Università degli Studi Firenze
DISEI
Dipartimento di Scienze per l'Economia e l'Impresa

Working Papers - Economics

A State-Space Approach for Time-Series Prediction of an Heterogeneous Agent Model

Filippo Gusella, Giorgio Ricchiuti

Working Paper N. 20/2022

DISEI, Università degli Studi di Firenze
Via delle Pandette 9, 50127 Firenze (Italia) www.disei.unifi.it

A State-Space Approach for Time-Series Prediction of an Heterogeneous Agent Model

Filippo Gusella[a,b] Giorgio Ricchiuti[a,b]

[a] *Università degli Studi di Firenze*

[b] *Complexity Lab in Economics (CLE), Università Cattolica del Sacro Cuore, Milano*

Abstract

In this paper we apply the state-space model approach to evaluate and compare the forecasting performance of a small-scale heterogeneous agent model (HAM) with fundamentalists and contrarians. As in the tradition of HAMs, agents are heterogeneous in the expectations formation and forecast future prices based on the deviations of previous values with respect to the fundamental value. Moreover, our agents have two specifications for the asset's fundamental value, formalized as a random walk (RW) or with the Gordon model (GM). We examine the models' performance at various forecast horizons (short vs. long horizon) and different frequency-time (monthly and quarterly). Overall, GM statistically outperforms RW specification at the long horizon with statistical significance, while RW and GM are statistically indifferent in the short horizon.

Keywords: Heterogeneous expectations, forecasting, RW, Gordon model, state-space model

JEL codes: C13, C50, G10, G12, G15, E32

Corresponding author: Giorgio Ricchiuti, Via delle Pandette, 32 - 50127 Firenze, giorgio.ricchiuti@unifi.it

Acknowledgment: An earlier version of this paper was presented at the 6th International Workshop on "Financial Markets and Nonlinear Dynamics" (FMND), Paris; the 4th Behavioral Macroeconomics Workshop, Bamberg; and the 25th WEHIA Conference, Catania. We are grateful to the participants for their helpful comments and suggestions which led to a considerable improvement of the paper. We also thank Fabrizio Cipollini and Gabriele Fiorentini for useful comments, which allowed us to improve the quality of the paper, and Yuri Basile, for suggestions on data. The usual disclaimer applies. Filippo Gusella is grateful for financial support from the Università di Firenze and Complexity Lab in Economics (CLE), Università Cattolica del Sacro Cuore, under the joint project "Modelli economico finanziari con agenti eterogenei e learning: approccio teorico e empirico".

1 Introduction

From a theoretical point of view the heterogeneous agent models (HAMs) are mature. The hypothesis that groups of agents with different strategies, expectations, and beliefs interact in financial markets following heuristic decisions is accepted in the literature (Hommes 2021). In the last two decades, scholars have devoted much effort to reconciling empirical and theoretical analysis, seeking evidence of intrinsic and pervasive heterogeneity in the data. Indeed, there is a growing empirical literature on HAMs, even if authors use different estimation procedures, showing no convergence on the methodology to be adopted (Lux and Zwinkels 2018; Ter Ellen and Verschoor 2018).

Following Franke and Westerhoff (2017), two approaches - direct and indirect - may be recognized. The first method employs surveys to measure the heuristic behavior of traders while the second method considers a model as a whole and strives to estimate its parameters with various inference methods (Kukacka and Barunik 2017). On one side, key structural features of HAMs can be estimated straightaway (Wan and Kao, 2009; Chiarella et al. 2012; Goldbaum and Zwinkels 2014; Hommes and in't Veld 2017; Ter Ellen et al. 2021; Lux 2021) through liner/nonlinear least square (Boswijk et al. 2007; Lof 2015) and maximum/quasi-maximum likelihood (De Jong et al. 2010; Chiarella et al. 2014). On the other side, through the method of simulated moments (Franke and Westerhoff 2011; Grazzini and Richiardi 2015; Chen and Lux 2018), it is possible to depict patterns which come from behavioral biases, such as volatility clustering and long-memory effects (Schmitt 2021)[1].

Recently, starting from Lux (2018, 2021b), some authors have used the state-space approach. As Lux pointed out, this methodology could be optimal for HAMs because it distinguishes between latent components (behavioral heuristics) and observable variables (asset prices). In this line, Gusella and Stockhammer (2021) propose an empirical test for Minskyan financial cycles in equity and housing prices. Gusella and Ricchiuti (2022), within a HAM's framework, estimate the presence of endogenous cycles that emerge as a consequence of opposite forces reflected in the agents' behavior. The HAM model boils down to a state-space system of two first-order difference equations, whose coefficients are estimated with a Kalman filter. The fundamental value employed follows a random walk (RW), or, as a second specification, it is identified through the Gordon model (1959).

Based on this background, three main reasons motivate this paper: despite its importance, forecasting procedures are still little explored in this literature; moreover, state-space models have not been used to study the forecasting performance of heterogeneous agent models. Finally, within HAM literature, the fundamental value follows either a RW, as in Franke (2008), or it is constructed from the Gordon growth model as in Chiarella et al. (2012). However, the forecasting power of the two specifications has never been compared: we believe that a comparison between the forecasting power of RW and Gordon approach is needed.

[1]In addition to these line of researches, asset pricing laboratory experiments is becoming increasingly popular. Controlled lab experiments confirm how the single representative rational hypothesis can be substituted with the idea that heterogeneous agents interact in the market following simple rules of thumb (Hommes et al. 2005; 2008; Bao et al. 2020; 2021).

To the best of our knowledge, empirical works within HAM tradition have focused more on finding heterogeneity in data, while few works, especially for stock market prices, have devoted attention to the possibility of extrapolating forecasting power through the models used. We may divide between forecasting analysis of stock prices and foreign exchange rates. Among the latter, we can mention Manzan and Westerhoof (2007), De Jong et al. (2010) and Jongen et al. (2012) for exchange rates market; while Chiarella et al. (2012), Lof (2012), Recchioni et al. (2015) are focused on the stock market and finally Ghonghadze and Lux (2016) forecast asset price volatility for both stock and foreign exchange markets as well as the price of gold.

Following, we briefly report the main elements of these papers. Chiarella et al. (2012) use US stock market prices from January 2000 to June 2010 to compare the forecasting accuracy of a heterogeneous behavioral model under Markov regime-switching with that developed in Boswijk et al. (2007), estimated with the nonlinear least-squares technique. While the log-likelihood and Akaike's information criterion over the sample period and in terms of in-sample estimation suggest that the models share similar explanatory power, the regime-dependent beliefs model possesses better forecasting accuracy than the nonlinear model for forecasting horizon of 1 up to 12 months. Lof (2012) considers the classical chartists/fundamentalists division to explain the price-dividends and price-earnings ratios of S&P 500. He applies the smooth-transition univariate and multivariate autoregressive (STAR) models and compares their forecasting power with a linear autoregressive model. The mean absolute errors and root mean square errors suggest that the model with a multivariate transition function outperforms both the univariate case and the classical linear alternative. Recchioni et al. (2015) compare the one and two days ahead forecast values of S&P 500, the Nikkey 225, CSI 300, and the Euro Stoxx 50 indices using the Brock and Hommes (1998) model (BH), an adjusted calibrated version of Brock and Hommes model (ABH) and the RW model (RW). With one day ahead forecast, the ABH matches the observed values better than the RW and the BH model for the S&P 500 and the Nikkey 225. In the case of the two days ahead forecasts, both the BH and ABH specifications are able to track the trend of the observed index but with weaker performance than the RW specification.

Moving to exchange rates markets, HAMs works are inspired by the seminal paper of Meese and Rogoff (1983), who show how, in out-of-sample prediction, structural models perform worse than the RW specification in terms of root mean squared forecast errors[2]. De Jong et al. (2010), starting from the approach proposed by De Grauwe and Grimaldi (2005; 2006), show that for seven monthly bilateral exchange rates from March 1979 through December 1998, the heterogeneous agent model with a switching mechanism outperforms the static version and the RW benchmark with better performance for the longer horizon. Jongen et al. (2012), using survey-based exchange rate forecasts for the period 1989:10-1995:2 and 1995:11-2009:12, study the HAM forecast ability in out-of-sample period compared to the RW model. The obtained results indicate better forecasting results for HAM than a RW, but the same forecasting

[2]See for example Kouwenberg et al. (2017) and Cheung et al. (2019) on this point.

power between static and the switching model. Manzan and Westerhoff (2007) detect the HAM's predictability against the RW alternative, finding short-term unpredictability and long-term predictability for major currencies vis a vis the US dollar for the period 1974-1998.

Finally, in a macro setting, an interesting attempt to produce forecasting exercises for an agent-based model (ABM) is the paper by Delli Gatti and Grazzini (2020). They propose a simulation-based forecasting technique using medium-scale macroeconomic ABM to forecast US data and compare its forecasting power with a VAR model. Using the Diebold Mariano test, they find that the forecasts produced by the VAR approach are more accurate than the forecasts generated by the simulation-based methodology for ABM. However, the ABM is able to forecast macro variables with promising results so as to open the research for future analysis.

Given these elements, in this paper, using state-space approach, we evaluate the forecasting power of an HAM to make a step forward by studying how these two specifications (RW vs. Gordon) compete in forecasting market price dynamics and fit better future output.

To perform it, our approach proceeds in different steps (see Fig. 1). As in the tradition of HAMs, we firstly formalize a structural theoretical model with two heterogeneous classes of agents, fundamentalists and contrarians. Because of the different definitions for the fundamental price, two structural theoretical specifications can be formalized. At the same time, considering the agents' beliefs as unobserved components, the two specifications can be translated into two state-space reduced-form models. In performing this transformation, we preserve the essential causality of HAMs. Indeed, the outlined method allows us to model the evolution of observable variables (asset price) driven by unobservable forces (heuristics). Finally, the state-space model can be used to forecast future observable variables and compare the out-of-sample predictability.

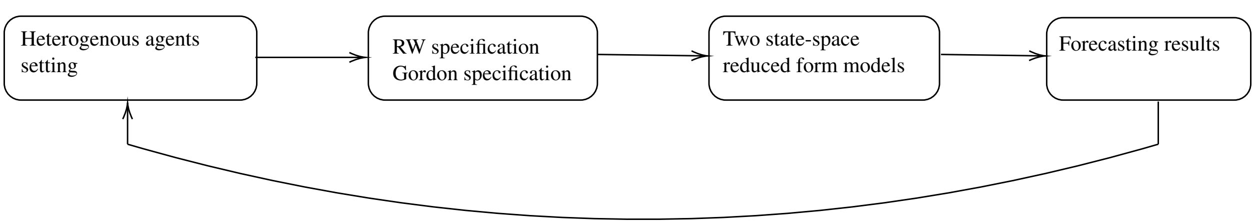

Figure 1: From the heterogeneous agents setting to the state-space model.

For the empirical analysis, we consider S&P500 data from 1990 to 2020 for monthly and quarterly data at different investment horizons: $\{h = 1, 2, ...24\}$ for monthly data and $\{h = 1, 2, ...8\}$ for quarterly data.[3] Finally, once the

[3] About the important connection between heterogeneity in time horizons and heterogeneity in expectations, see Anufriev and Bottazzi (2012).

forecasting procedure has been implemented, the root mean square errors (RMSE) and the mean absolute error (MAE) are calculated to compare the different forecasting results.

This paper will be structured as follows. Section 2 reviews the state-space approach in the context of heterogeneous agent models and formulates the forecasting procedure. Section 3 presents the forecasting results. Finally, section 4 concludes with suggestions on possible future developments.

2 Methodology

Before comparing the predictive power of our specifications, following Gusella and Ricchiuti (2022), we briefly review our state-space model approach for HAMs. Later, the forecasting procedure through the state-space model is examined.

2.1 The model

The market is populated by two groups of investors, fundamentalists (f) and contrarians (c), who predict future prices by extrapolating past price deviation from its fundamental value.

P_t, the asset price at time t, depends on the fundamental price, P_t^f, and agents' excess of demand, which is a weighted average of agents' beliefs:

$$P_t = P_t^f + \gamma B_t^f \left(P_{t-1}^f, P_{t-1}\right) + (1-\gamma) B_t^c \left(P_{t-1}^f, P_{t-1}\right) \tag{1}$$

where $B_t^f \left(P_{t-1}^f, P_{t-1}\right)$ and $B_t^c \left(P_{t-1}^f, P_{t-1}\right)$ are the functions which describe how fundamentalists and chartists form their beliefs, and γ with $1-\gamma$ are, respectively, the shares of the two groups.[4]

In general, the belief function, B_t^i (with $i = c, f$), can be specified as the difference between agents' expected price and the base price P^b, which is the price that enters in the formalization of their heuristic expectations[5]:

$$B_t^i = E_t^i \left(P_{t+1}\right) - P_{t-1}^b$$

The base price can be the fundamental value $\left(P_{t-1}^f\right)$ or the current asset price (P_{t-1}) at time $t-1$. As a consequence, the structure of price expectations can have two different forms:

$$E_t^i \left(P_{t+1}\right) = P_{t-1} + B^i \left(P_{t-1}^f, P_{t-1}\right), \qquad with \quad P_{t-1}^b = P_{t-1} \tag{2}$$

or

[4]Henceforth chartists and contrarians will be used as synonyms.

[5]It is worth noting, we do not impose any specific base price. This assumption is motivated by the fact that expectations of the agents are unobserved variables.

$$E_t^i\left(P_{t+1}\right) = P_{t-1}^f + B^i\left(P_{t-1}^f, P_{t-1}\right), \qquad with \quad P_{t-1}^b = P_{t-1}^f \tag{3}$$

The heterogeneous expectations influence the belief functions of the agents; fundamentalists believe that asset price will revert to the fundamental value:

$$E_t^f\left(P_{t+1}\right) = \tilde{\alpha} P_{t-1}^f + \left(1 - \tilde{\alpha}\right) P_{t-1} \qquad \left(0 \leq \tilde{\alpha} \leq 1\right) \tag{4}$$

where $\tilde{\alpha}$ measures the speed of mean-reversion of the market price to the fundamental value.

With respect to the belief function B^f and referring to (2) and (3), the expectation formation can be rewritten as:

$$B_t^f = \alpha\left(P_{t-1}^f - P_{t-1}\right) \tag{5}$$

If the base price is equal to the current asset price at a previous time ($P_{t-1}^b = P_{t-1}$), with respect to (4) we have $\alpha = \tilde{\alpha}$ with $0 \leq \alpha \leq 1$. If the base price is equal to the fundamental price at a previous time ($P_{t-1}^b = P_{t-1}^f$), with respect to (4), we have $\alpha = \tilde{\alpha} - 1$ with $-1 \leq \alpha \leq 0$.

While chartists speculate on the persistence deviation of asset price from the fundamental value, acting as contrarians:

$$E_t^c\left(P_{t+1}\right) = \tilde{\beta} P_{t-1}^f + \left(1 - \tilde{\beta}\right) P_{t-1} \qquad \left(\tilde{\beta} \leq 0\right) \tag{6}$$

Given B^c and referring to (2) and (3), respectively, the expectation formation can be rewritten as:

$$B_t^c = \beta\left(P_{t-1}^f - P_{t-1}\right) \tag{7}$$

As before, if the base price is equal to the observed asset price at time before ($P_{t-1}^b = P_{t-1}$), with respect to (6) we have that $\beta = \tilde{\beta}$ with $\beta \leq 0$. If the base price is equal to the fundamental price at time before ($P_{t-1}^b = P_{t-1}^f$), with respect to (6), we have that $\beta = \tilde{\beta} - 1$ with $\beta \leq -1$.

In the HAM literature, P_t^f follows a RW process or it is constructed through the Gordon model. In the first case, we have:

$$P_t{}^f = P_{t-1}^f + \varepsilon_t \qquad \varepsilon_t \sim \mathrm{N}\left(0, \sigma_\varepsilon^2\right) \tag{8}$$

where ε_t is the individual disturbance term that drives the evolution of the state vector.

In the second case, defining g as the average growth rate of dividends, r the average required return, and d_t the dividend flow, the fundamental value of asset price can be defined as:

$$P_t^f = u_t = d_t \frac{(1+g)}{(r-g)}$$

Following Chiarella et al. (2012), we assume that r is equal to the sum of the average dividend yield $\tilde{y}$ and the average rate of capital gain $\tilde{x}$. The Gordon growth model implies that x is equal to g, so as to obtain:

$$u_t = d_t \frac{(1+g)}{\tilde{y}} \tag{9}$$

Since the agents' expectations are unobserved variables that influence the observed price dynamics over time, the estimation procedure can be implemented through a state-space model. To construct it, we substitute (1) in (5) and (7). When the fundamental value follows Eq. (8), the stochastic state-space form in a matrix-vector formulation assumes the following form:

$$P_t = \begin{pmatrix} 1 & \gamma & 1-\gamma \end{pmatrix} \begin{pmatrix} {P_t}^f \\ B_t^f \\ B_t^c \end{pmatrix}$$

with the following transition equation:

$$\begin{pmatrix} B_t^f \\ B_t^c \\ {P_t}^f \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & 0 \\ a_{21} & a_{22} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} B_{t-1}^f \\ B_{t-1}^c \\ P_{t-1}^f \end{pmatrix} + \begin{pmatrix} \varphi_t \\ \eta_t \\ \varepsilon_t \end{pmatrix}$$

φ_t and η_t are the noise processes normally distributed with mean zero and variance σ_φ^2 and σ_η^2 respectively.

As the opposite, when the fundamental value is formalized through the Gordon model (${P_t}^f = u_t$), the observed equation assumes the following form:

$$P_t = u_t + \begin{pmatrix} \gamma & 1-\gamma \end{pmatrix} \begin{pmatrix} B_t^f \\ B_t^c \end{pmatrix}$$

with the transition equation equal to:

$$\begin{pmatrix} B_t^f \\ B_t^c \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} B_t^f \\ B_t^c \end{pmatrix} + \begin{pmatrix} \varphi_t \\ \eta_t \end{pmatrix}$$

The system's dynamics is given by the transition equation, which describes the evolution of the vector of unknown latent variables. In this way, differently from what is predicted by the efficient market theory in a representative rational agent framework, the observed price movements are not the consequence of exogenous shock to the system but are endogenously determined in the market.

2.2 Forecasting procedure

Once the model is in a state-space form, the unobserved beliefs of the agents together with the proportion of heterogeneous traders can be extracted by maximum likelihood via the Kalman filter over the state-space model. This procedure involves the prediction error decomposition approach with three steps: 1) initialization, 2) prediction and 3) correction (see Appendix A). The goal is to minimize the mean square prediction error of the unobserved state vector conditional on the observation of P_t (Harvey 1989).

Concerning the forecasting procedure, the prediction error decomposition approach is repeated for all the out-of-sample periods in the forecasting exercise (Harvey 2006). We divide our sample (1990-2020) into two segments; the in-sample estimation and the out-of-sample forecast. The out-of-sample period coincides with the last two years (2019-2020) so as to ensure many forecast observations to conduct inference. To perform the forecast, after running the Kalman filter up to time T , we have the starting values for the forecast of the observed series from the current estimate of the state vector. In particular, we use the final state predictor $Z_{t\,|\,t}$ together with the measurement and transition equation to construct $P_{T+h|T}$ where $h = 1, 2, ...24$ (for monthly data) and $h = 1, 2, ...8$ (for quarterly data) (Durbin and Koopman 2012). At the same time, since the sample separation is arbitrary, we perform a state-space rolling window estimation forecasting analysis with a fixed sample size to forecast the h period ahead. In this way, we aim to mimic the behavior of financial traders who update their forecasting when new information arrives in each period (see Fig. 2).

3 Results

In this section we evaluate the forecast power to understand which of the two model specifications fit better future output. First, we test the robustness of our analysis by looking at each specification through the Mincer and Zarnowitz test (Mincer and Zarnowitz, 1969). In this way, we can evaluate a possible systematic bias in the historical forecasts (see Appendix B). Second, we compare the two forecasting performances by looking at the ratios of root mean square error (RMSE) and mean absolute error (MAE) for different forecasting horizons. Finally, the Diebold-Mariano test is implemented to determine whether the two specification forecasts are significantly different (Diebold and Mariano, 1995)[6]. In order to have a deeper picture of the effect of heuristic beliefs, the analysis has been repeated firstly for monthly and then for quarterly data.

[6] see Appendix C.

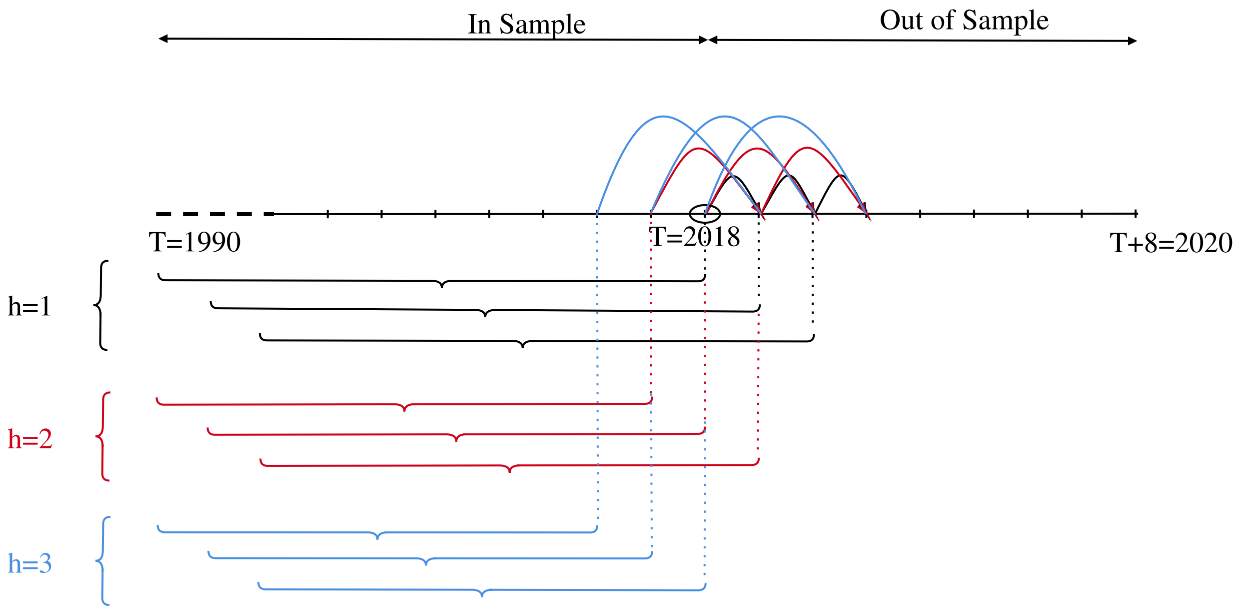

Figure 2: State-space rolling-window forecasting procedure for quarterly data.

Monthly Results. Following the Mincer-Zarnowitz test, to have unbiased forecasts, the regression of actual value on the ex-ante forecast should have a zero intercept and a coefficient of one. Tables 1 and 2 illustrate the obtained results for Gordon and RW, respectively. Overall, for both specifications, coefficients are different from zero at one percent level and values are very close to one, indicating systematic unbiased in the historical forecasts.

Table 1: Mincer-Zarnowitz regression test (Gordon model - monthly frequency)

h	1 m. h.	2 m. h.	3 m. h.	4 m. h.	5 m. h.	6 m. h.	7 m. h.	8 m. h.
	1.0019***	1.0031***	1.0037***	1.004***	1.0047***	1.0048***	1.0052***	1.0053 ***
	(0.00154)	(0.00222)	(0.00237)	(0.00264)	(0.00291)	(0.00294)	(0.00286)	(0.00300)
h	9 m. h.	10 m. h.	11 m. h.	12 m. h.	13 m. h.	14 m. h.	15 m. h.	16 m. h.
	1.0054***	1.0047 ***	1.0044***	1.0046***	1.005***	1.0054***	1.0058***	1.0063***
	(0.00303)	(0.00231)	(0.00214)	(0.00222)	(0.00216)	(0.00201)	(0.00192)	(0.00206)
h	17 m. h.	18 m. h.	19 m. h.	20 m. h.	21 m. h.	22 m. h.	23 m. h.	24 m. h.
	1.0067***	1.0072***	1.0078***	1.0079***	1.0084***	1.0088***	1.0097***	1.0096
	(0.00212)	(0.00252)	(0.00274)	(0.00262)	(0.00253)	(0.00242)	(0.00233)	(0.00227)

Notes: The table presents the results of the Mincer-Zarnowitz regression test. The columns represent the results for the Gordon model. h represents the forecast month horizons with standard error in parentheses. $*$, $**$, $***$ denote statistical significance at the 10%, 5%, and 1% levels respectively.

Once we have the forecast results, we want to analyze the predictive accuracy. Our approach is to select the forecast with the smaller error measurement based on the forecasting errors. Specifically, we compare the forecast-

Table 2: Mincer-Zarnowitz regression test (RW - monthly frequency)

h	1 m. h.	2 m. h.	3 m. h.	4 m. h.	5 m. h.	6 m. h.	7 m. h.	8 m. h.
	1.0021***	1.0035***	1.0044***	1.0052***	1.0062***	1.007***	1.0077***	1.0083***
	(0.00145)	(0.00206)	(0.00214)	(0.00234)	(0.00247)	(0.00243)	(0.00229)	(0.00247)

h	9 m. h.	10 m. h.	11 m. h.	12 m. h.	13 m. h.	14 m. h.	15 m. h.	16 m. h.
	1.0088***	1.0087***	1.0092***	1.0099***	1.0109***	1.0118***	1.0127***	1.01353***
	(0.00242)	(0.00185)	(0.00186)	(0.00196)	(0.00186)	(0.00168)	(0.00153)	(0.00164)

h	17 m. h.	18 m. h.	19 m. h.	20 m. h.	21 m. h.	22 m. h.	23 m. h.	24 m. h.
	1.0145***	1.0154 ***	1.0165***	1.0172***	1.0183***	1.0193***	1.0202***	1.0211***
	(0.00170)	(0.00212)	(0.00230)	(0.00223)	(0.0021)	(0.00203)	(0.00197)	(0.0019)

Notes: The table presents the results of the Mincer-Zarnowitz regression test. The columns represent the results for the RW model. h represents the forecast month horizons with standard error in parentheses. $*$, $**$, $***$ denote statistical significance at the 10%, 5%, and 1% levels respectively.

ing performance using the root mean square error (RMSE) and mean absolute error (MAE) for different forecasting horizons. We take the ratio between the Gordon and the RW model: a value greater than one means a better forecasting performance of the RW specification, while a value lower than one means a better forecasting performance of the Gordon model specification. Moreover, using the Diebold Mariano test, we determine whether the difference in forecasting performance is significantly different between the two specifications.

From tables 3, we observe that for forecasting horizons $\{h = 1, 2, 3, ..., 9\}$ months, the RMSE is smaller in the RW model than in the Gordon model. These results suggest that the RW model has a better forecasting performance in the short horizon than the Gordon model. Conversely, for the forecasting horizons $\{h = 10, 11, 12, ..., 24\}$ months, RMSE is smaller with Gordon specification compared to the RW hypothesis. These results are confirmed by Figure 3. With respect to MAE, we obtain similar results; for the forecasting horizons $\{h = 1, 2, 3, ..., 8\}$ months, RMA is smaller with random specification compared to the Gordon hypothesis. The opposite for $\{h = 9, 10, 11, ..., 24\}$. To summarize, the obtained results suggest the RW is always better in the short period up to month 9, while for a longer period, the Gordon specification is better. However, the Diebold-Mariano test suggests an important result: the two model specifications are not statistically different from each other in the period, while they are different in the long run.

With respect to the existing literature, even if the comparison regards the structural model with a pure RW model, our results are consistent with these previous studies where, at the short horizon, the RW model outperforms the structural model according to the mean squared error criterion. In the same line, structural models outperform RW characterization at longer horizons with a statistically significant amount. In particular, the obtained results are in line with the results obtained by De Jong et al. (2010); however, in their model, the RW model performs worst than the structural model also in the short run.

Table 3: Out-of-sample forecast results (monthly frequency)

h	1 m. h.	2 m. h.	3 m. h.	4 m. h.	5 m. h.	6 m. h.	7 m. h.	8 m. h.
$RMSE$	1.0497*	1.0558	1.065	1.0769*	1.1040*	1.1374*	1.0972	1.0607
	(-1.74)	(-1.54)	(-1.62)	(-1.96)	(-1.90)	(-1.76)	(-1.34)	(-0.92)
MAE	1.0271	1.0426	1.0388	1.0385	1.0579	1.0908	1.1094	1.0120
	(-0.82)	(-1.16)	(-0.87)	(-0.71)	(-1.00)	(-1.34)	(-1.32)	(-0.14)

h	9 m. h.	10 m. h.	11 m. h.	12 m. h.	13 m. h.	14 m. h.	15 m. h.	16 m. h.
$RMSE$	1.0681	0.9707	0.8177	0.8554*	0.8256**	0.7623***	0.7500***	0.6638***
	(-0.64)	(-0.32)	(-1.46)	(-1.95)	(-2.49)	(-3.32)	(-5.56)	(-4.32)
MAE	0.8870	0.9239	0.8149**	0.8258*	0.6821***	0.5667***	0.6208***	0.6991***
	(-1.25)	(-0.77)	(-2.32)	(-1.95)	(-2.66)	(-4.09)	(-7.23)	(-4.92)

h	17 m. h.	18 m. h.	19 m. h.	20 m. h.	21 m. h.	22 m. h.	23 m. h.	24 m. h.
$RMSE$	0.7420***	0.7668***	0.8032***	0.7405***	0.7113***	0.6798***	0.6540***	0.6380***
	(-4.55)	(-3.77)	(-3.83)	(-4.53)	(-6.06)	(-7.32)	(-9.15)	(-9.87)
MAE	0.6679***	0.6981***	0.6604***	0.7113***	0.6272***	0.5628***	0.5282***	0.5316***
	(-4.14)	(-4.93)	(-4.10)	(-4.67)	(-6.02)	(-8.02)	(-8.73)	(-8.83)

Notes: The table presents the results of the forecast analysis. The columns represent the ratio of root mean squared error (RMSE) and the ratio of the mean absolute forecast error (MAE) of the Gordon model to that of the RW. A number greater then 1 represents a better performance of the RW hypothesis. h represents the forecast month horizons and the Diebold-Mariano t-statistics is in parentheses. The two forecasts have the same accuracy under the null hypothesis and $*$, $**$, $***$ denote statistical significance at the 10%, 5%, and 1% levels respectively.

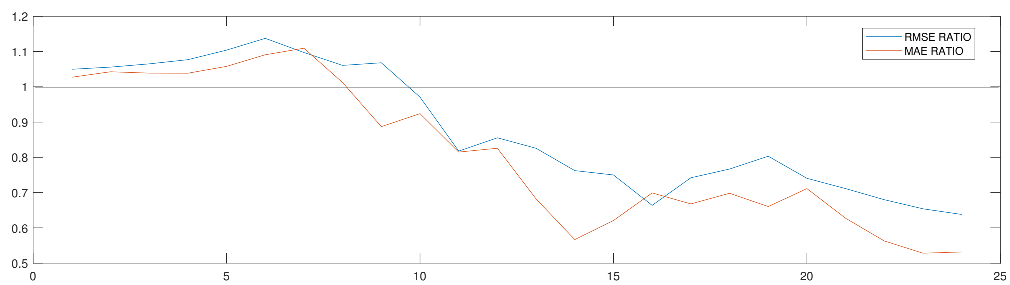

Figure 3: The blue line represents the ratio of RMSE from the Gordon to RMSE from the RW, while the red line represents the ratio of MAE from the Gordon to MAE from the RW. Values below unity indicate better accuracy of point forecasts with Gordon. The forecast horizon is expressed in months.

Quarterly Results. Considering the quarterly data, as we can see from tables 4 and 5, the Mincer-Zarnowitz regression test indicates systematic unbiased in the historical forecasts. In particular, the coefficient is one for all the forecasting horizon specifications, and results do not change between Gordon (table 4) and the RW specification (table 5).

Based on this, as we did for monthly data, we compare the forecasting power of the two model specifications. The comparison of forecasting performance based on RMSE and MAE is summarized in table 6 and in Figure 4. The obtained results confirm the previous results for monthly data: up to 4 quarterly horizon forecast (month 12), both the RMSE and the MAE results suggest that RW and Gordon specifications are not significantly different. Conversely, the Gordon model exhibits a more accurate forecasting ability between 5 and 8 quarterly horizon. However, it is worth noting a nonsignificant result of the Diebold-Mariano test at the six and seven quarterly horizons. One possible explanation is the explosion covid-19 and its temporary effect on the dividend process since march 2020 (the beginning of quarter six in our out-of-sample). In general, our analysis reveals that the forecasting performances of both specifications are indifferent in the short run horizon but not in the long run, when the Gordon approach is superior.

We believe that the results achieved follow from the intrinsic approach adopted. The RW is strongly linked to the trend of the current value of the S&P500; therefore, it fails to capture the long-term changes also because it does not carry relevant information. On the other hand, considering the dividends distributed, the Gordon specification captures the profit prospects of the medium-long term and, consequently, allows better predictive performance.

Table 4: Mincer-Zarnowitz regression test (Gordon - quarterly frequency)

h	1 q. h.	2 q. h.	3 q. h.	4 q. h.	5 q. h.	6 q. h.	7 q. h.	8 q. h.
	1.0012***	1.0022***	1.0007***	1.0025***	1.0044***	1.0065***	1.0079***	1.0088***
	(0.00649)	(0.00648)	(0.00438)	(0.00481)	(0.00211)	(0.00515)	(0.00503)	(0.00369)

Notes: The table presents the results of the Mincer-Zarnowitz regression test. The regression of the actual value on the ex-ante forecast should have a a coefficient of 1. The columns represent the results for the Gordon model. h represents the forecast quarterly horizons with standard error in parentheses. $*$, $**$, $***$ denote statistical significance at the 10%, 5%, and 1% levels respectively. Our results indicate systematic unbiased in the historical forecasts.

Table 5: Mincer-Zarnowitz Regression (RW - quarterly frequency)

h	1 q. h.	2 q. h.	3 q. h.	4 q. h.	5 q. h.	6 q. h.	7 q. h.	8 q. h.
	1.0029***	1.0042***	1.004***	1.0069***	1.0096***	1.0128***	1.0157***	1.0175***
	(0.00576)	(0.00549)	(0.00404)	(0.00417)	(0.00194)	(0.00435)	(0.00422)	(0.00422)

Notes: The table presents the results of the Mincer-Zarnowitz regression test. The regression of the actual value on the ex-ante forecast should have a a coefficient of 1. The columns represent the results for the random model. h represents the forecast month horizons with standard error in parentheses. $*$, $**$, $***$ denote statistical significance at the 10%, 5%, and 1% levels respectively. Our results indicate systematic unbiased in the historical forecasts.

Table 6: Out-of-sample forecast results (quarterly frequency)

h	1 q. h.	2 q. h.	3 q. h.	4 q. h.	5 q. h.	6 q. h.	7 q. h.	8 q. h.
$RMSE$	1.1117	1.1456	1.0209	1.0019	0.6563***	0.8809	0.8104*	0.6838**
	(-1.38)	(-1.56)	(-0.13)	(-0.01)	(-3.41)	(-0.68)	(-1.83)	(-2.29)
MAE	1.0616	1.1400	0.9664	1.0048	0.6175***	0.8340	0.7716*	0.8719**
	(-0.81)	(-1.57)	(-0.25)	(-0.02)	(-2.68)	(-1.10)	(-1.65)	(-2.15)

Notes: The table presents the results of the forecast analysis. The columns represent ratio of root mean squared error (RMSE) and the ratio of the mean absolute forecast error (MAE) of the Gordon model to that of the RW. A number greater then 1 represents a better performance of the RW hypothesis. h represents the forecast month horizons and the Diebold-Mariano t-statistics is in parentheses. The two forecasts have the same accuracy under the null hypothesis and $*$, $**$, $***$ denote statistical significance at the 10%, 5%, and 1% levels respectively.

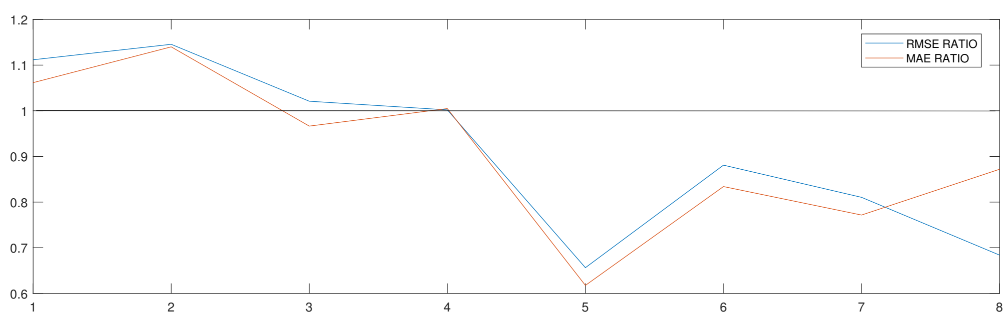

Figure 4: The blue line represents the ratio of RMSE from the Gordon to RMSE from the RW, while the red line represents the ratio of MAE from the Gordon to MAE from the RW. Values below unity indicate better accuracy of point forecasts with Gordon. The forecast horizon is expressed in quarterly.

4 Concluding remarks

Notwithstanding the recent progress made in the field of the empirical validation of HAMs, only a few empirical works analyze the forecasting performance in out-of-sample analysis. At the same time, the state-space model approach has not been used for this purpose; we aimed to fill this gap. To do this, we combine the state-space model approach with the heterogeneous based forecast rules to investigate the forecasting power of an heterogeneous agent model with two different fundamental specifications.

The key findings of our paper are the followings. Overall, our results suggest that the model has out-of-sample predictive power. More precisely, in terms of RMSE and MAE ratio, the RW and Gordon are not statistically different in the short horizon. While, in the long horizon, the Gordon approach statistically outperforms the model with the RW hypothesis. In other words, the two specifications are significantly different only in the long run. This means that the model's forecasting power with Gordon vis a vis the RW improves as the forecasting horizon increases.

Future research can be proposed. Theoretically, an interesting extension would be the inclusion of agent heterogeneity not only in the expectations but also in the investment horizons, as in the work by Anufriev and Bottazzi (2012). Empirically speaking, the present work can be extended by introducing different asset markets. The difficulty in considering and defining a fundamental value is particularly true in international economics, when agents have to estimate the fundamental value of the exchange rate. For example, many different models have been used to analyze the foreign exchange market, such as the purchasing power parity theory (PPP) or the balance of payment approach (BP). In this sense, future analysis can help in examining the different forecasting power of different exchange rate fundamental specifications inside the heterogeneous agent models.

In conclusion, our analysis shows that the state-space approach for HAMs not only is able to explain the in-sample mechanism beyond the price dynamics but also provides a reasonable out-of-sample output prediction with different prediction powers for different fundamental values specified in the literature. We believe that this relationship can represent an important area for future research.

References

Anufriev M, Bottazzi G (2012). Asset pricing with heterogeneous investment horizons. Studies in Nonlinear Dynamics & Econometrics 16:4.

Bao T, Hennequin M, Hommes C, Massaro D (2020). Coordination on bubbles in large-group asset pricing experiments. Journal of Economic Dynamics and Control 110:103702.

Bao T, Cars H , Jiaoying P. (2021) Expectation formation in finance and macroeconomics: A review of new experimental evidence. Journal of Behavioral and Experimental Finance 32:100591.

Boswijk H P, Hommes C H, Manzan S (2007) Behavioral heterogeneity in stock prices. Journal of Economic Dynamics and Control 31:1938-1970.

Brock W A, Hommes C H (1998) Heterogeneous beliefs and routes to chaos in a simple asset pricing model. Journal of Economic Dynamics and Control 22:1235-1274.

Chen Z, Lux T (2018) Estimation of sentiment effects in financial markets: A simulated method of moments approach. Computational Economics 52:711-744.

Cheung Y W, Chinn M D, Pascual A G, Zhang Y (2019) Exchange rate prediction redux: New models, new data, new currencies. Journal of International Money and Finance 95:332-362.

Chiarella C, He X Z, Huang W, Zheng H (2012) Estimating behavioural heterogeneity under regime switching Journal of Economic Behavior and Organization 83:446-460.

Chiarella C, He X Z, Zwinkels R C J (2014) Heterogeneous expectations in asset pricing: Empirical evidence from the S&P500. Journal of Economic Behavior and Organization 105:1-16.

Delli Gatti D, Grazzini J (2020) Rising to the challenge: Bayesian estimation and forecasting techniques for macroeconomic Agent Based Models. Journal of Economic Behavior and Organization 178:875-902.

De Grauwe, P, Grimaldi, M, 2005. Heterogeneity of agents, transaction costs and the exchange rate. Journal of Economic Dynamics and Control 29:691-719.

De Grauwe, P, Grimaldi, M, 2006. Exchange rate puzzles, a tale of switching attractors. European Economic Review 50:1-33.

De Jong E, Verschoor W F, Zwinkels R C (2010) Heterogeneity of agents and exchange rate dynamics: Evidence from the EMS. Journal of International Money and Finance 29:1652-1669.

Diebold FX, Mariano RS (1995) Comparing predictive accuracy. Journal of Business & Economic Statistics 13:253-63.

Durbin J, Koopman S J (2012) Time series analysis by state space methods. OUP, Oxford.

Franke R (2008) Models of asset price dynamics from Beja-Goldman to Brock-Hommes: A synopsis of specification details. University of Kiel.

Franke R, Westerhoff F (2011) Estimation of a structural stochastic volatility model of asset pricing. Computational Economics 38:53-83.

Franke R, Westerhoff F (2017) Taking stock: rigorous modeling of animal spirits in macroeconomics. Journal of Economic Surveys 31:1152-1182.

Ghonghadze J, Lux T (2016) Bringing an elementary agent-based model to the data: Estimation via GMM and an application to forecasting of asset price volatility Journal of Empirical Finance 37:1-19.

Goldbaum D, Zwinkels R C J (2014) An empirical examination of heterogeneity and switching in foreign exchange markets. Journal of Economic Behavior and Organization 107:667-684.

Gordon M J (1959) Dividends, earnings, and stock prices. The Review of Economics and Statistics 41:99-105.

Grazzini J, Richiardi M (2015) Estimation of ergodic agent-based models by simulated minimum distance. Journal of Economic Dynamics and Control 51:148-165.

Gusella F, Stockhammer E (2021) Testing fundamentalist-momentum trader financial cycles. An empirical analysis via the Kalman filter. Metroeconomica 72:758-797.

Gusella F, Ricchiuti G (2022) State Space Model to Detect Cycles in Heterogeneous Agents Models. Available at SSRN 3856471.

Harvey A C (1989) Forecasting, structural time series models and the Kalman filter. Cambridge University Press, Cambridge.

Harvey A C (2006) Forecasting with unobserved components time series models. Handbook of economic forecasting 1:327-412.

Hommes C (2021) Behavioral and experimental macroeconomics and policy analysis: A complex systems approach. Journal of Economic Literature 59:149-219.

Hommes C, in't Veld D (2017) Booms, busts and behavioural heterogeneity in stock prices. Journal of Economic Dynamics and Control 80:101-124.

Hommes C., Sonnemans J., Tuinstra J., Van de Velden H., 2005. Coordination of Expectations in Asset Pricing Experiments. The Review of Financial Studies, 18(3), 955-980.

Hommes C., Sonnemans J., Tuinstra J., Van de Velden H., 2008. Expectations and bubbles in asset pricing experiments. Journal of Economic Behavior and Organization, 67(1), 116-133.

Jongen R, Verschoor W F, Wolff C C, Zwinkels, R C (2012) Explaining dispersion in foreign exchange expectations: A heterogeneous agent approach. Journal of Economic Dynamics and Control 36:719-735.

Kouwenberg R, Markiewicz A, Verhoeks R, Zwinkels, R C (2017) Model uncertainty and exchange rate forecasting. Journal of Financial and Quantitative Analysis 52:341-363.

Kukacka J, Barunik J (2017) Estimation of financial agent-based models with simulated maximum likelihood. Journal of Economic Dynamics and Control 85:21-45.

Lof M (2012) Heterogeneity in stock prices: A STAR model with multivariate transition function. Journal of Economic Dynamics and Control 36:1845-1854.

Lof M (2015) Rational Speculators, Contrarians, and Excess Volatility. Management Science 61:1889-1901.

Lux T (2018) Estimation of agent-based models using sequential Monte Carlo methods. Journal of Economic Dynamics and Control 91:391-408.

Lux T (2021) Can heterogeneous agent models explain the alleged mispricing of the S&P 500? Quantitative Finance 21:1413-1433.

Lux T (2021b) Bayesian estimation of agent-based models via adaptive particle Markov chain Monte Carlo. Computational Economics, 1-27.

Lux T, Zwinkels R C J (2018) Empirical validation of agent-based models. Handbook of computational economics 4:437-488.

Manzan S, Westerhoff F (2007) Heterogeneous expectations, exchange rate dynamics and predictability. Journal of Economic Behavior and Organization 64:111-128.

Meese R A, Rogoff K (1983) Empirical exchange rate models of the seventies, do they fit out of sample? Journal of International Economics 14:3-24.

Mincer J, Zarnowitz V (1969) The Evaluation of Economic Forecasts. In J. Mincer (Ed.), Economic Forecasts and Expectations. New York: National Bureau of Economic Research, 1-46.

Recchioni M C, Tedeschi G, Gallegati M A (2015) A calibration procedure for analyzing stock price dynamics in an agent-based framework. Journal of Economic Dynamics and Control 60:1-25.

Schmitt N (2021) Heterogeneous expectations and asset price dynamics. Macroeconomic Dynamics 25:1538-1568.

Ter Ellen S, Verschoor W (2018) Heterogeneous beliefs and asset price dynamics: a survey of recent evidence. In: Jawadi, F. (Eds.). Essays in honor of Georges Prat, Springer, Nature Switzerland, pp 53-79.

Ter Ellen S, Hommes C, Zwinkels R C J (2021) Comparing behavioural heterogeneity across asset classes. Journal of Economic Behavior and Organization 185:747-769.

Wan, J Y, Kao C W (2009). Evidence on the contrarian trading in foreign exchange markets. Economic Modelling 26:1420-1431.

Appendix A

In this appendix we explain the Kalman filtering procedure.

The optimal forecasting rule has the form

$$Z_{t\,|\,t} = Z_{t\,|\,t-1} + K_t\left(P_t - P_{t\,|\,t-1}\right)$$

where K is a weight that changes as new information becomes available and Z is the state vector containing the unobserved beliefs. $Z_{t\,|\,t}$ denotes the forecast of the state variable once P_t is realized while $Z_{t\,|\,t-1}$ and $P_{t\,|\,t-1}$ denote respectively the forecast of variables Z_t and P_t before P_t is realized.

Now we can select the optimal value of K_t to minimize the mean square prediction error at time t

$$\min_{k_t} E_t\big(Z_t - Z_{t\,|\,t}\big)^2 = \min_{k_t} E_t\big[Z_t - \big(Z_{t\,|\,t-1} + K_t\left(P_t - P_{t\,|\,t-1}\right)\big)\big]^2$$

Using Eq. (1) and indicating with H the vector containing the quotas of the groups of agents, we obtain

$$\min_{k_t} E_t\big[Z_t - \big(Z_{t\,|\,t-1} + K_t\left(HZ_t - HZ_{t\,|\,t-1}\right)\big)\big]^2$$

Optimizing with respect to K_t we get

$$-2H\left(I - HK_t\right)E_t\big(Z_t - Z_{t\,|\,t-1}\big)^2 = 0$$

Indicating with $\Gamma_{t\,|\,t-1} = E_t\big(Z_t - Z_{t\,|\,t-1}\big)^2$, we obtain

$$-2H\left(I - HK_t\right)\Gamma_{t\,|\,t-1} = 0$$

Solving for K_t we obtain

$$K_t = \frac{H\;\Gamma_{t\,|\,t-1}}{H\;\Gamma_{t\,|\,t-1}H'}$$

Now we can regroup the equations. Indicating the transition matrix with A, we obtain that

$$Z_{t\,|\,t-1} = AZ_{t-1\,|\,t-1} \tag{10}$$

$$\Gamma_{t\,|\,t-1} = A\,\Gamma_{t-1\,|\,t-1}A' + Q \tag{11}$$

$$P_{t\,|\,t-1} = HP_{t-1\,|\,t-1}$$

Eqs. (10) and (11) are the so-called prediction equations in the Kalman filtering. The other equations we need are the three updating equations which are

$$K_t = \Gamma_{t\,|\,t-1} H' (\psi_t)^{-1} \tag{12}$$

with

$$\psi_t = H \Gamma_{t\,|\,t-1} H'$$

$$Z_{t\,|\,t} = Z_{t\,|\,t-1} + K_t \left(P_t - P_{t\,|\,t-1} \right) \tag{13}$$

$$\Gamma_{t\,|\,t} = (I - K_t H)\, \Gamma_{t\,|\,t-1} \tag{14}$$

In this case, the inference about Z_t is updated using the observed value of P_t.

We start with a specification information set with initial conditions $Z_{0\,|\,0}$ and $\Gamma_{0\,|\,0}$. Then we use the prediction Eqs. (10) and (11) to obtain $Z_{1\,|\,0}$ and $\Gamma_{1\,|\,0}$. Once we observe P_1 we use the updating Eqs. (12), (13), and (14) to obtain $Z_{1\,|\,1}$, $\Gamma_{1\,|\,1}$ and $P_{1\,|\,1}$. We next use this information to form $Z_{2\,|\,1}$ and $\Gamma_{2\,|\,1}$, then forecasts are updated and we continue to repeat this process until the end of the dataset.

Given the vector prediction errors $\mu_t = P_t - P_{t\,|\,t-1}$ and the variance-covariance matrix ψ_t, the log-likelihood function can be maximized.

$$\log l = -\tfrac{T}{2} \ln (2\pi) - \tfrac{1}{2} \sum_{t=1}^{T} \ln \left(\left| \psi_{t\,|\,t-1} \right| \right) - \tfrac{1}{2} \sum_{t=1}^{T} \mu_t{}' \left(\psi_{t\,|\,t-1} \right)^{-1} \mu_t$$

Appendix B

The Mincer-Zarnowitz test verifies that forecast errors have zero means and are uncorrelated with any other variable at the time of the forecast. To perform the test for out-of-sample predictions, we run the following regression:

$$p_{T+h} = b_0 + b_1 p_{\,T+h|T} + \varepsilon_{\,T+h|T}$$

.

where p_{T+h} and $p_{\,T+h|T}$ are the actual and predicted values of asset prices respectively.

In order to have:

$$\varepsilon_{\,T+h|T} = p_{T+h} - p_{\,T+h|T}$$

These two conditions should hold:

$$b_0 = 0, \quad b_1 = 1$$

i.e., the regression of the actual value on the ex-ante forecast should have a zero intercept and a coefficient of 1. Technically, the null hypothesis is:

$$H_0 : b_0 = 0,\ b_1 = 1$$

against the alternative one:

$$H_1 : b_0 \neq 0,\ b_1 \neq 1$$

If the null hypothesis is rejected, it indicates systematic bias or inefficiency in the forecasts.

Appendix C

We perform the Diebold-Mariano test to compare the predictive accuracy of different forecasts. We define the forecast errors with Gordon specification as:

$$\varepsilon^G_{T+h|T} = p_{T+h} - p^G_{T+h|T}$$

while for the RW as:

$$\varepsilon^{RW}_{T+h|T} = p_{T+h} - p^{RW}_{T+h|T}$$

The loss associated with Gordon forecast and RW forecast is assumed to be a function k of the forecast errors, $k\left(\varepsilon^G_{T+h|T}\right)$ and $k\left(\varepsilon^{RW}_{T+h|T}\right)$ respectively. We denote these functions with the squared-error loss and the absolute value.

In this way, for the Gordon specification, we obtain:

$$k\left(\varepsilon^G_{T+h|T}\right) = \left(\varepsilon^G_{T+h|T}\right)^2 \quad and \quad k\left(\varepsilon^{RW}_{T+h|T}\right) = \left(\varepsilon^{RW}_{T+h|T}\right)^2$$

For the specification with the RW, we obtain:

$$k\left(\varepsilon^G_{T+h|T}\right) = \left|\varepsilon^G_{T+h|T}\right| \quad and \quad k\left(\varepsilon^{RW}_{T+h|T}\right) = \left|\varepsilon^{RW}_{T+h|T}\right|$$

The loss differential between the two forecasts is:

$$d_{T+h} = k\left(\varepsilon^{G}_{T+h|T}\right) - k\left(\varepsilon^{RW}_{T+h|T}\right)$$

The two forecasts have equal accuracy if the loss differential has zero expectation for all $T+h$. The null hypothesis states that the Gordon and RW forecasts have the same predictive power:

$$H_0 : E\left(d_{T+h}\right) = 0 \quad \forall \left(T+h\right)$$

The alternative hypothesis states that the Gordon and RW forecasts have different levels of accuracy:

$$H_1 : E\left(d_{T+h}\right) \neq 0$$

Under H_0, the Diebold-Mariano test statistics is:

$$\frac{\bar{d} - u}{\sqrt{\sigma^2/h}} \to N\left(0, 1\right)$$