International Journal of Forecasting 35 (2019) 485–501

Contents lists available at ScienceDirect

International Journal of Forecasting

journal homepage: www.elsevier.com/locate/ijforecast

Forecasting cryptocurrencies under model and parameter instability

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

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

ARTICLE INFO

Keywords:
Cryptocurrency
Bitcoin
Forecasting
Density forecasting
VAR
Dynamic model averaging

ABSTRACT

This paper studies the predictability of cryptocurrency time series. We compare several alternative univariate and multivariate models for 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 the selection of multivariate models. Both schemes deliver sizable directional predictability.

1. Introduction

Bitcoin was the first decentralized cryptocurrency, having been created in 2009 and documented by Nakamoto (2009). It has been gaining increasing amounts of attention from the media, academics, the finance industry ever since its introduction, and recent months have seen the global interest in Bitcoin and cryptocurrencies spike dramatically. There are numerous reasons for this intensified interest, and we mention only 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 cryptocurrencies (Bloomberg, 2017c); and a large number of companies and banks created the Enterprise Ethereum Alliance[1] to make use of 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) and Tokyo Financial Exchange (2017).

This interest has been reflected in the market capitalization of the cryptocurrencies, which exploded from around 19 billion in February 2017 to around 800 billion in December 2017, and more than 1000 cryptocurrencies. Although Bitcoin is a relatively new currency, there has already been some initial analysis of it. 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) measured the volatility of the Bitcoin exchange rate against those of six major currencies. Chu, Nadarajah, and Chan (2015) provided a statistical analysis of the log-returns of the exchange rate of Bitcoin versus the USD. Catania and Grassi (2018) analysed the main characteristics of the cryptocurrency volatility. Hotz-Behofsits, Huber, and Zorner (2018) used a time-varying parameter VAR with t-distributed measurement errors and stochastic volatility to model cryptocurrencies.

However, despite all this effort, no detailed analysis of the forecasting performances of different models on this

* Corresponding author at: Faculty of Economics and Management, Free University of Bozen-Bolzano, Italy.
E-mail address: francesco.ravazzolo@unibz.it (F. Ravazzolo).

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

https://doi.org/10.1016/j.ijforecast.2018.09.005
0169-2070/

series has yet been provided. The present paper tries to fill this gap by comparing a large set of different models for point and density forecasting of four of the most capitalised cryptocurrencies, namely Bitcoin, Litecoin, Ripple and Ethereum. We compare univariate autoregressive models to univariate linear regression models based on a large set of crypto-predictors. These predictors include commodity prices, other financial assets such as stock and bond prices, and volatility indices for proxying market sentiments, following evidence by Bianchi (2018) that returns on cryptocurrencies are moderately correlated (in-sample analysis) with commodities and a 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, Kárnỳ, and Ettler (2010). DMS and DMA have been found to provide forecasting gains in macroeconomic applications (see for example Koop & Korobilis, 2011, and Koop & Korobilis, 2012), but have not yet been applied to cryptocurrencies. Then, we generalise the exercise to multivariate models where we predict the four series jointly using vector autoregressive (VAR), Bayesian VAR, and time–varying parameter VAR models with stochastic volatility as per Koop and Korobilis (2013), selecting and averaging these models with different degrees of smoothness and different sets of predictors; see among others 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 an application of model averaging to stock price prediction and Johannes, Korteweg, and Polson (2014) for the use of time-varying parameter 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.

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

Our results show that the 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. These gains are economically and statistically significant for both Bitcoin (up to 11% at the one-day-ahead horizon) and Ethereum (up to 14% at the one-day-ahead horizon). The evidence for Litecoin and Ripple is weaker. When focusing on density forecasting, the gains when predicting Bitcoin and Ethereum disappear, even if moderate improvements emerge for Litecoin and Ripple. Thus, combining a large set of predictors increases the point forecast accuracy for the major currencies, but does not improve the density forecasting.

When we focus on multivariate models we get conflicting evidence, as only a few models provide marginally more accurate point predictions, and then for longer horizons. In general, point predictability does not emerge. However, most of the multivariate schemes offer statically significant gains at all horizons when the complete distribution is predicted. In particular, the selection of time-varying VARs with different sets of predictors and different levels of smoothness provides the largest gains. Our findings corroborate and extend the evidence presented by Hotz-Behofsits et al. (2018), who found sizable improvements from the use of their time-varying parameter VAR in density forecasting but not in point forecasting, and by Catania, Grassi, and Ravazzolo (2018), who showed that (more) sophisticated volatility models can improve volatility predictions at different forecast horizons, relative 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 remainder of the paper proceeds as follows. Section 2 provides details of 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 on the cryptocurrencies. The crypto-market is open 24 h a day, seven days a week; hence, when computing returns, we use the closing price at midnight (UTC). These data are available freely from CoinMarketcap.[2]

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

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 all 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 has many advantages, such as being free from government control and regulation, but critics often argue that there is nobody overlooking the whole system apart from its users, and that the value of Bitcoin is unfounded. In spite of this criticism, though, it touched 20,000 USD in December 2017 and is above 10,000 USD at the time of writing, a steep increase from its starting point of a few cents in 2010.

2 https://coinmarketcap.com/.

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.

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
Minimum	−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

Notes: The table reports the names of the coins, their creation dates, the maximum numbers of coins in millions and billions, and the market capitalization in December 2017, as reported at https://coinmarketcap.com/. Ethereum has a total supply of 18 million coins per year; the other three have pre-set total amounts.

Ethereum is a decentralised platform which features smart contract functionality that facilitates online contractual agreement applications that eliminate any possibility of downtime, censorship, fraud, or third party interference. Ethereum also provides cryptocurrency tokens 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 both 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 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 is used by Bitcoin, for which reason it is often considered Bitcoin's leading rival. It has one main feature that distinguishes it from Bitcoin: it conducts transactions significantly faster, and is therefore particularly attractive in time-critical situations; see Litecoin (2014).

We collect data in the sample between August 8, 2015, and 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. The series are far from being normally distributed, as was documented in Chu et al. (2015), and display a high volatility, non-zero skewness, very high kurtosis and several spikes; see Fig. 1.

2.2. Crypto-predictors

At present, cryptocurrencies are mainly considered as an alternative investment, since their use for 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 and Stoxx Europe 600); commodity prices (gold and silver prices); interest rates and CDS (the 5-year Europe credit default swap and the 1-month and 10-year US Treasury rates); a volatility index and the VIX closing price. See Table A.1 in Appendix A for the data transformations.

We also apply lags of each of the four cryptocurrencies, as well as a transformation of the previous day's cryptocurrencies, labelled in the paper as "crypto-explicative", to account for intra-day patterns by taking the difference between the highest and lowest prices as a proxy for the volatility of the cryptocurrencies.

3. Competing models

This section introduces our different models for producing daily forecasts for the cryptocurrencies. As anticipated, we consider univariate and multivariate models with and without exogenous variables, and selections and combinations of such models. The full lists of univariate and multivariate models are reported in Tables 2 and 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

3.1.1. 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 the "kitchen sink" (KS). We also consider a restricted version where we include only lags of the cryptocurrencies and crypto-explicatives, labelled KS-NR.

3.1.2. Model combinations

The previous linear models can suffer from massive model uncertainty. Indeed, a model with 16 predictors and up to three lags of the dependent variable results in more

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 the best results.

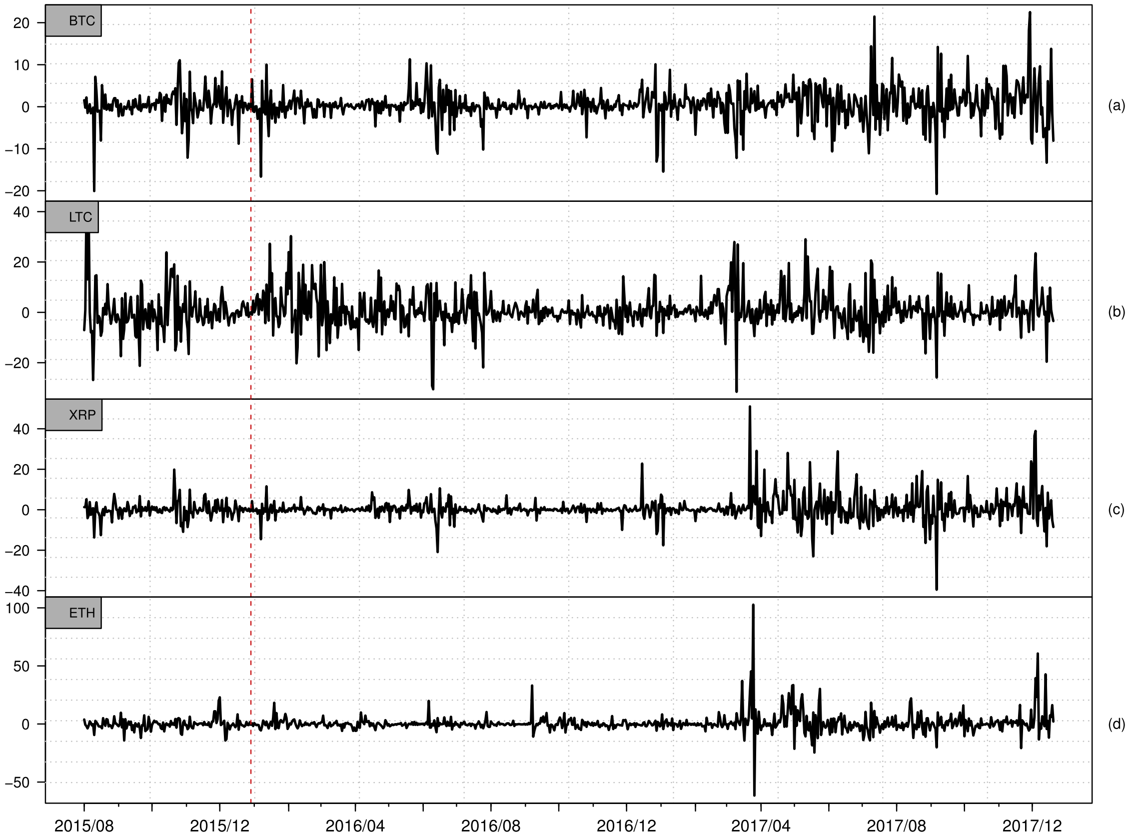


Fig. 1. The plots show the daily percentage log returns of the four cryptocurrencies considered in this study: Bitcoin (BTC) in panel (a), Litecoin (LTC) in panel (b), Ripple (XRP) in panel (c) and Ethereum (ETH) in panel (d). The dashed vertical red line indicates the beginning of the out-of-sample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

Table 2
Univariate models considered in the forecasting exercise.

Abbreviation	*Full description*
AR(1)–EWMA	Autoregressive model of order one with EWMA variance and $\kappa = 0.96$; benchmark model.
KS	Kitchen sink specification, i.e., a linear multiple regression that includes all variables.
KS–NR	Kitchen sink specification with only the lagged values of the series as covariates.
DMA	Dynamic model averaging across all model and forgetting factor combinations; see (Dangl & Halling, 2012).
DMS	Dynamic model selection, selecting the best model from among all model and forgetting factor combinations at each time t; 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.

Notes: The first column is the abbreviation of the model, while the second column provides a brief description of each individual model.

than 524,288 possible combinations.[4] To mitigate this issue, we propose to apply model combination techniques.

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

To provide more details 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 as a dynamic linear model (DLM), see West and Harrison (1999) and Raftery et al. (2010). In particular, let $\mathbf{F}_t^{(i)}$ denote a $p \times 1$ vector based on a given combination of our total predictors, $\mathbf{F}_t = \left(1, y_{t-1}, y_{t-2}, y_{t-3}, x_{1,t-1}, \ldots, x_{16,t-1}\right)'$. Then, we can express our ith DLM as:

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

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

Table 3
The multivariate models considered in the forecasting exercise plus the benchmark model.

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

Notes: The first column is the abbreviation of the model, while the second column provides a brief description of each model.

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 per Prado and West (2010). We employ the strategy of Prado and West (2010) and introduce an additional forgetting factor, $\kappa \in (0, 1)$, to induce time variation in V_t. The 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 by Dangl and Halling (2012). This additional forgetting factor, labelled 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 $t = 1, \ldots, T$, $\boldsymbol{\beta}_t^{(i)}$ is constant over time. Thus, Eq. (1) nests the specification of the constant regression coefficients. For $\mathbf{Q}_t^{(i)} \neq \mathbf{0}$, $\boldsymbol{\beta}_t^{(i)}$ varies according to Eq. (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 depends on the data at hand.

DMA then averages the 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 quickly the performances of past models determine the weights of the models. Common choices are $\alpha = 0.99$ and $\alpha = 1.0$; see equations (15) and (16) of Koop and Korobilis (2012). In addition to averaging, we can also use the model that receives the highest probability among all model combinations to forecast; 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 by 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 the predictive likelihood at each point.

3.2. Multivariate models

3.2.1. Constant parameter VARs

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

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

3.2.2. Time-varying parameter specifications

Cryptocurrencies are subject to several instabilities, both in the 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 by Koop and Korobilis (2013, henceforth KK), who provide a new approach to the estimation of large-dimensional TVP-VAR. Focusing on the case where the VAR has one lag and the intercepts are suppressed, the TVP-VAR(1) model can be written as:

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

with $\mathbf{y}_t$ being an $(M \times 1)$ vector containing observations on M time series at time t, T_t being 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, then rewrite the model in state space form:

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

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

KK propose that the computational burden be reduced by making 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 $\boldsymbol{\Sigma}_t$, and KK's idea 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). The latent state innovation variance $\mathbf{Q}_t$ typically 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 & 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 $\boldsymbol{\Sigma}_t$ with 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 was 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 $\boldsymbol{\Sigma}_0$); however, we refer to KK for an extensive discussion of the problem.

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

4. Empirical application

Our results are based on the $h = 1, \ldots, 7$-day-ahead forecasting process using an expanding window and an initial in-sample period of 146 days for Bitcoin, Litecoin, Ripple and Ethereum. Multi-step-ahead predictions are obtained through direct forecasting, see Marcellino, Stock, and Watson (2006). Hence, our forecast evaluation period is from January 1, 2016, to December 28, 2017. In each analysis we use the AR(1)–EWMA model as a benchmark for the univariate and multivariate models described in the previous section. We discuss forecast metrics in Section 4.1, univariate results in Section 4.2 and multivariate results in Section 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 performances of our forecasts using different point and density metrics. For the accuracy of point forecasts, we use the mean squared errors (MSEs) for each of the forecast horizons that we consider, $h = 1, \ldots, 7$.[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 forecast 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. We evaluate the density forecasts using the predictive log score (LS), which 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

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

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

More specifically, we report the MSEs and LSs for the benchmark model. For the other models, we report the ratio of each model's MSE to that of the baseline. Entries smaller than one indicate that the given model yields forecasts that are more accurate than those from the baseline and has scores that differ from those of the baseline, such that a positive number indicates a model that beats the baseline. We assess the differences among alternative models statistically by applying Diebold and Mariano (1995) t-tests for equality of the average loss (with the loss defined as the 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, Lunde, and Nason (2011) using the R package MCS detailed by Bernardi and Catania (2016) to compare all predictions jointly. The differences are tested separately for each horizon h.

Finally, we provide an economic evaluation of our forecasts by studying the directional predictability of cryptocurrency returns. As was pointed out by Christoffersen and Diebold (2006), for example, sign predictability may exist even in the absence of mean predictability, and complements 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 predicts the sign of future returns correctly.[9] A success rate of one means that a model predicts

6 Separate statistics for each currency are reported in Appendix B.

7 Appendix B also reports mean absolute deviations (MADs).

8 We also compute the Amisano and Giacomini (2007) test for differences in log score performances. Note 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 sample of data, rather than an expanding sample as in our case. The evidence is similar qualitatively.

9 We do not perform sign predictability tests because, as was indicated by Christoffersen and Diebold (2006), similar "tests that rely only on

Table 4
Mean squared error (MSE), computed over the forecast horizon.

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

Notes: The results reported are relative to the benchmark specification (AR(1)–EWMA), for which the absolute score is reported. The models are described 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 that the model belongs to the superior set of models delivered by the model confidence set procedure at a confidence level of 10%.

the correct sign for all forecasts, whereas a success rate of zero means that the model never predicts the correct sign. Similarly to the MSE evaluation, we report an average success rate for multivariate models across the four cryptocurrencies for each forecast horizon.

the sign sequence omit important information about volatility dynamics, which is potentially valuable for detecting sign predictability''. In regard to this background, we leave for future research market timing tests and portfolio strategies.

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 the daily horizon and Ethereum at several horizons. For Bitcoin, DMA gives statistically significant reductions

[10] See Table B.1 in Appendix B for MAD scores.

of 11% at one-day-ahead. Even when excluding crypto-predictors and applying combinations or selections of lags, the forecast accuracies of the strategies labelled DMA–NR and DMS–NR improve statistically. Considering the large volatility of the series and our focus on the daily forecast horizon, the gains are sizeable economically, especially 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 though DMA provides lower MSEs 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 outperforms the benchmark statistically at horizon four, but produces enormous losses at other horizons; see for example the statistics at the second horizon. BMA also outperforms the benchmark at horizon four, but is inferior at all other horizons. Therefore, methodologies other than DMA do not seem robust across horizons.

When focusing on Ethereum, we find evidence of statistically superior predictability of the alternative models relative 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 those of other currencies is a more valuable strategy for Ethereum than for Bitcoin, somewhat confirming its central role as a major exchanger of currencies in the crypto-system. Ethereum is closely connected to several currencies, 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 MSEs at one and four days ahead, but the differences are not significant statistically. KS and BMA improve the predictability at some horizons, but also perform very poorly at other horizons, 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 still indicates considerable uncertainty among predictors. Only DMA-NR outperforms the benchmark at two, five, six and seven days ahead. The lower capitalizations of Litecoin and Ripple relative to Bitcoin and Ethereum result in lower correlations with other assets and less accurate predictions. New predictors based on crypto-market sentiments might be considered for investigating the heterogeneity across cryptocurrencies.

Table 5 provides the log score results for the univariate models. The evidence differs from that for point forecasting: the gains vanish for the prediction of Bitcoin and Ethereum, with no models providing higher scores than the benchmark, and in the case of Ethereum, the AR model being the only specification that is included in the model confidence set at all horizons. Weaker evidence is found for Litecoin, but only at longer horizons with DMA-NR and DMS-NR. However, the 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 seems to fail to capture the higher moments dynamics of cryptocurrencies, even if it allows for time-varying volatility. Catania et al. (2018) find that sophisticated univariate models are required to produce accurate forecasts of the cryptocurrency volatility.

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.

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

Table 7 provides the average inclusion probabilities, while Fig. 2 shows how the weights of the corresponding five variables for the first horizon have evolved over time. 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 a tripling over time of the values of the variables Ethereum and Ripple high minus Ripple low, though with some drops. The interconnection with other markets, see the variables Nikkei 225 and Silver, also increases over our sample. However, the linkage to stock market risk, see VIX, decreases substantially from 2016Q2. Thus, our results suggest that Bitcoin has become more interconnected with other assets, but has also developed its own risk that is intrinsic to cryptocurrencies and seems idiosyncratic to other markets.

The evidence for the other cryptocurrencies is somewhat different: despite large changes in the middle of our sample, most of the coefficients assume similar values at the beginning and end of the sample. This is particularly evident for the coefficients of 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 of over 0.8, decreases to a minimum of almost zero, and by the end of the sample is back to a value of around 0.5. A slightly different pattern is seen for the coefficients of the other financial variables, with the coefficient of the Stock Europe 600 increasing and

Table 5
Log score (LS), computed over the forecast horizon.

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

Notes: The results reported are relative to the benchmark specification (AR(1)–EWMA), for which the absolute score is reported. The models are described 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 that the model belongs to the superior set of models delivered by the model confidence set procedure at a confidence level of 10%.

doubling to a final value of 0.4 for all three cryptocurrencies. Thus, Litecoin, Ripple and Ethereum all experienced a period of large changes in the second half of 2016 and the first half of 2017, but the relationships have stabilised and become more similar to Bitcoin since the second half of 2017, when the entire crypto-market experienced an impressive increase in value.

4.3. Multivariate forecasting results

Tables 8 and 9 report the MSEs and predictive log scores 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

Table 6
Success rate (SR), computed over the forecast horizon.

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

Note: The models are described in Table 2.

long horizons, though several models provide large gains when density forecasting cryptocurrencies. Focusing on the MSE results, the simpler constant-parameter VAR and BVAR specifications, labelled $\mathcal{M}_2$ and $\mathcal{M}_3$ respectively, are very imprecise at short horizons, with losses of up to 20%, but they perform more similarly to the AR(1)–EWMA at longer horizons, with slight improvements for $\mathcal{M}_2$ at long horizons, as was discussed above. Time-varying specifications provide more similar performances across horizons but 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 scores, several of the models in Table 9 provide forecasts that are statistically superior to that of the benchmark at almost all horizons. Models $\mathcal{M}_9$ (selection between a model with only cryptocurrencies and a model that also includes crypto-explicatives), $\mathcal{M}_{13}$ (selection between a model with only cryptocurrencies and a model that also includes macroeconomic variables), and $\mathcal{M}_{17}$ (selection among a model with only cryptocurrencies, a model that also includes crypto-explicatives, a model that also includes macro predictors and a model that includes 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 at all horizons, and is the only model in the confidence set at the two-day-ahead horizon. As in the univariate case, the crypto-explicatives and macro and financial predictors improve the forecast accuracy. However, in contrast to the univariate case, the selection of models that contain clusters of them provides the largest gains, rather than averaging predictors. We speculate that the importance of 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, Grassi, Ravazzolo, & van Dijk, 2018).

When focusing on the performance over time, Fig. 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 the TVP-VAR models provides constant gains relative to the AR(1)–EWMA benchmark over the entire out-of-sample period, with the increase being very large in some cases, such as at the end of March 2017, when all currencies experienced a break in volatility; see Fig. 1. The models based on DMA also outperform the

Table 7
Top five crypto-predictors for different cryptocurrencies and forecast horizons.

$h = 1$	$h = 2$	$h = 3$	$h = 4$	$h = 5$	$h = 6$	$h = 7$
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)

Note: The numbers in brackets are the average (%) inclusion probabilities of the selected predictors over the forecasting period.

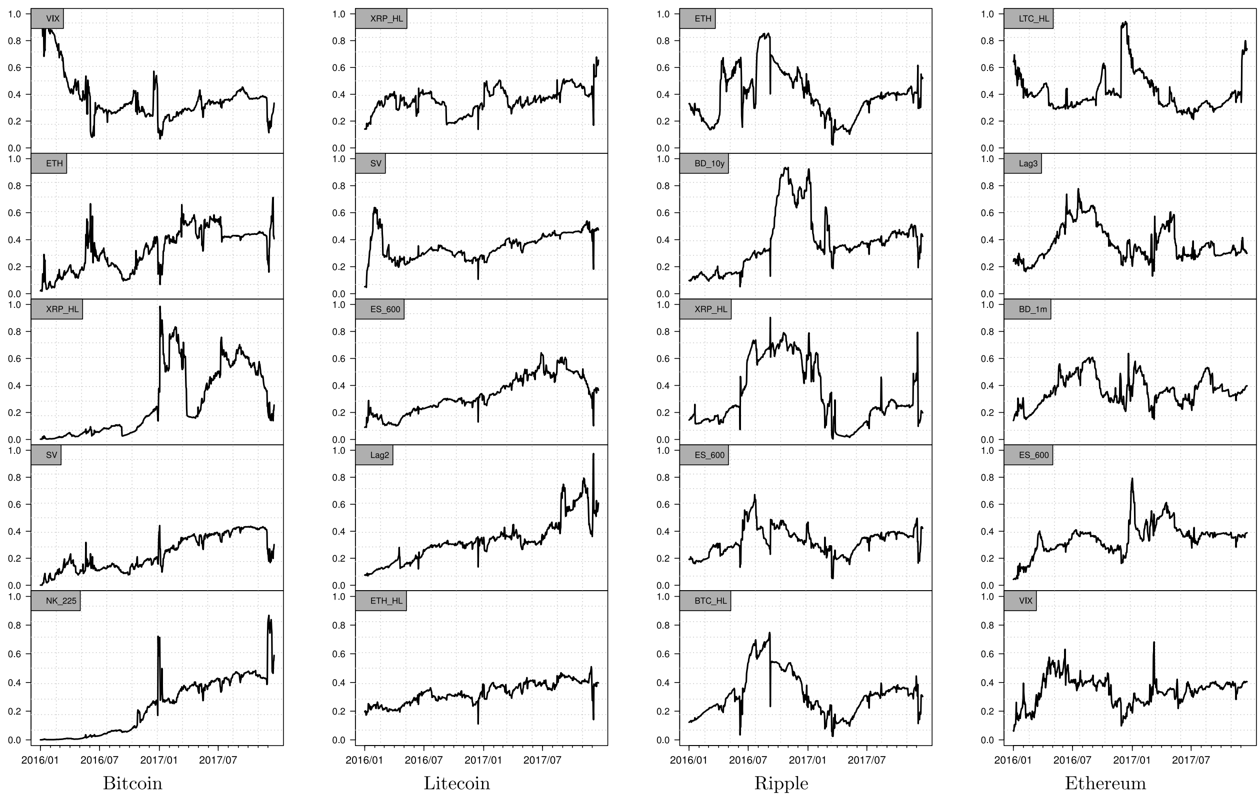


Fig. 2. Inclusion probabilities. The plots show inclusion probabilities of the top five predictors over time, computed over a one-day forecast horizon, for the four cryptocurrencies.

Table 8
(Multivariate) mean squared error, computed over the forecast horizon.

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

Notes: The results reported are relative to the benchmark specification (AR(1)–EWMA), for which the absolute score is reported. The models are described 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 that the model belongs to the superior set of models delivered by the model confidence set procedure at a confidence level of 10%.

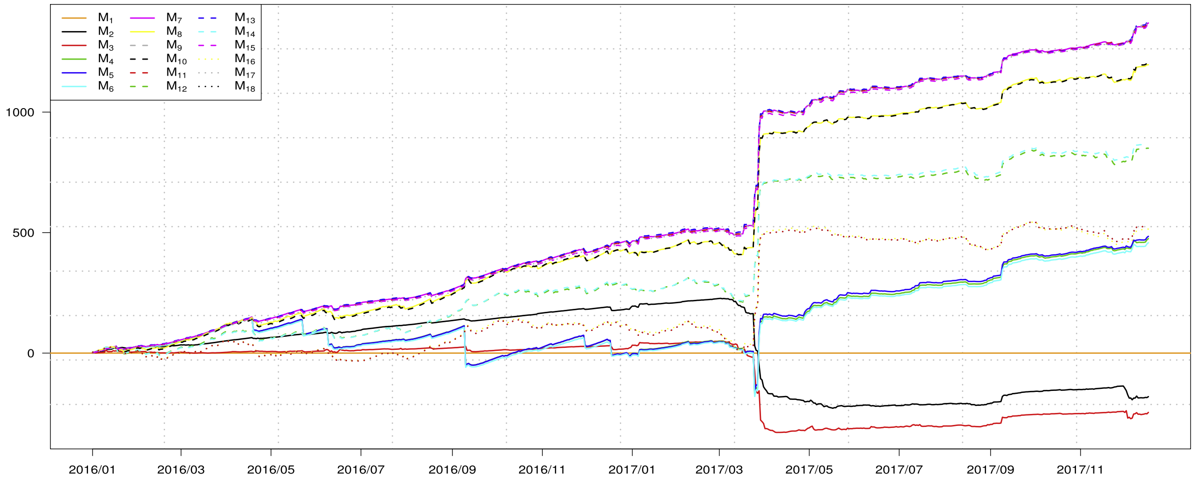
Table 9
(Multivariate) log score (LS), computed over the forecast horizon.

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

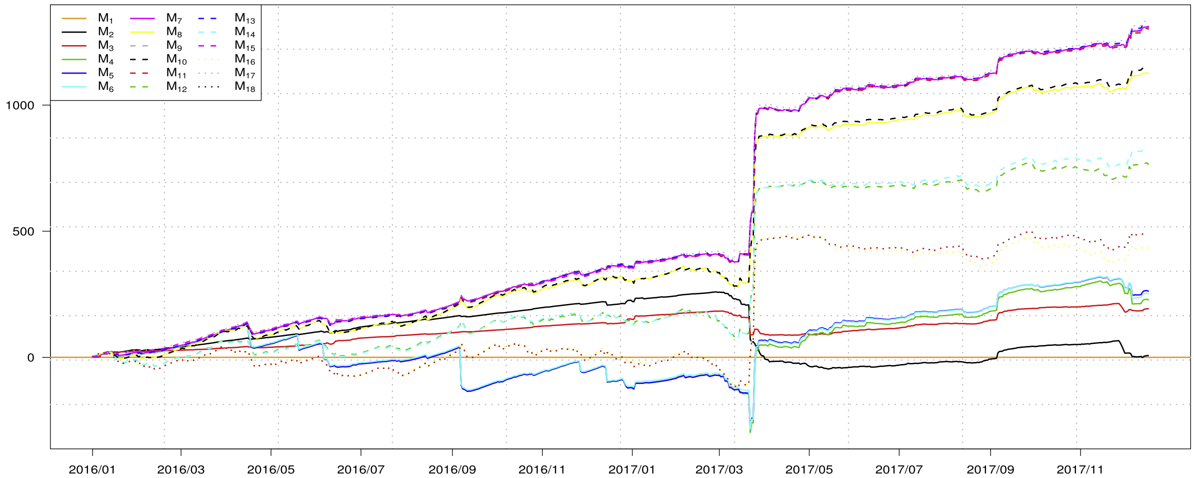
Notes: The results reported are relative to the benchmark specification (AR(1)–EWMA), for which the absolute score is reported. The models are described in Table 3. Values in **bold** indicate rejection of the null hypothesis of equal predictive ability between each model and the benchmark according to the Diebold–Mariano test at the 5% confidence level. Grey cells indicate that the model belongs to the superior set of models delivered by the model confidence set procedure at a confidence level of 10%.

benchmark, but the gains are smaller. VAR ($\mathcal{M}_2$) and BVAR ($\mathcal{M}_3$) provide some gains in the initial part of the sample, but their performances deteriorate substantially with the instability in March 2017. Only the BVAR at long horizons maintains part of its predictive power, but the scores are substantially lower than those of the DMS alternatives.

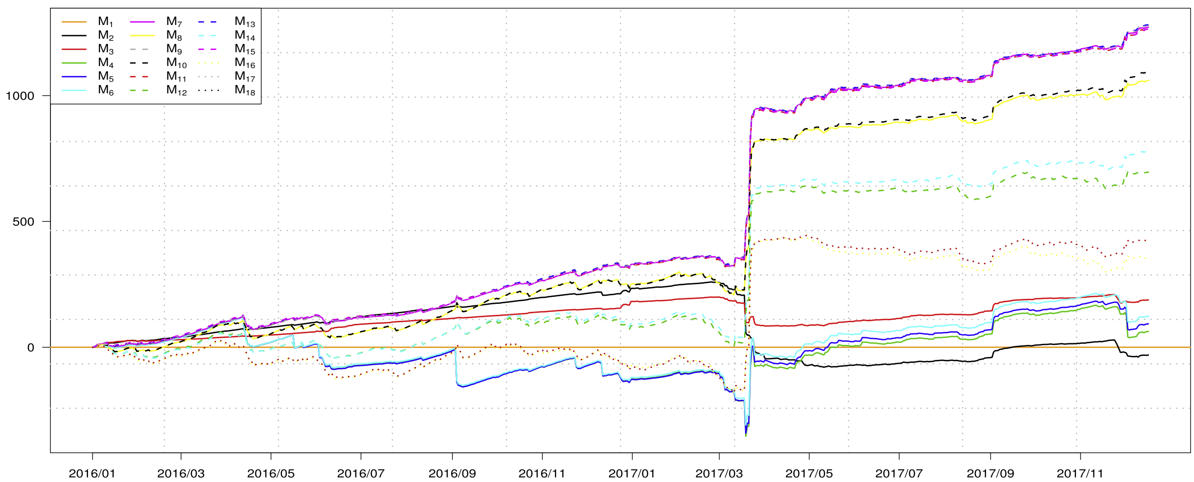
Tables B.3–B.6 in Appendix B report log score results for multivariate models when predicting each cryptocurrency



(a) $h = 1$



(b) $h = 4$



(c) $h = 7$

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

Table 10
(Multivariate) average success rate (SR), computed over the forecast horizon.

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

Note: The models are described in Table 3.

separately. The previous evidence is confirmed for each currency, and the gains are not a result of the models performing well just for a subset of assets.

Table 10 reports the directional predictability of the multivariate models. The success rate is over 50% in all cases, with higher numbers of almost 60% at the two- and three-day-ahead horizons. The results are very similar across models. When investigating the directional predictability for each currency individually, see Tables B.3–B.6, the highest gains are found for predicting Ripple, with success rates well above 60%. On the other hand, Ethereum seems the most difficult to predict in terms of sign movements. Considering the large gains that we find for log score evaluation, more advanced investment strategies with more weight on higher moments rather than only on the return direction could be investigated.

5. Robustness: forgetting factor comparison

Our combination schemes in Eqs. (1) and (3) depend on the selection of three parameters: (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 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. The next two subsections report the results for different values of α and κ in the univariate and multivariate schemes.[11]

5.1. Robustness: univariate models

Tables C.1–C.9 in Appendix C report the MSEs, log scores and success rates 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 period, 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 each for $\alpha = \{0.01, 0.99, 1.00\}$ and $\kappa = \{0.96, 0.98, 1.00\}$.

The results are qualitatively similar to the main case for small variations of α and κ. In contrast, the results deteriorate substantially when $\alpha = 0.01$, indicating that inducing persistence in the selection of the models improves the forecasts. Interestingly, the MSEs and success rates improve in several cases for $\alpha = 1.00$, but the log score decreases for the same setting.

5.2. Robustness: multivariate models

Tables C.4–C.9 in Appendix C report the MSEs, log scores and success rates 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 period, and for the other specifications relative to them. For the success rate, the full sample statistics for all models are given, and we consider $\kappa = 0.96$, $\kappa = 1.00$.

The results are qualitatively similar to those in Tables 8–10, even if marginally inferior in several cases, and therefore fixing $\kappa = 0.98$ seems the better strategy. We also try to vary $\alpha = \{0.01, 1.00\}$ but find the results to deteriorate, especially 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 the abilities of several alternative univariate and multivariate models to predict four of the most capitalised cryptocurrencies: Bitcoin, Litecoin, Ripple, and Ethereum. A set of crypto-predictors is applied and both univariate and multivariate model combinations are proposed for combining these predictors. The results show large, statistically significant improvements in the point forecasting of Bitcoin and Ethereum when using combinations of univariate models, and in density forecasting for all four 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 the prediction of cryptocurrencies. For example,

[11] See McCormick, Raftery, Madigan, and Burd (2012) and Bork and Møller (2015) for an extension to recursive estimation of the forgetting factors.

flexible multivariate combination schemes that allow for different weights across series could improve the point and density forecast accuracy. Moreover, new predictors based on crypto-market sentiments might be considered for investigating heterogeneity across cryptocurrencies, and could result in (point) forecast gains across all series.

Acknowledgments

The authors would like to thank the editor, the associate editor, two anonymous referees, and seminar and conference participants at the 2018 European Seminar on Bayesian Econometrics in New Orleans, the 2018 Mathematical and Statistical Methods for Actuarial Sciences and Finance (MAF 2018) in Madrid, R/Finance 2018 in Chicago for valuable comments. The usual disclaimers apply.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.ijforecast.2018.09.005.

References

Amisano, G., & Giacomini, R. (2007). Comparing density forecasts via weighted likelihood ratio tests. *Journal of Business & 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., & Catania, L. (2016). Portfolio optimisation under flexible dynamic dependence modelling. ArXiv e-prints.

Bianchi, D. (2018). *Cryptocurrencies as an asset class? An empirical assessment*: *Tech. rep., 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., & Møller, S. V. (2015). Forecasting house prices in the 50 states using dynamic model averaging and dynamic model selection. *International Journal of Forecasting*, *31*, 63–78.

Casarin, R., Grassi, S., Ravazzolo, F., & van Dijk, H. K. (2018). *Predictive density combinations with dynamic learning for large data sets in economics and finance*: *Tech. rep., Tinbergen institute discussion paper 15-084/III*.

Catania, L., & Grassi, S. (2018). *Modelling Crypto–currencies financial time–series*: *Tech. rep., CEIS working paper no. 417*.

Catania, L., Grassi, S., & Ravazzolo, F. (2018). In M. Corazza, M. Durbán, A. Grané, C. Perna, & M. Sibillo (Eds.), *Predicting the volatility of cryptocurrency time-series*. In *Mathematical and statistical methods for actuarial sciences and finance*, Charm: Springer.

Catania, L., & Nonejad, N. (2018). Dynamic model averaging for practitioners in economics and finance: The eDMA package. *Journal of Statistical Software*, *84*(11).

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., & Diebold, F. (2006). Financial asset returns, direction-of-change forecasting, and volatility dynamics. *Management Science*, *52*, 1273–1287.

Chu, J., Nadarajah, S., & 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., & Halling, M. (2012). Predictive regressions with time–varying coefficients. *Journal of Financial Economics*, *106*, 157–181.

Diebold, F., & Mariano, R. (1995). Comparing predictive accuracy. *Journal of Business and Economic Stastistics*, *13*, 253–263.

Durbin, J., & Koopman, S. (2012). *Time series analysis by state space methods* (2nd ed.). Oxford, UK: Oxford University Press.

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., & Amisano, G. (2010). Comparing and evaluating bayesian predictive distributions of asset returns. *International Journal of Forecasting*, *26*, 216–230.

Hansen, P. R., Lunde, A., & Nason, J. M. (2011). The model confidence set. *Econometrica*, *79*, 453–497.

Hencic, A., & Gourieroux, C. (2014). Noncausal autoregressive model in application to bitcoin/USD exchange rate. In *Proceedings of the 7th financial risks international forum* (pp. 1–25).

Hotz-Behofsits, C., Huber, F., & Zorner, T. O. (2018). Predicting cryptocurrencies using sparse non–gaussian state space models. *Journal of Forecast*, *36*(6), 627–640.

Johannes, M., Korteweg, A., & 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., & Korobilis, D. (2010). Forecasting inflation using dynamic model averaging. *Bayesian Multivariate Time Series Methods for Empirical Macroeconomics*, *3*, 267–358.

Koop, G., & 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., & Korobilis, D. (2012). Forecasting inflation using dynamic model averaging. *International Economic Review*, *53*, 867–886.

Koop, G., & 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., & 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., & Burd, R. S. (2012). Dynamic logistic regression and dynamic model averaging for binary classification. *Bioemtrics*, *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., & Stambaugh, R. F. (2000). Comparing asset pricing models: an investment perspective. *Journal of Financial Economics*, *56*, 335–381.

Prado, R., & West, M. (2010). Time series: Modeling, computation, and inference. Boca Raton: CRC Press.

Raftery, A. E., Kárný, M., & 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., & Kokkinaki, A. (2014). Bitcoin is volatile! Isn't that right? *Business Information Systems Workshops*, 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., & Harrison, J. (1999). Bayesian forecasting & dynamic models. Berlin: Springer–Verlag.

Leopoldo Catania is an Assistant Professor in the Department of Economics and Business Economics at Aarhus University. Leopoldo's research interests focus on Financial Econometrics and Time Series Analysis. His works concern the development and estimation of univariate and multivariate econometrics models applied to quantitative risk management, density predictions of financial returns, time-varying dependence, and volatility. He has published articles in leading journals such as Journal of Applied Econometrics; Journal of Statistical Software; and the European Journal of Finance.

Stefano Grassi is a senior Assistant Professor (RTD-B) in the Department of Economics and Finance at University of Rome 'Tor Vergata'. His research interests is time series analysis, state space models, Bayesian analysis, financial econometrics, empirical finance, sequential Monte Carlo methods, spectral analysis, computational economics and DSGE models. He has published articles in leading journals, such as International Journal of Forecasting; Journal of Applied Econometrics; Journal of Statistical Software; and Journal of Empirical Finance.

Francesco Ravazzolo is a full professor in the Faculty of Economics and Management at Free University of Bozen-Bolzano. His research focuses on modelling and (density) forecasting of financial and macroeconomic time series, Bayesian inference, sequential Monte Carlo methods, VAR models. He is member of the editorial board of several international academic journals and international societies, including *International Journal of Forecasting, Journal of Applied Econometrics and Annals of Applied Statistics*. He has published articles in several leading journals, including *International Journal of Forecasting; Journal of American Statistical Association; Journal of Applied Econometrics; Journal of Business and Economic Statistics; Journal of Econometrics; and Economic Journal*.