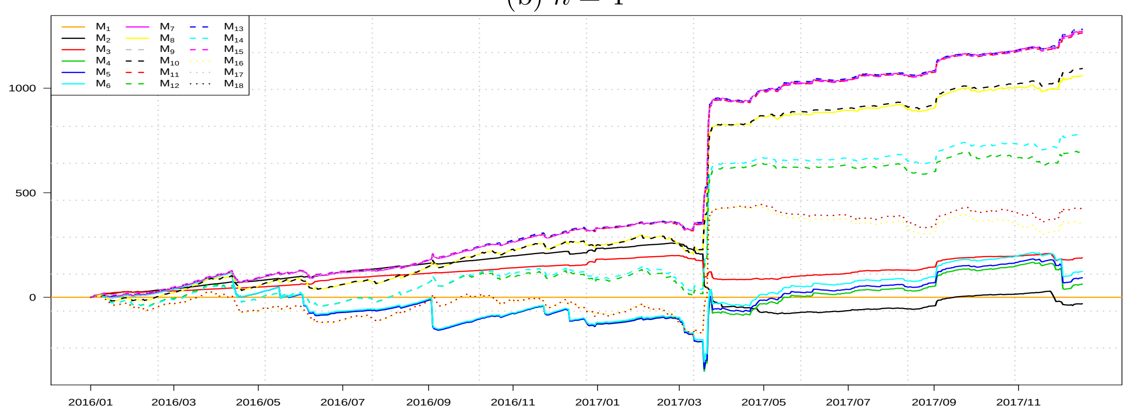
Figure 3: Cumulative Log Score. *The plots show cumulative log score relative to the AR(1)–EWMA benchmark ($\mathcal{M}_1$), computed over a one day forecast horizon (panel (a)), four day forecast horizon (panel (b)) and seven day forecast horizon (panel (c)).*



(a) $h = 1$



(b) $h = 4$



(c) $h = 7$

References

Amisano, G. and Giacomini, R. (2007). Comparing Density Forecasts via Weighted Likelihood Ratio Tests. *Journal of Business and Economic Statistics*, 25:177–190.

Barberis, N. (2000). Investing for the Long Run when Returns Are Predictable. *The Journal of Finance*, 55:225 – 264.

Bernardi, M. and Catania, L. (2016). Portfolio Optimisation Under Flexible Dynamic Dependence Modelling. *ArXiv e-prints*.

Bianchi, D. (2018). Cryptocurrencies As an Asset Class? An Empirical Assessment. Technical report, SSRN Working Paper.

Bloomberg (2017a). Japan's BITPoint to Add Bitcoin Payments to Retail Outlets. https://www.bloomberg.com/news/articles/2017-05-29/japan-s-bitpoint-to-add-bitcoin-payments-to-100-000s-of-outlets.

Bloomberg (2017b). Nasdaq Plans to Introduce Bitcoin Futures. https://www.bloomberg.com/news/articles/2017-11-29/nasdaq-is-said-to-plan-bitcoin-futures-joining-biggest-rivals.

Bloomberg (2017c). Some Central Banks Are Exploring the Use of Cryptocurrencies. https://www.bloomberg.com/news/articles/2017-06-28/rise-of-digital-coins-has-central-banks-considering-e-versions.

Bork, L. and Møller, S. V. (2015). Forecasting House Prices in the 50 States using Dynamic Model Averaging and Dynamic Model Selection. *International Journal of Forecasting*, 31(1):63–78.

Casarin, R., Grassi, S., Ravazzolo, F., and van Dijk, H. K. (2018). Predictive Density Combinations with Dynamic Learning for Large Data Sets in Economics and Finance. Technical report, Tinbergen Institute Discussion Paper 15-084/III.

Catania, L. and Grassi, S. (2018). Modelling Crypto–Currencies Financial Time–Series. Technical report, CEIS Working Paper No. 417.

Catania, L., Grassi, S., and Ravazzolo, F. (2018). Predicting the Volatility of Cryptocurrency Time-Series. Technical report, MIMEO.

Catania, L. and Nonejad, N. (2018). Dynamic Model Averaging for Practitioners in Economics and Finance: The eDMA Package. *Journal of Statistical Software, (forthcoming)*.

Chicago Mercantile Exchange (2017). CME Group Announces Launch of Bitcoin Futures. http://www.cmegroup.com/media-room/press-releases/2017/10/31/cme_group_announceslaunchofbitcoinfutures.html.

Christoffersen, P. and Diebold, F. (2006). Financial Asset Returns, Direction-of-Change Forecasting, and Volatility Dynamics. *Management Science*, 52:1273–1287.

Chu, J., Nadarajah, S., and Chan, S. (2015). Statistical Analysis of the Exchange Rate of Bitcoin. *PloS one*, 10:1–27.

Cointelegraph (2017). South Korea Officially Legalizes Bitcoin, Huge Market For Traders. https://cointelegraph.com/news/south-korea-officially-legalizes-bitcoin-huge-market-for-traders.

Dangl, T. and Halling, M. (2012). Predictive Regressions with Time–Varying Coefficients. *Journal of Financial Economics*, 106:157–181.

Diebold, F. and Mariano, R. (1995). Comparing Predictive Accuracy. *Journal of Business and Economic Stastistics*, 13:253–263.

Durbin, J. and Koopman, S. (2012). *Time Series Analysis by State Space Methods*. Oxford University Press, Oxford, UK., 2nd edition.

Ethereum (2014). Ethereum Wiki. https://github.com/ethereum/wiki/wiki/White-Paper/.

Forbes (2017). Emerging Applications For Blockchain. https://www.forbes.com/sites/forbestechcouncil/2017/07/18/emerging-applications-for-blockchain.

Geweke, J. and Amisano, G. (2010). Comparing and Evaluating Bayesian Predictive Distributions of Asset Returns. *International Journal of Forecasting*, 26:216–230.

Hansen, P. R., Lunde, A., and Nason, J. M. (2011). The Model Confidence Set. *Econometrica*, 79:453–497.

Hencic, A. and Gourieroux, C. (2014). Noncausal Autoregressive Model in Application to Bitcoin/USD Exchange Rate. *Proceedings of the 7th Financial Risks International Forum*, pages 1–25.

Hotz-Behofsits, C., Huber, F., and Zorner, T. O. (2018). Predicting Crypto-currencies Using Sparse Non–Gaussian State Space Models. *Journal of Statistical Software, (forthcoming).*

Johannes, M., Korteweg, A., and Polson, N. (2014). Sequential Learning, Predictive Regressions, and Optimal Portfolio Returns. *Journal of Finance*, 69:611–644.

Koop, G. (2003). *Bayesian Econometrics*. Wiley.

Koop, G. and Korobilis, D. (2010). Forecasting Inflation Using Dynamic Model Averaging. *Bayesian Multivariate Time Series Methods for Empirical Macroeconomics*, 3:267–358.

Koop, G. and Korobilis, D. (2011). UK Macroeconomic Forecasting with Many Predictors: Which Models Forecast Best and When Do They Do So? *Economic Modelling*, 28:2307–2318.

Koop, G. and Korobilis, D. (2012). Forecasting Inflation Using Dynamic Model Averaging. *International Economic Review*, 53:867–886.

Koop, G. and Korobilis, D. (2013). Large Time-Varying Parameter VARs. *Journal of Econometrics*, 177:185–198.

Litecoin (2014). Litecoin Wiki. https://litecoin.info/Litecoin/.

Lutkepohl, H. (2007). *New Introduction to Multiple Time Series Analysis*. Springer Publishing Company, Incorporated.

Marcellino, M., Stock, J. H., and Watson, M. W. (2006). A Comparison of Direct and Iterated Multistep AR Methods for Forecasting Macroeconomic Time Series . *Journal of Econometrics*, 135:499 – 526.

McCormick, T. H., Raftery, A. E., Madigan, D., and Burd, R. S. (2012). Dynamic Logistic Regression and Dynamic Model Averaging for Binary Classification. *Biometrics*, 68:23–30.

Nakamoto, S. (2009). Bitcoin: A Peer-to-Peer Electronic Cash System. https://Bitcoin.org/Bitcoin.pdf.

Pastor, L. (2000). Portfolio Selection and Asset Pricing Models. *Journal of Finance*, 55:179–223.

Pastor, L. and Stambaugh, R. F. (2000). Comparing Asset Pricing Models: an Investment Perspective. *Journal of Financial Economics*, 56:335–381.

Prado, R. and West, M. (2010). *Time series: Modeling, Computation, and Inference*. CRC Press, Boca Raton.

Raftery, A. E., Kárnỳ, M., and Ettler, P. (2010). Online Prediction Under Model Uncertainty via Dynamic Model Averaging: Application to a Cold Rolling Mill. *Technometrics*, 52:52–66.

Ripple (2012). Welcome to Ripple. https://ripple.com/.

Sapuric, S. and Kokkinaki, A. (2014). Bitcoin is Volatile! Isn't That Right? *Business Information Systems Workshops*, pages 255–265.

Stambaugh, R. F. (1999). Predictive regressions. *Journal of Financial Economics*, 54:375 – 421.

Tokyo Financial Exchange (2017). Tokyo Financial Exchange Plans for Bitcoin Futures Launch. https://www.cryptocoinsnews.com/breaking-tokyo-financial-exchange-plans-bitcoin-futures-launch/.

West, M. and Harrison, J. (1999). *Bayesian Forecasting & Dynamic Models*. Springer–Verlag, Berlin.

Online Supplementary Material

This online supplementary material contains additional material for the paper "Forecasting Cryptocurrencies Under Model and Parameter Instability". It contains detailed information on data in Appendix A; forecasting results for alternative score metrics in Appendix B; and for different forgetting factors in Appendix C.

Appendix A. Overview of the data series

Data Overview		
Abbreviation	Full name	Transformation
Cryptocurrencies time series		
BTC	Bitcoin	First difference of Log
ETH	Ethereum	First difference of Log
XRP	Ripple	First difference of Log
LTC	Litecoin	First difference of Log
Additional crypto–explicative time series		
BTC_HL	Bitcoin High minus Bitcoin Low	Log
ETH_HL	Ethereum High minus Ethereum Low	Log
XRP_HL	Ripple High minus Ripple Low	Log
LTC_HL	Litecoin High minus Litecoin Low	Log
Additional financial and macro time series		
CDS_5y	Europe credit default swap index 5 years	First difference of Log
ES_600	Stoxx Europe 600 - Price Index	First difference of Log
GLD	Gold Bullion LBM	First difference of Log
NK_225	Nikkei 225 Stock Average - Price Index	First difference of Log
SP_500	S&P 500 Composite - Price Index	First difference of Log
SV	Silver Handy & Harman Base Price	First difference of Log
BD_1m	1-Month US Treasury Constant Maturity Rate	First difference
BD_10y	10-Year US Treasury Constant Maturity Rate	First difference
VIX	VIX closing price	Log

Table A.1: *This table provides an overview of the data we use in the paper. The table reports the main four cryptocurrencies used in the paper. The series are available over the period August 8, 2015 to December 28, 2017. The table also reports the additional cryptocurrency time series, e.g. Bitcoin High minus Bitcoin Low together with the macro and financial time series used in the study. For each series the table reports the abbreviation we use in the paper, the full name of the series, and the transformation applied on the raw data series.*

Appendix B. Results for alternative forecasting metrics

h	1	2	3	4	5	6	7
				Bitcoin			
AR(1)–EWMA	42.49	42.28	41.54	41.62	41.55	41.42	41.12
KS	1.52	12.25	1.84	**0.96**	1.06	1.06	1.12
KS–NR	2.07	7.03	1.61	1.02	1.01	1.03	1.01
DMA–NR	**0.97**	**0.98**	0.99	**0.98**	0.99	**0.98**	**0.98**
DMS–NR	**0.98**	0.99	1.00	1.00	1.00	1.00	1.00
DMA	**0.95**	**0.98**	1.01	0.99	0.98	0.99	1.03
BMA	1.19	1.04	1.11	**0.93**	1.00	1.03	1.02
DMS	0.99	1.01	1.03	0.99	0.99	1.01	1.07
				Litecoin			
AR(1)–EWMA	134.27	132.88	133.05	133.43	133.60	133.25	131.71
KS	1.02	1.17	7.64	1.01	1.17	1.11	1.09
KS–NR	**0.96**	1.03	1.88	1.00	1.01	1.02	1.00
DMA–NR	1.00	1.02	1.07	1.01	1.00	1.03	1.02
DMS–NR	1.01	1.04	1.08	1.00	1.01	1.02	1.01
DMA	0.98	1.02	1.06	1.01	1.06	1.03	1.07
BMA	**0.97**	0.99	1.18	**0.99**	1.00	**0.97**	1.00
DMS	1.00	1.07	1.12	1.02	1.07	1.03	1.04
	0.88	**0.89**					
				Ripple			
AR(1)–EWMA	224.02	221.31	222.02	221.13	218.93	219.62	219.45
KS	1.11	1.24	1.27	1.10	1.76	1.21	2.01
KS–NR	1.03	1.04	1.02	1.01	1.08	1.00	1.02
DMA–NR	1.00	0.99	1.00	1.01	1.00	1.00	1.00
DMS–NR	1.01	0.99	1.01	1.00	1.01	1.02	1.02
DMA	1.03	1.03	1.07	1.06	1.02	1.05	1.08
BMA	1.03	0.99	1.01	1.02	0.99	1.00	1.00
DMS	1.08	1.06	1.12	1.11	1.07	1.10	1.10
				Ethereum			
AR(1)–EWMA	180.57	174.99	175.61	175.56	175.79	175.90	174.08
KS	1.05	12.72	1.09	1.01	1.02	1.67	1.09
KS–NR	1.01	3.40	1.02	1.00	1.00	1.01	1.00
DMA–NR	**0.97**	1.00	**0.97**	**0.98**	0.99	0.99	0.99
DMS–NR	0.98	1.02	0.99	0.99	0.99	1.00	1.00
DMA	**0.96**	1.01	1.03	1.01	1.03	1.03	1.03
BMA	1.00	1.19	1.00	1.00	1.00	1.00	1.00
DMS	**0.97**	1.03	1.05	1.02	1.04	1.06	1.03

Table B.2: *Mean absolute deviation (MAD), computed over the forecast horizon. Results are reported relative to the benchmark specification (AR(1)–EWMA), for which the absolute score is reported. The description of the models is reported in Table 2. Values in* ***bold****, indicate rejection of the null hypothesis of Equal Predictive Ability between each model and the benchmark according to the Diebold–Mariano test at the* 5% *confidence level. Grey cells indicate those models that belong to the Superior Set of Models delivered by the Model Confidence Set procedure at confidence level* 10%.

h	1	2	3	4	5	6	7
$\mathcal{M}_1$	3.06	3.06	3.07	3.08	3.09	3.10	3.12
$\mathcal{M}_2$	1.07	1.05	1.03	1.01	1.00	1.00	1.00
$\mathcal{M}_3$	1.12	1.04	1.01	1.01	1.00	1.00	1.00
$\mathcal{M}_4$	1.01	1.01	1.01	1.00	1.00	1.00	1.00
$\mathcal{M}_5$	1.01	1.01	1.01	1.01	1.00	1.00	1.00
$\mathcal{M}_6$	1.02	1.02	1.02	1.00	1.00	1.00	1.00
$\mathcal{M}_7$	1.01	1.01	1.01	1.01	1.00	1.00	1.00
$\mathcal{M}_8$	1.01	1.01	1.01	1.01	1.00	1.00	1.00
$\mathcal{M}_9$	1.02	1.01	1.01	1.01	1.00	1.00	1.00
$\mathcal{M}_{10}$	1.01	1.01	1.01	1.01	1.00	1.00	1.00
$\mathcal{M}_{11}$	1.02	1.02	1.01	1.01	1.01	1.00	1.00
$\mathcal{M}_{12}$	1.01	1.01	1.01	1.00	1.00	1.00	1.00
$\mathcal{M}_{13}$	1.02	1.01	1.01	1.01	1.00	1.00	1.00
$\mathcal{M}_{14}$	1.02	1.01	1.01	1.01	1.00	1.00	1.00
$\mathcal{M}_{15}$	1.02	1.01	1.01	1.01	1.01	1.01	1.01
$\mathcal{M}_{16}$	1.01	1.01	1.01	1.01	1.01	1.00	1.00
$\mathcal{M}_{17}$	1.02	1.02	1.01	1.01	1.01	1.01	1.00
$\mathcal{M}_{18}$	1.02	1.02	1.02	1.01	1.01	1.01	1.01

Table B.3: *(Multivariate) Mean Absolute Error, computed over the forecast horizon. Results are reported relative to the benchmark specification (AR(1)–EWMA), for which the absolute score is reported. The description of the models is reported in Table 3. Values in* ***bold****, indicate rejection of the null hypothesis of Equal Predictive Ability between each model and the benchmark according to the Diebold–Mariano test at the* 5% *confidence level. Grey cells indicate those models that belong to the Superior Set of Models delivered by the Model Confidence Set procedure at confidence level* 10%.

h	1	2	3	4	5	6	7
$\mathcal{M}_1$	−799.65	−802.52	−803.13	−811.52	−818.94	−822.67	−828.14
$\mathcal{M}_2$	−83.62	−70.45	−77.41	−63.28	−59.95	−59.96	−58.21
$\mathcal{M}_3$	−139.80	−98.78	−103.01	−82.20	−83.30	−75.40	−71.47
$\mathcal{M}_4$	1.35	0.72	−0.10	−0.18	−0.20	0.85	−2.98
$\mathcal{M}_5$	−1.57	1.94	−3.69	−5.03	−2.71	−3.05	−1.43
$\mathcal{M}_6$	1.90	−4.98	−12.14	−0.50	−1.03	0.35	−2.35
$\mathcal{M}_7$	0.02	−5.17	−2.20	−6.57	0.69	−4.44	−5.72
$\mathcal{M}_8$	−39.98	−54.13	−55.50	−57.50	−54.20	−77.98	−79.71
$\mathcal{M}_9$	−5.88	−2.23	−8.38	−3.75	3.48	−4.03	−2.84
$\mathcal{M}_{10}$	−45.75	−40.01	−41.64	−50.35	−55.64	−67.51	−68.28
$\mathcal{M}_{11}$	**3.39**	−2.72	−5.40	−5.29	0.77	0.75	−4.64
$\mathcal{M}_{12}$	−174.03	−193.94	−209.12	−211.17	−229.21	−244.83	−262.44
$\mathcal{M}_{13}$	−5.20	−9.10	−4.01	−7.95	−1.18	−1.68	0.38
$\mathcal{M}_{14}$	−158.77	−172.02	−180.94	−197.67	−203.57	−227.86	−238.87
$\mathcal{M}_{15}$	−5.71	−2.81	−5.90	−5.20	0.82	1.28	−4.29
$\mathcal{M}_{16}$	−356.51	−356.41	−384.19	−397.08	−407.88	−423.10	−441.40
$\mathcal{M}_{17}$	−5.40	−5.99	−10.19	−4.43	−0.58	0.85	−3.07
$\mathcal{M}_{18}$	−333.63	−356.43	−373.80	−378.88	−421.76	−411.15	−413.44

Table B.4: *Bitcoin. Predictive Log Score (LS), computed over the forecast horizon. Results are reported relative to the benchmark specification (AR(1)–EWMA). The description of the models is reported in Table 2. Values in* ***bold****, indicate rejection of the null hypothesis of Equal Predictive Ability between each model and the benchmark according to the Diebold–Mariano test at the* 5% *confidence level. Grey cells indicate those models that belong to the Superior Set of Models delivered by the Model Confidence Set procedure at confidence level* 10%.

h	1	2	3	4	5	6	7
$\mathcal{M}_1$	−655.37	−668.20	−672.74	−683.92	−687.56	−687.58	−691.48
$\mathcal{M}_2$	−115.30	−98.65	−97.49	−84.69	−91.12	−84.15	−81.58
$\mathcal{M}_3$	−184.93	−120.47	−93.94	−82.56	−72.94	−76.97	−76.64
$\mathcal{M}_4$	**1.78**	−1.54	−2.31	1.09	−1.72	0.86	**1.89**
$\mathcal{M}_5$	−2.75	−2.41	−1.62	1.06	−2.12	−1.12	1.61
$\mathcal{M}_6$	−0.70	−4.62	−2.26	0.83	−3.56	−3.06	−0.22
$\mathcal{M}_7$	−0.55	−5.16	−5.08	−1.14	−0.19	−3.90	2.17
$\mathcal{M}_8$	8.38	−4.00	−22.39	−36.53	−29.80	−39.11	−22.60
$\mathcal{M}_9$	−3.20	−1.24	−7.01	−6.35	0.66	−5.37	−0.06
$\mathcal{M}_{10}$	5.25	1.18	−8.43	−29.13	−15.32	−19.62	−19.90
$\mathcal{M}_{11}$	−3.69	−4.50	−2.06	−2.14	−5.27	−4.49	2.07
$\mathcal{M}_{12}$	−85.93	−95.24	−122.83	−161.37	−189.99	−170.87	−161.15
$\mathcal{M}_{13}$	−5.17	−1.46	−4.59	−6.62	−3.21	−2.07	−1.97
$\mathcal{M}_{14}$	−80.80	−78.70	−131.03	−140.47	−157.03	−135.93	−140.20
$\mathcal{M}_{15}$	−3.28	−9.91	−1.47	−2.12	−6.84	−3.89	−1.83
$\mathcal{M}_{16}$	−184.54	−244.85	−269.80	−325.59	−334.94	−300.74	−327.37
$\mathcal{M}_{17}$	−4.39	−9.56	−8.12	−4.20	−3.16	−3.18	−0.86
$\mathcal{M}_{18}$	−216.98	−214.67	−271.82	−291.12	−305.30	−332.57	−307.79

Table B.5: *Litecoin. Predictive Log Score (LS), computed over the forecast horizon. Results are reported relative to the benchmark specification (AR(1)–EWMA). The description of the models is reported in Table 2. Values in* ***bold****, indicate rejection of the null hypothesis of Equal Predictive Ability between each model and the benchmark according to the Diebold–Mariano test at the* 5% *confidence level. Grey cells indicate those models that belong to the Superior Set of Models delivered by the Model Confidence Set procedure at confidence level* 10%.

h	1	2	3	4	5	6	7
$\mathcal{M}_1$	−543.42	−550.18	−561.32	−563.28	−564.31	−563.93	−562.63
$\mathcal{M}_2$	−123.66	−115.56	−104.62	−51.87	−37.91	−43.33	−46.92
$\mathcal{M}_3$	−131.35	−91.38	−64.09	−68.03	−57.56	−56.58	−58.58
$\mathcal{M}_4$	**9.02**	−1.67	−1.95	−2.71	−0.10	−1.27	−3.84
$\mathcal{M}_5$	4.25	0.05	−3.57	−3.23	−2.54	−2.39	−3.14
$\mathcal{M}_6$	−4.74	0.31	−5.76	−0.29	−1.99	−2.32	−0.39
$\mathcal{M}_7$	2.88	1.80	−1.56	−2.74	−3.65	−5.80	−8.27
$\mathcal{M}_8$	**37.95**	23.51	7.10	0.75	4.46	−4.84	−7.22
$\mathcal{M}_9$	−2.67	−6.52	−7.12	−6.82	−7.21	−6.01	−6.76
$\mathcal{M}_{10}$	**43.66**	21.45	9.23	1.24	13.12	6.74	8.30
$\mathcal{M}_{11}$	**6.49**	−1.13	−1.05	−0.30	−2.64	−3.20	−2.16
$\mathcal{M}_{12}$	−41.57	−57.06	−79.43	−86.50	−80.21	−87.45	−95.10
$\mathcal{M}_{13}$	−4.98	−6.61	−9.65	−8.98	−3.50	−1.73	−6.98
$\mathcal{M}_{14}$	−8.04	−46.50	−75.38	−75.31	−56.27	−57.43	−72.59
$\mathcal{M}_{15}$	2.02	−6.63	1.69	−1.46	−3.99	−2.29	−9.77
$\mathcal{M}_{16}$	−145.04	−154.63	−181.16	−201.94	−201.72	−188.66	−193.79
$\mathcal{M}_{17}$	−2.35	−6.13	−6.67	−4.75	−7.11	−5.78	−10.53
$\mathcal{M}_{18}$	−123.18	−162.74	−182.94	−182.59	−198.53	−160.03	−171.59

Table B.6: *Ripple. Predictive Log Scores (LS), computed over the forecast horizon. Results are reported relative to the benchmark specification (AR(1)–EWMA). The description of the models is reported in Table 2. Values in* ***bold****, indicate rejection of the null hypothesis of Equal Predictive Ability between each model and the benchmark according to the Diebold–Mariano test at the* 5% *confidence level. Grey cells indicate those models that belong to the Superior Set of Models delivered by the Model Confidence Set procedure at confidence level* 10%.

h	1	2	3	4	5	6	7
$\mathcal{M}_1$	−730.06	−738.61	−744.18	−748.66	−752.92	−759.41	−758.45
$\mathcal{M}_2$	−106.58	−90.70	−76.36	−60.81	−55.77	−48.66	−47.51
$\mathcal{M}_3$	−140.05	−101.17	−79.11	−78.87	−74.19	−66.43	−68.17
$\mathcal{M}_4$	2.93	**5.78**	**4.84**	3.16	3.85	**9.15**	**8.55**
$\mathcal{M}_5$	2.17	**6.02**	**6.08**	3.68	3.64	**8.78**	4.73
$\mathcal{M}_6$	−7.52	−1.43	3.13	0.99	3.95	**8.39**	**6.99**
$\mathcal{M}_7$	**6.11**	1.85	4.24	2.43	1.91	5.40	3.23
$\mathcal{M}_8$	−40.38	−52.34	−52.78	−67.46	−62.37	−67.25	−67.01
$\mathcal{M}_9$	0.46	1.09	6.51	−0.35	−0.33	4.50	**6.94**
$\mathcal{M}_{10}$	−39.79	−38.25	−39.47	−40.67	−54.89	−57.10	−54.49
$\mathcal{M}_{11}$	0.22	−8.44	2.41	−6.82	−5.19	−2.13	5.11
$\mathcal{M}_{12}$	−191.65	−199.68	−201.81	−212.25	−226.38	−223.10	−236.68
$\mathcal{M}_{13}$	−6.06	1.05	5.07	−3.12	1.92	4.47	2.53
$\mathcal{M}_{14}$	−176.02	−175.90	−187.46	−188.05	−200.68	−199.58	−199.79
$\mathcal{M}_{15}$	−6.79	−9.76	−6.13	−13.35	−5.12	−3.09	−1.29
$\mathcal{M}_{16}$	−352.33	−361.88	−368.92	−410.72	−399.24	−412.16	−401.72
$\mathcal{M}_{17}$	−10.67	−7.81	2.64	−7.10	0.36	−8.22	−2.43
$\mathcal{M}_{18}$	−326.46	−329.77	−369.17	−372.67	−367.76	−382.22	−376.20

Table B.7: *Ethereum. Predictive Log Score (LS), computed over the forecast horizon. Results are reported relative to the benchmark specification (AR(1)–EWMA). The description of the models is reported in Table 2. Values in* ***bold****, indicate rejection of the null hypothesis of Equal Predictive Ability between each model and the benchmark according to the Diebold–Mariano test at the* 5% *confidence level. Grey cells indicate those models that belong to the Superior Set of Models delivered by the Model Confidence Set procedure at confidence level* 10%.

h	1	2	3	4	5	6	7
$\mathcal{M}_1$	0.53	0.53	0.53	0.53	0.53	0.53	0.52
$\mathcal{M}_2$	0.53	0.54	0.57	0.56	0.54	0.53	0.52
$\mathcal{M}_3$	0.51	0.50	0.56	0.53	0.51	0.53	0.52
$\mathcal{M}_4$	0.52	0.53	0.54	0.55	0.53	0.54	0.54
$\mathcal{M}_5$	0.55	0.55	0.56	0.56	0.55	0.56	0.56
$\mathcal{M}_6$	0.52	0.53	0.53	0.55	0.54	0.55	0.54
$\mathcal{M}_7$	0.52	0.52	0.54	0.55	0.52	0.53	0.53
$\mathcal{M}_8$	0.54	0.56	0.55	0.54	0.53	0.52	0.54
$\mathcal{M}_9$	0.55	0.54	0.57	0.55	0.54	0.55	0.54
$\mathcal{M}_{10}$	0.55	0.57	0.55	0.56	0.55	0.54	0.55
$\mathcal{M}_{11}$	0.54	0.54	0.53	0.53	0.51	0.54	0.52
$\mathcal{M}_{12}$	0.53	0.54	0.53	0.54	0.53	0.55	0.53
$\mathcal{M}_{13}$	0.55	0.55	0.57	0.54	0.55	0.55	0.55
$\mathcal{M}_{14}$	0.56	0.56	0.59	0.55	0.54	0.56	0.54
$\mathcal{M}_{15}$	0.53	0.53	0.55	0.52	0.53	0.53	0.52
$\mathcal{M}_{16}$	0.53	0.52	0.53	0.54	0.54	0.54	0.53
$\mathcal{M}_{17}$	0.55	0.53	0.57	0.56	0.55	0.57	0.52
$\mathcal{M}_{18}$	0.55	0.56	0.57	0.57	0.54	0.55	0.55

Table B.8: *Bitcoin. Success Rate (SR), computed over the forecast horizon. The description of the models is reported in Table 2.*

h	1	2	3	4	5	6	7
$\mathcal{M}_1$	0.55	0.58	0.58	0.58	0.58	0.58	0.57
$\mathcal{M}_2$	0.56	0.58	0.60	0.58	0.58	0.58	0.57
$\mathcal{M}_3$	0.52	0.58	0.56	0.57	0.58	0.58	0.57
$\mathcal{M}_4$	0.58	0.58	0.58	0.58	0.58	0.58	0.57
$\mathcal{M}_5$	0.59	0.58	0.58	0.58	0.58	0.58	0.58
$\mathcal{M}_6$	0.57	0.55	0.57	0.58	0.58	0.58	0.57
$\mathcal{M}_7$	0.57	0.57	0.58	0.58	0.59	0.57	0.57
$\mathcal{M}_8$	0.58	0.58	0.58	0.58	0.58	0.57	0.57
$\mathcal{M}_9$	0.57	0.60	0.58	0.57	0.58	0.56	0.57
$\mathcal{M}_{10}$	0.58	0.59	0.58	0.57	0.58	0.57	0.57
$\mathcal{M}_{11}$	0.56	0.57	0.57	0.57	0.57	0.58	0.57
$\mathcal{M}_{12}$	0.58	0.58	0.58	0.58	0.58	0.58	0.57
$\mathcal{M}_{13}$	0.58	0.59	0.58	0.58	0.58	0.57	0.58
$\mathcal{M}_{14}$	0.58	0.58	0.58	0.58	0.58	0.58	0.57
$\mathcal{M}_{15}$	0.58	0.56	0.57	0.60	0.59	0.57	0.57
$\mathcal{M}_{16}$	0.58	0.58	0.58	0.58	0.58	0.58	0.58
$\mathcal{M}_{17}$	0.57	0.58	0.57	0.57	0.57	0.57	0.57
$\mathcal{M}_{18}$	0.57	0.59	0.58	0.58	0.58	0.58	0.58

Table B.9: *Litecoin. Predictive Log Score (LS), computed over the forecast horizon. Success Rate (SR), computed over the forecast horizon. The description of the models is reported in Table 2.*

h	1	2	3	4	5	6	7
$\mathcal{M}_1$	0.64	0.65	0.65	0.64	0.64	0.64	0.64
$\mathcal{M}_2$	0.61	0.63	0.63	0.64	0.65	0.65	0.63
$\mathcal{M}_3$	0.60	0.62	0.61	0.63	0.65	0.65	0.64
$\mathcal{M}_4$	0.64	0.65	0.65	0.64	0.64	0.64	0.64
$\mathcal{M}_5$	0.63	0.65	0.65	0.64	0.64	0.64	0.64
$\mathcal{M}_6$	0.63	0.65	0.65	0.65	0.65	0.64	0.64
$\mathcal{M}_7$	0.64	0.65	0.64	0.64	0.64	0.64	0.64
$\mathcal{M}_8$	0.64	0.65	0.64	0.64	0.64	0.64	0.64
$\mathcal{M}_9$	0.62	0.64	0.65	0.64	0.64	0.64	0.64
$\mathcal{M}_{10}$	0.63	0.64	0.65	0.64	0.65	0.64	0.64
$\mathcal{M}_{11}$	0.64	0.65	0.65	0.65	0.65	0.64	0.65
$\mathcal{M}_{12}$	0.64	0.65	0.64	0.65	0.65	0.65	0.64
$\mathcal{M}_{13}$	0.63	0.66	0.63	0.64	0.65	0.64	0.64
$\mathcal{M}_{14}$	0.63	0.64	0.65	0.64	0.65	0.64	0.64
$\mathcal{M}_{15}$	0.64	0.64	0.65	0.64	0.65	0.64	0.64
$\mathcal{M}_{16}$	0.65	0.65	0.65	0.64	0.64	0.64	0.64
$\mathcal{M}_{17}$	0.63	0.64	0.65	0.64	0.65	0.63	0.65
$\mathcal{M}_{18}$	0.63	0.64	0.64	0.65	0.65	0.64	0.65

Table B.10: *Ripple. Success Rate (SR), computed over the forecast horizon. The description of the models is reported in Table 2.*

h	1	2	3	4	5	6	7
$\mathcal{M}_1$	0.55	0.56	0.56	0.55	0.56	0.55	0.55
$\mathcal{M}_2$	0.47	0.48	0.48	0.49	0.46	0.46	0.45
$\mathcal{M}_3$	0.51	0.50	0.53	0.49	0.47	0.46	0.45
$\mathcal{M}_4$	0.50	0.50	0.50	0.50	0.50	0.50	0.50
$\mathcal{M}_5$	0.49	0.49	0.50	0.49	0.50	0.50	0.50
$\mathcal{M}_6$	0.50	0.49	0.50	0.50	0.51	0.50	0.49
$\mathcal{M}_7$	0.52	0.49	0.52	0.51	0.49	0.49	0.49
$\mathcal{M}_8$	0.51	0.49	0.49	0.51	0.50	0.48	0.50
$\mathcal{M}_9$	0.50	0.51	0.53	0.50	0.51	0.51	0.52
$\mathcal{M}_{10}$	0.49	0.50	0.51	0.51	0.50	0.53	0.50
$\mathcal{M}_{11}$	0.49	0.47	0.49	0.50	0.48	0.51	0.47
$\mathcal{M}_{12}$	0.49	0.48	0.49	0.49	0.48	0.49	0.49
$\mathcal{M}_{13}$	0.50	0.51	0.53	0.50	0.51	0.51	0.52
$\mathcal{M}_{14}$	0.49	0.49	0.51	0.50	0.51	0.49	0.51
$\mathcal{M}_{15}$	0.49	0.46	0.50	0.49	0.49	0.47	0.44
$\mathcal{M}_{16}$	0.47	0.46	0.46	0.47	0.48	0.46	0.46
$\mathcal{M}_{17}$	0.48	0.47	0.51	0.53	0.48	0.49	0.50
$\mathcal{M}_{18}$	0.48	0.49	0.49	0.49	0.49	0.49	0.48

Table B.11: *Ethereum. Success Rate (SR), computed over the forecast horizon. The description of the models is reported in Table 2.*

Appendix C. Forgetting factor comparison

h	1	2	3	4	5	6	7
				Bitcoin			
$\alpha = 0.99,\ \kappa = 0.96$	17.03	16.14	16.83	16.39	15.91	15.68	15.63
$\alpha = 0.99,\ \kappa = 0.98$	1.05	1.03	1.01	1.01	1.01	1.05	1.04
$\alpha = 0.99,\ \kappa = 1.00$	1.04	1.04	1.03	1.13	1.15	1.14	1.17
$\alpha = 0.01,\ \kappa = 0.96$	1.10	1.20	1.10	1.11	1.05	1.05	1.05
$\alpha = 0.01,\ \kappa = 0.98$	1.10	1.26	1.10	1.11	1.04	1.05	1.07
$\alpha = 0.01,\ \kappa = 1.00$	1.08	1.41	1.10	1.14	1.04	1.05	1.08
$\alpha = 1.00,\ \kappa = 0.96$	0.96	1.01	0.99	0.97	0.98	1.00	0.99
$\alpha = 1.00,\ \kappa = 0.98$	0.96	1.01	1.01	0.98	0.98	1.01	0.99
$\alpha = 1.00,\ \kappa = 1.00$	0.96	0.99	1.04	0.99	0.98	1.03	1.00
				Litecoin			
$\alpha = 0.99,\ \kappa = 0.96$	40.71	42.89	44.11	42.02	47.79	44.29	47.16
$\alpha = 0.99,\ \kappa = 0.98$	1.03	1.01	1.04	1.04	1.14	1.04	1.02
$\alpha = 0.99,\ \kappa = 1.00$	1.13	1.03	1.09	1.63	1.81	1.10	1.38
$\alpha = 0.01,\ \kappa = 0.96$	1.17	1.15	1.12	1.10	0.99	1.16	1.16
$\alpha = 0.01,\ \kappa = 0.98$	1.16	1.12	1.11	1.09	0.97	1.12	1.06
$\alpha = 0.01,\ \kappa = 1.00$	1.12	1.10	1.13	1.08	0.95	1.10	0.96
$\alpha = 1.00,\ \kappa = 0.96$	1.01	0.96	1.09	0.99	0.87	0.95	0.90
$\alpha = 1.00,\ \kappa = 0.98$	1.01	0.96	1.05	0.99	0.90	0.96	0.93
$\alpha = 1.00,\ \kappa = 1.00$	1.02	1.00	1.03	0.99	1.40	1.05	1.08
				Ripple			
$\alpha = 0.99,\ \kappa = 0.96$	82.66	80.02	87.58	80.41	77.52	85.31	79.12
$\alpha = 0.99,\ \kappa = 0.98$	1.40	1.03	1.19	1.10	1.04	1.04	1.03
$\alpha = 0.99,\ \kappa = 1.00$	1.81	1.48	1.49	2.45	1.58	1.26	1.22
$\alpha = 0.01,\ \kappa = 0.96$	1.16	1.09	0.96	1.02	1.18	1.05	1.12
$\alpha = 0.01,\ \kappa = 0.98$	1.17	1.08	0.96	1.05	1.17	1.06	1.13
$\alpha = 0.01,\ \kappa = 1.00$	1.05	1.08	0.92	1.07	1.33	1.11	1.07
$\alpha = 1.00,\ \kappa = 0.96$	0.94	0.97	0.88	0.97	1.01	0.92	0.99
$\alpha = 1.00,\ \kappa = 0.98$	0.94	0.96	0.93	0.96	1.00	0.92	0.99
$\alpha = 1.00,\ \kappa = 1.00$	1.37	1.09	1.39	1.68	1.49	1.01	1.17
				Ethereum			
$\alpha = 0.99,\ \kappa = 0.96$	35.00	33.95	35.15	34.40	34.32	33.53	32.75
$\alpha = 0.99,\ \kappa = 0.98$	1.04	1.04	1.05	1.06	1.05	1.04	1.02
$\alpha = 0.99,\ \kappa = 1.00$	1.14	1.04	1.05	1.13	1.31	1.07	1.06
$\alpha = 0.01,\ \kappa = 0.96$	1.13	1.02	1.00	1.13	1.10	1.04	1.03
$\alpha = 0.01,\ \kappa = 0.98$	1.14	1.03	1.00	1.14	1.12	1.03	1.03
$\alpha = 0.01,\ \kappa = 1.00$	1.19	1.06	1.02	1.15	1.20	1.06	1.05
$\alpha = 1.00,\ \kappa = 0.96$	0.94	1.02	0.93	0.94	0.94	0.97	0.98
$\alpha = 1.00,\ \kappa = 0.98$	0.95	1.02	0.95	0.94	0.94	0.98	0.98
$\alpha = 1.00,\ \kappa = 1.00$	0.96	1.03	0.97	0.94	0.95	0.99	0.99

Table C.12: *Mean squared error (MSE), computed over the forecast horizon delivered by DMA. Results are reported for different choices of the κ and α forgetting factors and relative to the case $\alpha = 0.99$, $\kappa = 0.96$. The lines $\alpha = 0.99$, $\kappa = 0.96$ report the value computed over the out-of-sample.*

h	1	2	3	4	5	6	7
				Bitcoin			
$\alpha = 0.99,\ \kappa = 0.96$	-906.17	-898.87	-908.26	-904.52	-890.25	-889.08	-882.27
$\alpha = 0.99,\ \kappa = 0.98$	-34.78	-32.48	-38.26	-33.76	-28.80	-31.90	-27.50
$\alpha = 0.99,\ \kappa = 1.00$	-75.34	-77.69	-76.86	-104.76	-90.72	-82.86	-85.29
$\alpha = 0.01,\ \kappa = 0.96$	-31.70	-139.53	-39.73	-27.73	-27.09	-26.49	-25.14
$\alpha = 0.01,\ \kappa = 0.98$	-60.50	-276.01	-62.10	-53.31	-46.38	-42.67	-41.48
$\alpha = 0.01,\ \kappa = 1.00$	-173.94	-827.34	-143.54	-86.76	-64.03	-63.12	-54.13
$\alpha = 1.00,\ \kappa = 0.96$	-13.37	-13.79	-12.27	2.55	2.53	0.79	2.86
$\alpha = 1.00,\ \kappa = 0.98$	-51.58	-50.06	-48.36	-36.90	-36.16	-37.05	-36.74
$\alpha = 1.00,\ \kappa = 1.00$	-140.01	-125.47	-133.81	-127.37	-122.15	-117.92	-125.04
				Litecoin			
$\alpha = 0.99,\ \kappa = 0.96$	-822.13	-816.99	-841.70	-821.18	-804.94	-807.72	-803.30
$\alpha = 0.99,\ \kappa = 0.98$	-41.07	-36.66	-44.27	-42.59	-42.56	-37.27	-39.53
$\alpha = 0.99,\ \kappa = 1.00$	-187.69	-146.77	-97.64	-206.97	-195.80	-111.28	-194.18
$\alpha = 0.01,\ \kappa = 0.96$	-44.11	-39.74	-83.30	-23.41	-33.57	-29.39	-45.83
$\alpha = 0.01,\ \kappa = 0.98$	-69.07	-57.14	-163.71	-50.95	-55.45	-48.50	-52.32
$\alpha = 0.01,\ \kappa = 1.00$	-130.05	-111.07	-438.49	-111.87	-83.58	-64.50	-60.26
$\alpha = 1.00,\ \kappa = 0.96$	2.19	-3.10	-19.12	3.20	-6.68	-4.30	6.47
$\alpha = 1.00,\ \kappa = 0.98$	-50.19	-55.23	-63.86	-45.68	-55.10	-54.89	-41.93
$\alpha = 1.00,\ \kappa = 1.00$	-241.58	-252.10	-229.15	-241.10	-238.26	-223.95	-230.05
				Ripple			
$\alpha = 0.99,\ \kappa = 0.96$	-724.62	-741.91	-719.44	-721.60	-716.38	-718.92	-716.62
$\alpha = 0.99,\ \kappa = 0.98$	-52.85	-57.00	-57.03	-49.92	-54.12	-62.69	-57.58
$\alpha = 0.99,\ \kappa = 1.00$	-277.09	-176.80	-195.24	-299.45	-151.24	-248.40	-136.68
$\alpha = 0.01,\ \kappa = 0.96$	-39.08	-39.88	-37.36	-29.53	-55.98	-37.07	-41.04
$\alpha = 0.01,\ \kappa = 0.98$	-68.24	-75.64	-51.88	-61.88	-92.49	-74.05	-82.61
$\alpha = 0.01,\ \kappa = 1.00$	-82.20	-103.96	-37.66	-68.15	-96.93	-162.56	-78.69
$\alpha = 1.00,\ \kappa = 0.96$	-8.74	8.95	-11.56	-8.03	-12.10	2.26	-9.85
$\alpha = 1.00,\ \kappa = 0.98$	-77.68	-61.54	-67.78	-75.00	-78.63	-63.24	-79.40
$\alpha = 1.00,\ \kappa = 1.00$	-329.90	-289.33	-282.77	-312.51	-318.64	-310.23	-287.73
				Ethereum			
$\alpha = 0.99,\ \kappa = 0.96$	-817.83	-818.68	-813.02	-818.70	-808.63	-807.10	-806.69
$\alpha = 0.99,\ \kappa = 0.98$	-23.70	-21.83	-21.59	-22.45	-22.63	-22.14	-20.87
$\alpha = 0.99,\ \kappa = 1.00$	-66.40	-70.55	-57.45	-56.44	-60.81	-63.42	-58.29
$\alpha = 0.01,\ \kappa = 0.96$	-22.49	-59.33	-23.04	-28.79	-27.14	-28.00	-22.02
$\alpha = 0.01,\ \kappa = 0.98$	-40.63	-103.73	-35.06	-42.87	-42.87	-43.17	-38.09
$\alpha = 0.01,\ \kappa = 1.00$	-78.26	-408.79	-68.12	-78.94	-80.36	-86.98	-72.18
$\alpha = 1.00,\ \kappa = 0.96$	6.15	-10.59	7.15	11.36	1.91	-0.82	6.53
$\alpha = 1.00,\ \kappa = 0.98$	-20.95	-34.36	-19.31	-13.41	-22.24	-25.61	-18.67
$\alpha = 1.00,\ \kappa = 1.00$	-61.64	-73.20	-58.32	-49.08	-56.63	-60.41	-54.04

Table C.13: *Log Score (LS), computed over the forecast horizon delivered by DMA. Results are reported for different choices of the* κ *and* α *forgetting factors and in excess to the case* $\alpha = 0.99$, $\kappa = 0.96$. *The lines* $\alpha = 0.99$, $\kappa = 0.96$ *report the value computed over the out-of-sample.*

h	1	2	3	4	5	6	7
				Bitcoin			
$\alpha = 0.99,\ \kappa = 0.96$	0.53	0.55	0.55	0.57	0.56	0.55	0.53
$\alpha = 0.99,\ \kappa = 0.98$	0.53	0.55	0.53	0.57	0.57	0.55	0.52
$\alpha = 0.99,\ \kappa = 1.00$	0.54	0.54	0.53	0.56	0.56	0.56	0.52
$\alpha = 0.01,\ \kappa = 0.96$	0.53	0.55	0.54	0.55	0.57	0.54	0.53
$\alpha = 0.01,\ \kappa = 0.98$	0.54	0.54	0.54	0.55	0.56	0.53	0.53
$\alpha = 0.01,\ \kappa = 1.00$	0.52	0.53	0.54	0.57	0.56	0.54	0.53
$\alpha = 1.00,\ \kappa = 0.96$	0.54	0.56	0.57	0.58	0.55	0.56	0.56
$\alpha = 1.00,\ \kappa = 0.98$	0.55	0.57	0.56	0.57	0.55	0.54	0.54
$\alpha = 1.00,\ \kappa = 1.00$	0.57	0.56	0.54	0.56	0.55	0.55	0.53
				Litecoin			
$\alpha = 0.99,\ \kappa = 0.96$	0.51	0.52	0.55	0.50	0.52	0.55	0.53
$\alpha = 0.99,\ \kappa = 0.98$	0.52	0.52	0.54	0.49	0.52	0.55	0.52
$\alpha = 0.99,\ \kappa = 1.00$	0.54	0.53	0.56	0.53	0.52	0.55	0.55
$\alpha = 0.01,\ \kappa = 0.96$	0.51	0.51	0.56	0.51	0.51	0.52	0.50
$\alpha = 0.01,\ \kappa = 0.98$	0.50	0.51	0.57	0.51	0.51	0.52	0.51
$\alpha = 0.01,\ \kappa = 1.00$	0.52	0.52	0.56	0.53	0.52	0.53	0.52
$\alpha = 1.00,\ \kappa = 0.96$	0.57	0.58	0.54	0.57	0.56	0.55	0.57
$\alpha = 1.00,\ \kappa = 0.98$	0.58	0.57	0.54	0.56	0.56	0.53	0.57
$\alpha = 1.00,\ \kappa = 1.00$	0.57	0.57	0.55	0.56	0.56	0.55	0.57
				Ripple			
$\alpha = 0.99,\ \kappa = 0.96$	0.58	0.58	0.60	0.57	0.61	0.58	0.58
$\alpha = 0.99,\ \kappa = 0.98$	0.58	0.58	0.62	0.58	0.61	0.57	0.59
$\alpha = 0.99,\ \kappa = 1.00$	0.59	0.57	0.61	0.58	0.60	0.57	0.60
$\alpha = 0.01,\ \kappa = 0.96$	0.56	0.55	0.58	0.55	0.56	0.55	0.60
$\alpha = 0.01,\ \kappa = 0.98$	0.57	0.54	0.58	0.54	0.57	0.56	0.60
$\alpha = 0.01,\ \kappa = 1.00$	0.59	0.56	0.59	0.55	0.59	0.56	0.61
$\alpha = 1.00,\ \kappa = 0.96$	0.60	0.61	0.58	0.60	0.60	0.60	0.60
$\alpha = 1.00,\ \kappa = 0.98$	0.59	0.59	0.59	0.59	0.59	0.59	0.59
$\alpha = 1.00,\ \kappa = 1.00$	0.62	0.60	0.63	0.61	0.60	0.59	0.63
				Ethereum			
$\alpha = 0.99,\ \kappa = 0.96$	0.55	0.52	0.53	0.51	0.53	0.52	0.54
$\alpha = 0.99,\ \kappa = 0.98$	0.55	0.53	0.53	0.49	0.53	0.52	0.53
$\alpha = 0.99,\ \kappa = 1.00$	0.55	0.53	0.51	0.49	0.52	0.53	0.54
$\alpha = 0.01,\ \kappa = 0.96$	0.53	0.54	0.49	0.51	0.52	0.53	0.54
$\alpha = 0.01,\ \kappa = 0.98$	0.54	0.54	0.49	0.50	0.52	0.53	0.54
$\alpha = 0.01,\ \kappa = 1.00$	0.55	0.54	0.50	0.50	0.52	0.51	0.54
$\alpha = 1.00,\ \kappa = 0.96$	0.55	0.56	0.57	0.57	0.58	0.56	0.57
$\alpha = 1.00,\ \kappa = 0.98$	0.55	0.57	0.56	0.56	0.56	0.56	0.56
$\alpha = 1.00,\ \kappa = 1.00$	0.54	0.56	0.53	0.56	0.57	0.55	0.56

Table C.14: *Average success rate (SR), computed over the forecast horizon delivered by DMA. Results are reported for different choices of the κ and α forgetting factors.*

h	1	2	3	4	5	6	7
$\mathcal{M}_1$	21.35	21.36	21.51	21.63	21.84	22.00	22.22
$\mathcal{M}_2$	1.12	1.11	1.07	1.01	**0.99**	**0.99**	**0.99**
$\mathcal{M}_3$	1.22	1.08	1.02	1.02	1.00	1.00	1.00
$\mathcal{M}_4$	1.10	1.09	1.04	1.02	1.01	1.02	1.02
$\mathcal{M}_5$	1.14	1.11	1.08	1.06	1.04	1.04	1.05
$\mathcal{M}_6$	1.10	1.10	1.05	1.03	1.01	1.03	1.03
$\mathcal{M}_7$	1.11	1.09	1.10	1.07	1.04	1.04	1.04
$\mathcal{M}_8$	1.12	1.12	1.11	1.06	1.06	1.04	1.05
$\mathcal{M}_9$	1.16	1.14	1.12	1.12	1.11	1.11	1.10
$\mathcal{M}_{10}$	1.19	1.21	1.16	1.13	1.12	1.11	1.11
$\mathcal{M}_{11}$	1.13	1.12	1.11	1.08	1.05	1.04	1.02
$\mathcal{M}_{12}$	1.14	1.19	1.14	1.10	1.09	1.08	1.07
$\mathcal{M}_{13}$	1.18	1.20	1.16	1.12	1.12	1.11	1.09
$\mathcal{M}_{14}$	1.23	1.29	1.20	1.16	1.14	1.12	1.14
$\mathcal{M}_{15}$	1.15	1.12	1.11	1.07	1.05	1.08	1.05
$\mathcal{M}_{16}$	1.17	1.21	1.16	1.12	1.12	1.10	1.10
$\mathcal{M}_{17}$	1.23	1.21	1.17	1.13	1.14	1.15	1.12
$\mathcal{M}_{18}$	1.28	1.33	1.30	1.23	1.23	1.22	1.23

Table C.15: *(Multivariate) Mean Squared Error, computed over the forecast horizon as in Table 8. The multivariate specifications have* $\kappa = 1.0$*. Results are reported relative to the benchmark specification (AR(1)–EWMA), for which the absolute score is reported. The description of the models is reported in Table 3. Values in* ***bold****, indicate rejection of the null hypothesis of Equal Predictive Ability between each model and the benchmark according to the Diebold–Mariano test at the* 5% *confidence level. Grey cells indicate those models that belong to the Superior Set of Models delivered by the Model Confidence Set procedure at confidence level* 10%.

h	1	2	3	4	5	6	7
$\mathcal{M}_1$	21.35	21.36	21.51	21.63	21.84	22.00	22.22
$\mathcal{M}_2$	1.12	1.11	1.07	1.01	**0.99**	**0.99**	**0.99**
$\mathcal{M}_3$	1.22	1.08	1.02	1.02	1.00	1.00	1.00
$\mathcal{M}_4$	1.02	1.03	1.02	1.01	1.01	1.00	1.00
$\mathcal{M}_5$	1.03	1.04	1.03	1.01	1.00	1.01	1.00
$\mathcal{M}_6$	1.04	1.05	1.03	1.01	1.01	1.00	1.00
$\mathcal{M}_7$	1.03	1.03	1.03	1.01	1.01	1.01	1.00
$\mathcal{M}_8$	1.03	1.02	1.03	1.00	1.01	1.00	1.00
$\mathcal{M}_9$	1.03	1.05	1.04	1.02	1.01	1.02	1.01
$\mathcal{M}_{10}$	1.04	1.05	1.03	1.02	1.01	1.02	1.01
$\mathcal{M}_{11}$	1.03	1.04	1.02	1.00	1.01	1.01	1.01
$\mathcal{M}_{12}$	1.03	1.03	1.02	1.01	1.01	1.01	1.00
$\mathcal{M}_{13}$	1.04	1.05	1.03	1.02	1.02	1.02	1.01
$\mathcal{M}_{14}$	1.05	1.06	1.03	1.02	1.01	1.02	1.01
$\mathcal{M}_{15}$	1.03	1.05	1.01	1.02	1.00	1.01	1.00
$\mathcal{M}_{16}$	1.04	1.03	1.02	1.02	1.01	1.00	1.00
$\mathcal{M}_{17}$	1.04	1.05	1.05	1.03	1.01	1.02	1.02
$\mathcal{M}_{18}$	1.04	1.06	1.04	1.02	1.03	1.02	1.01

Table C.16: *(Multivariate) Mean Squared Error, computed over the forecast horizon as in Table 8. The multivariate specifications have $\kappa = 0.98$. Results are reported relative to the benchmark specification (AR(1)–EWMA), for which the absolute score is reported. The description of the models is reported in Table 3. Values in **bold**, indicate rejection of the null hypothesis of Equal Predictive Ability between each model and the benchmark according to the Diebold–Mariano test at the 5% confidence level. Grey cells indicate those models that belong to the Superior Set of Models delivered by the Model Confidence Set procedure at confidence level 10%.*

h	1	2	3	4	5	6	7
$\mathcal{M}_1$	−3032.71	−3128.53	−3194.86	−3214.03	−3245.41	−3269.08	−3320.47
$\mathcal{M}_2$	−265.77	−140.83	−69.17	−28.24	73.97	49.22	90.07
$\mathcal{M}_3$	−358.20	−154.34	17.43	16.86	56.69	67.90	108.48
$\mathcal{M}_4$	−736.81	−567.63	−500.48	−440.95	−455.39	−433.20	−332.04
$\mathcal{M}_5$	−869.07	−637.64	−498.03	−420.37	−368.26	−364.40	−309.10
$\mathcal{M}_6$	−792.13	−591.23	−530.42	−378.90	−450.36	−348.05	−325.55
$\mathcal{M}_7$	−593.58	−471.42	−403.16	−322.62	−357.77	−261.49	−203.60
$\mathcal{M}_8$	−666.34	−564.04	−512.40	−425.31	−467.74	−422.23	−324.92
$\mathcal{M}_9$	−632.02	−446.94	−314.36	−232.72	−249.70	−209.25	−121.76
$\mathcal{M}_{10}$	−702.48	−552.86	−409.33	−345.25	−384.04	−310.21	−227.45
$\mathcal{M}_{11}$	−573.59	−450.55	−333.45	−313.59	−312.26	−261.55	−191.55
$\mathcal{M}_{12}$	−949.36	−825.07	−732.61	−657.07	−741.33	−674.13	−573.28
$\mathcal{M}_{13}$	−649.55	−362.81	−199.26	−190.96	−205.97	−180.08	−75.07
$\mathcal{M}_{14}$	−933.55	−659.39	−482.93	−458.08	−483.62	−452.03	−355.83
$\mathcal{M}_{15}$	−571.60	−398.49	−256.56	−207.41	−266.80	−272.17	−173.03
$\mathcal{M}_{16}$	−1236.27	−1109.68	−1014.52	−915.93	−981.46	−927.81	−839.90
$\mathcal{M}_{17}$	−683.91	−250.55	−170.61	−177.44	−198.92	−196.40	−128.24
$\mathcal{M}_{18}$	−1261.80	−907.25	−696.98	−693.74	−697.02	−641.33	−569.57

Table C.17: *(Multivariate) Log Score (LS), computed over the forecast horizon. Results are reported relative to the benchmark specification (AR(1)–EWMA), for which the absolute score is reported. The multivariate specifications have $\kappa = 1.0$. The description of the models is reported in Table 3. Values in* ***bold****, indicate rejection of the null hypothesis of Equal Predictive Ability between each model and the benchmark according to the Diebold–Mariano test at the 5% confidence level. Grey cells indicate those models that belong to the Superior Set of Models delivered by the Model Confidence Set procedure at confidence level 10%.*

h	1	2	3	4	5	6	7
$\mathcal{M}_1$	−3032.71	−3128.53	−3194.86	−3214.03	−3245.41	−3269.08	−3320.47
$\mathcal{M}_2$	−265.77	−140.83	−69.17	−28.24	73.97	49.22	90.07
$\mathcal{M}_3$	−358.20	−154.34	17.43	16.86	56.69	67.90	108.48
$\mathcal{M}_4$	170.85	**216.56**	**245.96**	**258.68**	**269.56**	**245.17**	**290.46**
$\mathcal{M}_5$	175.71	**213.81**	**238.10**	**258.16**	**270.19**	**243.71**	**287.40**
$\mathcal{M}_6$	152.14	**218.58**	**234.23**	**259.14**	**273.67**	**268.26**	**300.06**
$\mathcal{M}_7$	176.56	**225.46**	**234.42**	**230.81**	**263.73**	**215.85**	**279.31**
$\mathcal{M}_8$	−41.21	−22.04	10.61	5.45	15.09	23.95	56.92
$\mathcal{M}_9$	176.05	**218.96**	**240.88**	**250.77**	**258.17**	**237.56**	**293.71**
$\mathcal{M}_{10}$	−9.56	1.07	19.13	29.44	36.61	52.54	65.99
$\mathcal{M}_{11}$	184.90	208.37	**234.98**	**237.17**	**241.50**	**225.30**	**284.28**
$\mathcal{M}_{12}$	−321.04	−296.99	−257.04	−257.81	−271.95	−255.62	−249.06
$\mathcal{M}_{13}$	179.40	203.64	**235.72**	**254.59**	**248.76**	**249.53**	**291.09**
$\mathcal{M}_{14}$	−278.02	−243.32	−206.15	−213.99	−217.68	−205.84	−184.97
$\mathcal{M}_{15}$	176.58	211.66	215.73	**229.20**	**244.29**	**233.80**	**264.50**
$\mathcal{M}_{16}$	−632.08	−608.86	−571.05	−590.59	−598.55	−595.82	−580.09
$\mathcal{M}_{17}$	176.43	**211.38**	**239.66**	**223.63**	**275.18**	**248.31**	**272.32**
$\mathcal{M}_{18}$	−596.38	−535.16	−493.37	−515.59	−517.89	−499.58	−498.91

Table C.18: *(Multivariate) Log Score (LS), computed over the forecast horizon. Results are reported relative to the benchmark specification (AR(1)–EWMA), for which the absolute score is reported. The multivariate specifications have $\kappa = 0.98$. The description of the models is reported in Table 3. Values in **bold**, indicate rejection of the null hypothesis of Equal Predictive Ability between each model and the benchmark according to the Diebold–Mariano test at the 5% confidence level. Grey cells indicate those models that belong to the Superior Set of Models delivered by the Model Confidence Set procedure at confidence level 10%.*

h	1	2	3	4	5	6	7
$\mathcal{M}_1$	0.57	0.58	0.58	0.58	0.58	0.57	0.57
$\mathcal{M}_2$	0.54	0.56	0.57	0.57	0.56	0.55	0.54
$\mathcal{M}_3$	0.54	0.55	0.56	0.55	0.55	0.55	0.55
$\mathcal{M}_4$	0.56	0.57	0.57	0.56	0.57	0.57	0.57
$\mathcal{M}_5$	0.57	0.57	0.57	0.56	0.56	0.56	0.56
$\mathcal{M}_6$	0.57	0.57	0.56	0.56	0.57	0.56	0.56
$\mathcal{M}_7$	0.56	0.55	0.56	0.56	0.56	0.56	0.55
$\mathcal{M}_8$	0.55	0.55	0.56	0.55	0.55	0.56	0.55
$\mathcal{M}_9$	0.57	0.57	0.56	0.56	0.56	0.57	0.56
$\mathcal{M}_{10}$	0.57	0.57	0.57	0.56	0.56	0.56	0.56
$\mathcal{M}_{11}$	0.55	0.55	0.55	0.55	0.57	0.55	0.56
$\mathcal{M}_{12}$	0.56	0.56	0.56	0.56	0.56	0.56	0.56
$\mathcal{M}_{13}$	0.56	0.56	0.57	0.57	0.56	0.56	0.56
$\mathcal{M}_{14}$	0.57	0.57	0.57	0.56	0.57	0.57	0.56
$\mathcal{M}_{15}$	0.56	0.56	0.56	0.56	0.55	0.57	0.56
$\mathcal{M}_{16}$	0.55	0.56	0.56	0.56	0.56	0.57	0.56
$\mathcal{M}_{17}$	0.56	0.56	0.56	0.56	0.56	0.56	0.56
$\mathcal{M}_{18}$	0.57	0.57	0.57	0.56	0.57	0.56	0.56

Table C.19: *(Multivariate) Average Success Rate (SR), computed over the forecast horizon. The multivariate specifications have $\kappa = 1.0$. The description of the models is reported in Table 3.*

h	1	2	3	4	5	6	7
$\mathcal{M}_1$	0.57	0.58	0.58	0.58	0.58	0.57	0.57
$\mathcal{M}_2$	0.54	0.56	0.57	0.57	0.56	0.55	0.54
$\mathcal{M}_3$	0.54	0.55	0.56	0.55	0.55	0.55	0.55
$\mathcal{M}_4$	0.57	0.57	0.57	0.57	0.57	0.56	0.56
$\mathcal{M}_5$	0.57	0.58	0.58	0.58	0.57	0.57	0.57
$\mathcal{M}_6$	0.56	0.56	0.57	0.57	0.56	0.57	0.56
$\mathcal{M}_7$	0.55	0.55	0.56	0.55	0.56	0.57	0.55
$\mathcal{M}_8$	0.57	0.56	0.56	0.56	0.57	0.57	0.56
$\mathcal{M}_9$	0.56	0.58	0.58	0.57	0.56	0.57	0.57
$\mathcal{M}_{10}$	0.56	0.58	0.57	0.57	0.57	0.57	0.57
$\mathcal{M}_{11}$	0.56	0.55	0.56	0.56	0.55	0.56	0.55
$\mathcal{M}_{12}$	0.56	0.56	0.56	0.57	0.57	0.56	0.56
$\mathcal{M}_{13}$	0.56	0.57	0.57	0.57	0.57	0.57	0.56
$\mathcal{M}_{14}$	0.57	0.58	0.57	0.57	0.58	0.58	0.57
$\mathcal{M}_{15}$	0.56	0.55	0.56	0.56	0.56	0.56	0.55
$\mathcal{M}_{16}$	0.56	0.56	0.57	0.56	0.56	0.56	0.55
$\mathcal{M}_{17}$	0.57	0.58	0.58	0.56	0.57	0.57	0.56
$\mathcal{M}_{18}$	0.56	0.57	0.57	0.57	0.58	0.57	0.57

Table C.20: *(Multivariate) Average Success Rate (SR), computed over the forecast horizon. The multivariate specifications have $\kappa = 0.98$. Models' description is reported in Table 3.*