Professional Survey Forecasts and Expectations in DSGE Models

RYCHALOVSKA, YULIYA; Slobodyan, Sergey; Wouters, Rafael

Publication date:
2023

Professional Survey Forecasts and Expectations in DSGE Models

Yuliya Rychalovska*, Sergey Slobodyan†, Rafael Wouters‡

November 1, 2023

Abstract

This paper proposes a strategy to exploit timely information from survey data in DSGE models. We include the Survey of Professional Forecasters data on consumption, investment, output, and inflation expectations in the set of observable variables to discipline the dynamics of model-based predictions and evaluate alternative belief specifications. Due to surveys, we are able to improve identification of fundamental shocks and models' predictive power by separating the sources of low and high frequency volatility. Models with an imperfectly-rational, time-varying expectation formation mechanism based on Adaptive Learning can reduce limitations implied by the Rational Expectation hypothesis and process survey data more effectively.

JEL Classification: C5, D84, E3

Keywords: Expectations, Survey data, Adaptive Learning, DSGE models

*University of Namur; email: yuliya.rychalovska@unamur.be

†CERGE-EI; email: sergey.slobodyan@cerge-ei.cz

‡National Bank of Belgium; email: rafael.wouters@nbb.be

Structural macro-models have become an important tool used by policymakers for economic analysis, forecasting and decision-making. The application of these models to describe the business cycle and generate accurate forecasts can enhance policy credibility, improve communication between the central bank and the public, and thus facilitate the achievement of better macroeconomic outcomes.

In the current dynamic environment, the use of timely information is crucial to develop reasonable economic assessments and generate reliable projections of the future state of the economy. The recent generation of macro-models traditionally rely on the use of realized data for macroeconomic analysis and forecasting. However, the role of expectations, which incorporate additional up-to-date information and judgement, can be substantial. Today, the availability of survey data, which contain timely information collected by experts and the public, expands the information set used in policy-oriented analysis.

Recent literature has highlighted various gains from the use of survey data in macroeconomics. In particular, an important feature of survey-based expectations is their rich information content and strong correlation with economic outcomes. A number of recent studies document the excellent forecasting performance of survey data (Ang et al., 2007) and illustrate the benefits of integrating surveys into economic forecasting with reduced form (Clark et al., 2017; Tallman 2020) and structural models (Milani 2011; Del Negro et al., 2013; Carvalho et al., 2023). Data on real-time expectations play an important role in the identification of monetary policy shocks (Romer and Romer, 2004) and the reaction function of central banks (Coibion and Gorodnichenko 2011; Bauer et al., 2022). Crump et al. (2022) explore survey expectations to estimate the parameters of the consumption Euler equation. Vast literature focuses on studying the features and macroeconomic implications of inflation expectations of households (Coibion et al., 2020; Malmendier and Nagel 2016) or professional forecasters (Andrade et al., 2013; Coibion and Gorodnichenko 2012; Mankiw et al., 2003).

A prominent strand of the literature emphasizes the usefulness of real-time expectations through survey data for understanding and modeling the expectation formation mechanism, which plays a central role in the transmission of shocks and determining the effects of monetary policy measures in the New-Keynesian models. In particular, numerous studies explore the statistical properties of survey-based forecasts in order to test the empirical validity of the Rational Expectation (RE) hypothesis, which is commonly assumed in the DSGE modelling literature and implies that agents' expectations are equivalent to model-consistent optimal forecasts. At the same time, the availability of survey datasets enables us to assess the extent to which observed expectations reflect the features of predictions based on full and perfect information. Various papers document convincing evidence on departures from Full Information Rational Expectations (FIRE) for different macroeconomic variables and forecast datasets, at the level of both the individual and aggregate forecasts. In particular, Coibion and Gorodnichenko (CG) (2015) test the FIRE hypothesis by assessing the predictability of the average forecast errors by forecast revisions obtained from the Survey of Professional Forecasters (SPF). They report the widespread evidence of systematic underreaction to news in inflation and other macro forecasts and propose models of information rigidities that are able to generate similar properties of expectations. Angeletos et al. (2020) confirm the finding of CG and document that, for inflation and unemployment, aggregate forecast errors are positively related to lagged forecast revisions. Bordalo et al. (2020) demonstrate inefficient adjustment of both the individual and consensus forecasts from the SPF and the Blue Chip Survey of macroeconomic and financial variables. Several studies explore the feature of imperfect rationality of survey expectations to explain the observed macroeconomic dynamics, and refine economic theories and modeling approach. Coibion et al. (2018) demonstrate how incorporating survey data on inflation expectations can address a number of otherwise puzzling features of FIRE Phillips curves such as the need for ad hoc lags and missing disinflation during the Great Recession.

The above literature generally agrees on the finding that survey forecasts depart from the FIRE hypothesis and that surveys represent useful sources of timely information, which can be used to improve macroeconomic analysis. However, determining the best approach to modelling the non-rational expectations of economic agents remains a subject to debate. Various studies advocate for alternative specifications of the expectation formation mechanism, which can potentially produce different macroeconomic effects of shocks and policy prescriptions.[1] An additional important question is how to incorporate and properly explore survey forecasts in structural macro models.

In this paper, we demonstrate the usefulness of survey data for macroeconomic analysis and propose a strategy to integrate and efficiently utilize information from surveys in the DSGE setup. We exploit the SPF forecasts to better understand the expectation formation and belief-updating processes. In particular, we evaluate the ability of a range of models with non-rational belief specifications to generate forecasts that reflect the actual properties of agents' expectations in line with extensively documented empirical evidence. We also investigate whether the proposed deviations from the complete rationality assumption can reduce the limitations faced by the model with RE.[2]

We develop our approach in several steps. Firstly, we empirically assess the degree of predictability in the survey nowcast errors for our data sample. We can confirm inefficient information processing and the underestimation problem in SPF forecasts, but only in the investment data. We then formulate and estimate the medium-scale DSGE model with alternative belief specifications that may deviate from the complete rationality assumption. We want to build a framework able to generate the model-based measures of expectations con-

[1] Early approaches that attempted to reconcile economic theory and empirical evidence maintained the RE hypothesis but assumed imperfect information and price adjustment (Sims 2003; Woodford 2003; Mankiw and Reis 2002; Gabaix 2019). Other contributions introduce a deeper departure from rationality as in Evans and Honkapohja (2001), Eusepi and Preston (2011), Milani (2007), and Hommes (2020).

[2] In a follow-up paper Rychalovska et al. (2023b), we focus on studying the responses of expectations to business cycle shocks. We also investigate whether the proposed deviation from the FIRE can alter the general macroeconomic effects of shocks in line with other competing theories of expectations.

sistent with both the behaviour of professional forecasters and the dynamics of the main macroeconomic variables. Our approach to modelling the imperfectly rational expectations is grounded in the Adaptive Learning (AL) mechanism as in Evans and Honkapohja (2001), Woodford (2013), Milani (2007, 2011), Eusepi and Preston (2011), Hommes (2019), and Molavi (2019). This approach implies that agents formulate their expectations based on estimated forecasting models that reflect the information available at the time of their decisions. More specifically, we assume the following alternative belief specifications: 1) Minimum State Variable (MSV) beliefs based on RE information set with flexible constant updating; 2) Restricted beliefs (RB), which imply that agents use a restricted information set to generate predictions; and 3) Heterogeneous beliefs (HB), which assume that agents can switch between belief models based on their past forecasting performance. We assume that agents update their beliefs gradually as soon as new information becomes available on the basis of Kalman filter learning algorithm. The learning models therefore imply the time varying transmission mechanism.

In order to discipline the behaviour of the expectation variables in the model and achieve consistency with the survey forecasts, we integrate the survey data on consumption, investment, output, and inflation expectations, as measured by the SPF, into the standard set of observable variables of the DSGE model. Therefore, this paper represents an extension and generalization of the results shown in Rychalovska, Slobodyan, and Wouters (*RSW*) (2023a), who focus on inflation expectations and integrate the inflation survey (nowcast) combined with the AL mechanism in the Smets-Wouters DSGE model. The study shows that inflation survey information helps to identify separately the innovations in the i.i.d. and the persistent markup shocks. In the model with AL, observing inflation SPF data improves the inflation forecast. The time-variation produced by the updating of beliefs captures the dynamics in the mean and the volatility of the inflation process. The model therefore provides a consistent interpretation of actual inflation developments, inflation expectations and the uncertainty around the future inflation developments.

In our study, we emphasize the usefulness of timely information contained in real survey forecasts for explaining business cycle fluctuations and improving the predictive power of macro-models. We illustrate that the integration of consumption, investment and output forecasts/nowcasts into the structural DSGE model solves an important information filtering problem and facilitates the identification of shocks playing a key role in driving business cycle fluctuations, such as the risk premium, investment specific technology and government spending shocks. Due to the information contained in the survey data, we are able to distinguish between the persistent and the i.i.d. components of these structural shocks, thus separating the sources of low and high frequency volatility. We also illustrate that shock decomposition into persistent and transitory components is essential to exploit optimally the information in the surveys and to achieve a significant improvement in the model fit and forecasting performance. In addition, we demonstrate that with our fundamental shock re-specification we can jointly explain the survey and realized real-time macro data without assuming additional 'exogenous' sentiment shocks, which are otherwise needed in order to fit the survey forecasts and generate the role for expectation factors in the model. In this respect, our paper is related to the work of Milani (2017), who also considers professional forecasts on inflation, investment and consumption growth and integrates them into the DSGE model with the AL mechanism of expectation formation. That study exploits the observed expectation data in order to identify the so-called sentiment shocks, which are assumed to be orthogonal to structural fundamental shocks. "Sentiment" is needed to capture the excesses of optimism and pessimism, which are observed in survey forecasts but could not be generated by the learning model to a sufficient degree. Milani (2017) shows that exogenous sentiment shocks are responsible for a significant portion of historical US business cycle fluctuations.

In the model comparison exercise, we contrast the empirical properties of the models with RE and alternative beliefs to see which approach is more consistent with the timely information contained in the surveys. We show that the models that deviate from the complete rationality assumption outperform in terms of the model fit. We can therefore validate the deviation from the RE hypothesis. In addition, our results indicate that AL model based on small autoregressive forecasting functions, which have been used extensively in the learning literature[3], cannot produce forecasts of the similar precision to the SPF. Therefore, our RB learning model relies on expanded information set, which implies that belief specification is augmented with innovations of the structural shocks. We interpret this result as an evidence that agents tend to use more complete information set in forming their beliefs.

Furthermore, we illustrate that our models with alternative beliefs can reduce some of the limitations imposed by the RE hypothesis. More specifically, we show that our learning models based on time-variation of agents' beliefs can generate the endogenous fluctuations in the long-term mean and volatility of the model variables. Therefore, the AL mechanism improves the ability of the model to deal with the trend breaks observed in the data and can generate time-varying responses to macroeconomic shocks. Finally, we illustrate that the model with adaptive learning can do better than RE in processing the signals contained in the surveys. By relaxing the constraint that expectations are formed consistently with the actual law of motion, we avoid the problem that model-based predictions inherit the inefficiencies that are present in the survey forecasts. As a result, the AL model shows lower predictability of forecast errors.

Therefore, the AL mechanism, which can successfully reconcile survey data, model-based expectations and realized macro variables, represents a competitive and flexible framework useful to study the nature of the expectation formation process and its implications for the business cycle and policy analysis.

This paper is organized as follows. We first present details on our dataset and discuss the motivation for the use of SPF-nowcasts as proxy variables for expectations. In section 2, we describe the RE model specification, and analyse

[3] Slobodyan and Wouters (2012), Hommes (2014).

its performance and limitations. Section 3 focuses on the models with alternative belief specifications. Section 4 describes the model comparison results. Section 5 demonstrates how time-varying expectation formation mechanims implied by AL can reduce the limitations of models based on the full rationality assumption. Section 6 concludes. The online Appendix contains more detailed estimation results and robustness check exercises.

1 Why SPF expectations?

In this project, we focus on the expectation measures published by the Survey of Professional Forecasters (SPF). SPF is a quarterly survey of forecasts for a variety of macroeconomic variables in the United States. For the period 1981q2-2019q2, we explore four SPF time series: inflation, consumption, output and investment (residential and non-residential).[4] Given the timing of the SPF forecast data, we consider nowcasts as a proxy for the expectations in our model. SPF nowcast is defined as a prediction formed in the middle of the period $t+1$ for the period $t+1$, given the information for period t. As a result, this concept is suitable to represent the model-based expectations $E_t y_{t+1}$.

To motivate the use of SPF-nowcasts as proxy variables for expectations and illustrate the role of survey data for the business cycle, we first examine the evidence obtained from the purely data-driven empirical model, which combines the standard real-time macro-variables and survey data but abstracts from any specific assumptions regarding the expectation formation mechanism. In particular, we estimate VAR models based on 7 standard (Smets-Wouters) macro-variables and one additional survey time series at a time. The data on actual outcomes includes: real GDP, real consumption, real investment, the GDP price index, real wage, hours (all expressed as the log difference), and the three-month TB rate. For the first six variables we consider the first release from the Philadelphia Fed Real-Time Data Set for Macroeconomists, which is described in more detail by Croushore and Stark (2001). It is important to present the model in a real-time data environment so that the estimation is based on a dataset that is consistent with the information available to agents that produce the survey data.[5]

Based on each of the 4 models, we evaluate the impact of the corresponding survey series on the macro variables. We order the survey last in the Cholesky decomposition to capture the fact that nowcasts contain timely information and do react contemporaneously to other shocks. As a result, we can identify the innovation in survey data that incorporates information not reflected in other macro variables. This set-up allows us to assess the importance of this shock for explaining the business cycle. Table 1 illustrates that, in a reduced-form

[4] The beginning of the sample corresponds to the first available observations for consumption and investment.

[5] Real-time data and SPF data are available on the Philadelphia FED web-site: https://www.philadelphiafed.org/research-and-data/real-time-center/. SPF data for consumption, investment, and output are available in levels. We therefore transform them into the growth rates.

VAR decomposition exercise, the survey nowcasts shocks explain a substantial fraction of the forecast error variance of our real-time macro variables. Thus SPF data is very informative, and including the nowcasts made by professionals in the empirical analysis should improve the explanatory power of the model.

Table 1: 5-year variance decomposition and the role of SPF

	Innovation in the SPF-nowcast for			
Fraction of variance explained in real-time data for:	Inflation	Consumption	Investment	GDP
Inflation	**0.19**	0.03	0.07	0.06
Consumption	0.09	**0.33**	0.12	0.11
Investment	0.02	0.24	**0.33**	0.19
GDP	0.04	0.21	0.17	**0.29**
Short rate	0.01	0.13	0.33	0.19
Nowcast	0.44	0.71	0.57	0.50

SPF nowcasts are generally considered precise and timely forecasts for macroeconomic aggregates. Professional forecasters utilize the information advantage from observing current high-frequency data, thus processing a large number of up-to-date economic and financial indicators and adjust their forecasts flexibly to account for changes in the macroeconomic dynamics. In addition, surveys capture some non-fundamental aspects, such as perceptions and sentiment, which also can be important drivers of the business cycle. On the other hand, macro-economic models can process only a limited set of macro-economic aggregates that are measured in real-time with considerable uncertainty and measurement error.

Table 2 demonstrates the impressive ability of survey nowcasts to predict the first release of key macroeconomic variables relative to the 1-step ahead forecasts from the Smets and Wouters (SmW) (2007) model estimated on the Real-Time data. We report the Root Mean Squared Forecast Errors (RMSFEs) and the results of the Diebold-Mariano (DM) test, which verifies the statistical significance of the difference in the RMSFEs. Table 2 illustrates that the predictions for the main macroeconomic variables in a standard model differ significantly from the survey forecasts. This suggests that ignoring the data on surveys may lead to pronounced divergence between the observed and the model-based measures of expectations. As a result, valuable information can be omitted from the macroeconomic analysis, and the role of expectation factors in driving the business cycle might be misspecified. Therefore, we argue that survey data represents a helpful source of timely information that can be integrated into the DSGE model and efficiently exploited.[6]

Table 2: RMSFE and DM test statistics: SPF nowcasts and RE model

[6] We do not find additional valuable information in the longer-term SPF forecasts. The quality of these predictions is comparable to the accuracy of the model-based forecasts. As a result, we limit our analysis to SPF nowcasts.

RMSE 1Q ahead	Inflation	Consumption	Investment	GDP
SPF - nowcast	0.21	0.43	1.49	0.35
RE_{noSPF}	1.1**	1.26***	1.19***	1.26***

Note: RMSFEs statistics is based on in-sample predictions over the sample period 1981Q2 - 2019Q2. RE_{noSPF} denotes the Smets and Wouters (2007) DSGE model estimated on the Real-Time data only, without including the SPF data. For the SPF, RMSFEs are reported in levels. For the RE_{noSPF} model, they are presented in terms of the ratios of the RMSFE for the DSGE model to the corresponding RMSFE for the SPF nowcast. ***, **, and * denote the rejection of the null of the DM test, stating that the RMSFEs from the SPF and DSGE do not differ significantly, at 1%, 5%, and 10% significance level, respectively. Therefore, values above 1 marked with stars indicate that SPF nowcasts are statistically better than the DSGE model predictions.

To get more insights into the properties of the SPF forecasts included in our sample, we follow the approach developed in CG (2015), which enables us to verify the deviations of forecasts from the FIRE hypothesis and understand whether the departures from the complete rationality assumption are due to under- or overreaction. More specifically, we run the following regression to test whether the SPF forecast errors are predictable from the forecasts revisions:

$$y_{t+1} - y_{t+1|t} = a + b \cdot (y_{t+1|t} - y_{t+1|t-1}), \tag{1}$$

where $y_{t+1|t}$ is the forecast made at quarter t about the next quarter value y_{t+1} of the variable. The term $(y_{t+1|t} - y_{t+1|t-1})$ therefore denotes one-period change in the forecast about y_{t+1}, i.e. the forecast revision. The values of b-coefficient significantly different from zero indicate that forecast errors can be predicted by the forecast revisions. This should not happen under FIRE. Moreover, a positive (negative) b-coefficient indicates underreaction (overreaction) of SPF forecasts to news relative to FIRE.

The results of the predictability test for the survey nowcast errors are presented in Table 3.

Table 3: Predictability test: SPF-nowcasts

	Inflation	Consumption	Investment	GDP
b-coefficient	-0.16	0.17	**0.49**	0.17
St.error	0.16	0.17	0.16	0.10
p-value	0.258	0.463	0.006	0.325

The test shows an important underestimation problem in the investment data: forecast error is significantly positively related to the revision in the forecasts. In other words, we detect the evidence of inefficient information processing because professional forecasters underreact to news available at the time when the forecast is made and revised. Our result is in line with the analysis presented in CG (2015) and Bordalo et al. (2020), who also document the severe predictability problem for investment SPF data but not for inflation, output growth and consumption growth for the sample starting from 1981.[7]

[7] Predictability problem was detected in inflation SPF forecasts for the sample starting 1968.

2 SPF forecasts in DSGE models with rational expectations

2.1 SPF forecasts and (re) specification of fundamental shocks

We estimate a version of the SmW (2007) DSGE model[8] on the combined Real-time and SPF dataset described in section I. We use the 1st (x_{r1}) and 2nd (x_{r2}) releases of the variables x available in real time. The second release is treated as final data. We use the SPF-nowcasts (x_{f0}) for consumption, investment, output and inflation as observables for the model-based measures of expectations. More specifically, we assume that SPF prediction is equal to the model expectations plus the measurement error. The corresponding measurement equations are as follows:

$$dx_{r1,t} = (x_t - x_{t-1}) + \overline{\gamma} + e_t^{x_{r1}} \tag{2}$$
$$dx_{r2,t} = (x_{t-1} - x_{t-2}) + \overline{\gamma} \tag{3}$$
$$dx_{f0,t} = (E_t x_{t+1} - x_t) + \overline{\gamma}_{x_{f0}} + e_t^{x_{f0}}, \tag{4}$$

where measurement errors $e_t^{x_{r1}}$, $e_t^{x_{f0}}$ are modelled as i.i.d. processes[9] and $\overline{\gamma}$ denotes the trend. We allow for trend parameter in the SPF measurement equation to vary from the trend parameter in the equation for the realized data to capture possible shifts in expectations. In addition, we allow for the independent trend parameters in the investment measurement equations (2) and (3) due to significant, systematic updates in the first and second releases of this time series.

The timing assumptions in our model are made in line with the actual data release. In particular, the $r1$ data for quarter t and the $r2$ for the quarter $t-1$ are normally published in the beginning of the quarter $t+1$. Based on this information, SPF participants make predictions and several weeks later the SPF forecasts are issued. As shown in Figure 1, the assumed timing in our setup implies that we evaluate the model at point X after the publication of SPF nowcast $f0$ for $t+1$ and given the $r1$ for time t (and $r2$ for $t-1$).

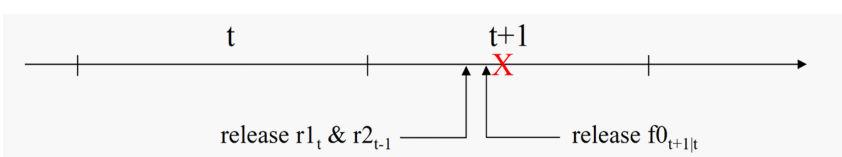

Figure 1: Timing assumptions

[8] For more details see Slobodyan and Wouters (2012, 2022).

[9] As a robustness check, we also estimated a version with AR(1) measurement errors. However, the estimates of the persistency parameters are rather small and do not influence the model fit and the conclusions of our paper.

Therefore, the information set at time t includes $r1$ for t, $r2$ for $t-1$ and $f0$ for $t+1$. Agents are assumed to make their decisions for time t and their expectations for time $t+1$ based on predetermined state vector for $t-1$, parameter beliefs formed at $t-1$, and structural shocks hitting the economy in time t. The model uses $r1$ and $f0$ – as published in the course of $t+1$ – as measurement variables for these decisions, and the Kalman filter delivers estimates for the implied structural shocks.

In order to integrate and properly explore the SPF data within the DSGE setup, we propose to modify the structure of the stochastic processes driving the macroeconomic dynamics. The model specification should be flexible enough to capture the signals contained in the surveys. *RSW* (2023a) apply this structure to the markup shocks. In particular, in *SmW* (2007), the price and wage markup shocks are modelled as $ARMA(1,1)$ processes. Such a specification implies that the same innovation is driving the high frequency MA component and low frequency AR component. Observing the survey inflation expectations provides useful and timely information to distinguish two independent – the transitory and the persistent – components of the innovations driving the markup process. *RSW* (2023a) demonstrate that such shock re-specification is necessary to effectively explore the content of the inflation survey[10]. In our paper, we extend this approach to the survey data on real variables, which allows us to separately identify the *i.i.d.* and persistent components of the fundamental shocks driving the real business cycle, such as the risk premium, and investment specific and government spending shocks.[11] In particular, we re-specify the shock processes as follows:

$$b_t = b_t^{ar} + b_t^{iid} \quad with \quad b_t^{ar} = \rho_b b_{t-1}^{ar} + \varepsilon_t^{b_{ar}} \quad and \quad b_t^{iid} = \varepsilon_t^{b_{iid}}, \tag{5}$$

where $\varepsilon_t^{b_{ar}}$ and $\varepsilon_t^{b_{iid}}$ are two *i.i.d.* innovations. Together with the TFP and monetary policy shocks, we have in total 12 structural innovations in the model. Compared to the original *SmW* (2007) that was estimated on final data, the use of real-time data up to the second release results in different estimates for shocks (later revisions have an important impact).[12]

2.2 Estimated RE model with re-specified shocks

We first estimate several RE versions of the model in order to illustrate the role of the shock re-specification in the presence of survey data. Table 4 reports the logarithm of the marginal likelihood for the estimated specifications.

Table 4. RE models: comparison of the model fit and predictive power

[10] RSW (2023a) show that observing the inflation survey data helps to reconcile the model-based and observed inflation expectations and also improves inflation forecasts.

[11] These shocks are modelled as $AR(1)$ processes in the original specification.

[12] See also Jacobs-Van Norden (2011), Bognanni (2016).

Model	1-q RMSFE				*Log(Marg. Lik)*
	Inflation	Consumption	Investment	GDP	
RE	1.09**	1.23***	1.17***	1.23***	-577.37
$RE_{\{\mu\}}$	1.00	1.20***	1.17***	1.23***	-536.63
$RE_{\{\mu,b\}}$	1.00	0.97*	1.09***	1.15***	-473.29
$RE_{\{\mu,b,qs\}}$	1.00	0.97	1.00	1.15***	-410.84
$RE_{\{all\}}$	1.00	0.98	1.00	1.00	-385.07

Note: Models are evaluated over the sample period 1981Q2 - 2019Q2 using the first four observations as a presample. All models are estimated on the same dataset which includes 17 observable variables: real-time dataset and 4 SPF variables for inflation, and growth rates of consumption, investment and output. Model notations: RE denotes the original SmW model, without any shock respecification; $RE_{\{\mu\}}$ is the RSW (2023a) model with re-specified price and wage markup shocks; $RE_{\{\mu,b\}}$ is the model with re-specified price, wage markup and risk premium shocks; $RE_{\{\mu,b,qs\}}$ is the model with re-specified price and wage markup, risk premium and investment technology shocks; $RE_{\{all\}}$ denotes the model with re-specification of all the fundamental shocks. RMSFEs are reported in terms of the ratios of the RMSFE for the DSGE model to the corresponding RMSFE for the SPF nowcast. ***, **, and * denote the rejection of the null of the DM test, stating that the RMSFEs from the SPF and DSGE do not differ significantly, at 1%, 5%, and 10% significance level, respectively. Therefore, values above 1 marked with stars indicate that SPF nowcasts are statistically better than the DSGE model predictions.

Table 4 clearly illustrates the gain from the separation of each of the fundamental shocks into the high and low frequency components. The results confirm an exceptional model fit of the $RE_{\{all\}}$ model, which explores the survey information in the most efficient way due to the re-specification of all the real and markup shocks.[13] The estimated measurement errors on observed nowcast variables are small and no longer have any predictive information, which can be verified by repeating our Reduced-Form Variance Decomposition exercise (see section I).

Figure 2 visualizes the ability of our model $RE_{\{all\}}$ to identify separately the shock components, which can be distinguished only through the observation of the nowcasts. As can be seen, the consumption and consumption expectations react very differently to the persistent and temporary components of the risk premium shock.

Furthermore, we examine the role of the real nowcasts and shock re-specification for the predictive power of the model. As previously illustrated in Table 2, the survey data have very good predictive power compared to the RE model estimated without observing the nowcasts. Table 4 shows that the RE model incorporating the SPF data and, in addition, separately identifying the i.i.d. and persistent components of all the fundamental shocks ($RE_{\{all\}}$), outperforms the original SmW model with the standard shock processes (RE) and also the specification with the 2-component shock structure only for markups ($RE_{\{\mu\}}$). Comparing the performance of the models relative to the SPF predictions, we observe that the RE model with two-component real and markup

[13] Table A1 in the Appendix, which presents the summary of the estimated parameters, indicates that the estimates of the structural parameters remain rather standard and in line with the previous literature.

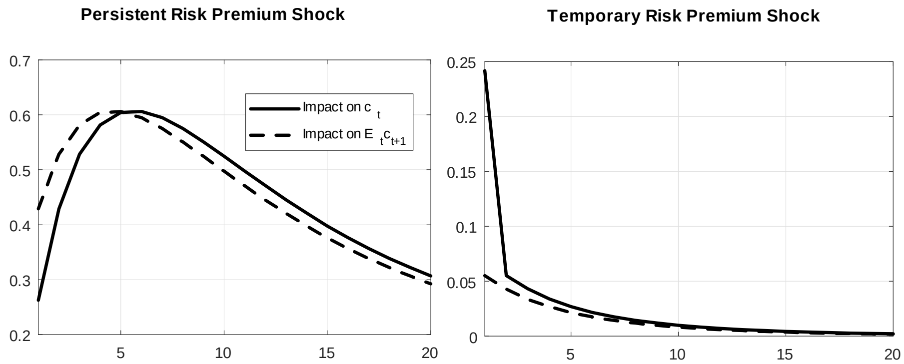

Figure 2: Impulse Response Functions for consumption and consumption expectations

shocks, $RE_{\{all\}}$, does either as well as the survey or better than the survey. In particular, the RE model is able to beat the survey in forecasting consumption growth. Additionally, Diebold-Mariano tests performed for up to 5-q ahead horizon forecasts (not reported in Table 2) show that the model predictions are not significantly different from the SPF-forecasts. Such a result implies that using information from the nowcasts is sufficient to capture the most informative content from the surveys.

Therefore, our results indicate that the SPF-surveys can be efficiently integrated within the model-consistent FIRE setup. The $RE_{\{all\}}$ version of the model, which separately identifies the i.i.d. and persistent components of the shocks, is able to process the survey information most efficiently and achieve a significant gain in terms of the fit and forecasting performance.

2.3 SPF forecasts: re-specification of fundamental shocks versus sentiment

To check the robustness of our set up with fundamental shock re-specification, we also examine an alternative approach used in the literature to explore the content of the survey data. In particular, we follow Milani (2017), who introduces exogenous sentiment shocks into the model with AL to allow for deviations from FIRE and for non-fundamental expectation shifts meant to capture the waves of optimism and pessimism of SPF agents. The approach implies that the forward-looking variables in the model are composed of the expectation variable itself and the exogenous sentiment shock component, which follows an autoregressive process. As a result, the framework allows for a more flexible dynamics of expectations in line with the observed forecasts. More specifically, we integrate the sentiment shocks in the following way:

$$\begin{aligned} dx_{f0,t} &= (E_t^S x_{t+1} - x_t) + \overline{\gamma}_{xf0} + e_t^{xf0} \\ E_t^S x_{t+1} &= E_t x_{t+1} + e_t^{xS} \\ e_t^{xS} &= \rho_S e_{t-1}^{xS} + \upsilon_t^{xS} \end{aligned} \quad (6)$$

where e_t^{xS} is the sentiment shock process and $E_t^S x_{t+1}$ is the expectation term, which substitutes forward variables in the model equations. Table 5 shows the estimation results for the specifications with sentiment.

Table 5. Models with sentiment shocks: comparison of the model fit and predictive power

Model	1-q RMSFE				*Log(Marg. Lik)*
	Inflation	Consumption	Investment	GDP	
RE^S	1.01	1.18***	1.06	1.22***	-477.88
$RE^S_{\{\mu\}}$	1.00	1.20***	1.05	1.22***	-488.96
$RE^S_{\{all\}}$	1.00	0.96	1.00	1.00	-388.17

Note: Models are evaluated over the sample period 1981Q2 - 2019Q2 using the first four observations as a presample. All models are estimated on the same dataset which includes 17 observable variables: real-time dataset and 4 SPF variables for inflation, growth rates of consumption, investment and output. Model notations: RE^S is the model with sentiment shocks for inflation, consumption and investment; $RE^S_{\{\mu\}}$ is the model with sentiment shocks and with the re-specification of price and wage markup shocks; $RE^S_{\{all\}}$ denotes the model which incorporates the sentiment shocks in addition to the re-specification of all the fundamental shocks. RMSFEs are reported in terms of the ratios of the RMSFE for the DSGE model to the corresponding RMSFE for the SPF nowcast. ***, **, and * denote the rejection of the null of the DM test, stating that the RMSFEs from the SPF and DSGE do not differ significantly, at 1%, 5%, and 10% significance level, respectively. Therefore, values above 1 marked with stars indicate that SPF nowcasts are statistically better than the DSGE model predictions.

We find that the models RE^S and $RE^S_{\{\mu\}}$, which assume standard real fundamental shocks but allow for the exogenous sentiment shocks, fit data much better compared to the models RE and $RE_{\{\mu\}}$ (shown in Table 4), which do not assume any exogenous waves of optimism and pessimism.[14] We can therefore confirm the result of Milani, who finds an important role of sentiment shocks, and, more specifically, investment sentiment shock, in explaining the observed expectations and real-time data. At the same time, the $RE_{\{all\}}$ model, which assumes the 2-component structure of all the real shocks, significantly outperforms the model with sentiment RE^S in terms of the fit and predictive power. Moreover, the results presented in Tables 4 and 5 also indicate that once we introduce the distinction between the transitory and persistent shock components, we no longer need to rely on the exogenous sentiment to fit the survey data. In the $RE^S_{\{all\}}$ model, estimated sentiment shocks are small and do not explain a significant portion of the business cycle.[15]

[14] Table A1 in the Appendix presents the details on the estimated parameters.

[15] We obtain a similar result in spirit and confirm the very limited role of sentiment shocks also in models with alternative, non-rational belief specifications, as illustrated later in Table 7.

2.4 Remaining issues with RE-hypothesis

In this section, we discuss the remaining problematic aspects of the framework based on the full rationality assumption. We focus on three main weaknesses of the RE-model: inability to deal with the shifts in macroeconomic trends observed in the data, stability of the transmission mechanism, and predictability of the forecast errors. We will therefore motivate the shift towards models based on alternative belief specifications. In the subsequent sections, we illustrate how our models with non-rational beliefs can overcome the limitations of the models based on RE discussed here.

2.4.1 Macroeconomic trend shifts

Modern macro data is generally characterised by time varying trends, which may display rather pronounced and long-lasting deviations from the mean. Such evidence exists for different variables and datasets, including our Real-time and SPF data. In particular, our sample contains the periods of persistently high and low productivity and macroeconomic growth, several deep recessions followed by slow recoveries (for example, the episode of the Great Financial Crisis (GFC)) as well as the structural break in the evolution of risk-free and natural rates in the post-financial crisis period. From the modelling perspective, such features of the actual data could be addressed by assuming the long run stochastic trends or, alternatively, by modelling the time-varying transmission mechanism.

In contrast to the observed evidence, models based on the full rationality assumption and fixed coefficients imply stabilizing expectations around a constant steady state and therefore abstract from the time varying trends and shifts in the long-run relations. The inability of the model to reproduce the dynamic nature of data will normally be reflected in the systematic pattern of the forecast errors. The first row of Figure 3 presents evidence of observed trend shifts in the mean of the first release of investment ($dinve_{r1,t}$) and consumption growth ($dc_{r1,t}$) as well as the corresponding SPF variables across 3 subperiods[16]. The figure also displays the co-movement between the trend changes in the data and in the RE-model longer-term forecast errors. In particular, both tend to increase in the second sub-period (1992-2006), which is marked by a persistent upward adjustment of the macroeconomic trends reflecting the acceleration of productivity growth, as considerable evidence suggests. Systematically positive RE-model forecast errors in period 2 indicate that the model, on average, underpredicts the realizations of consumption and investment. A similar in spirit conclusion can be seen from the second row of the Figure 3, which shows that declining interest rate trend (in particular in the post GFC period) is accompanied with greater tendency to over predict the interest rate dynamics. Therefore, the RE model struggles to adjust the forecasts to account for the long-term developments in the data.

[16] We define sub-periods on the basis of the visual inspection of developments in the data as well as complementary analysis of possible structural breaks. The subperiods are defined as follows: 1982q2-1992q3; 1992q4-2006q4; 2007q1-2019q2.

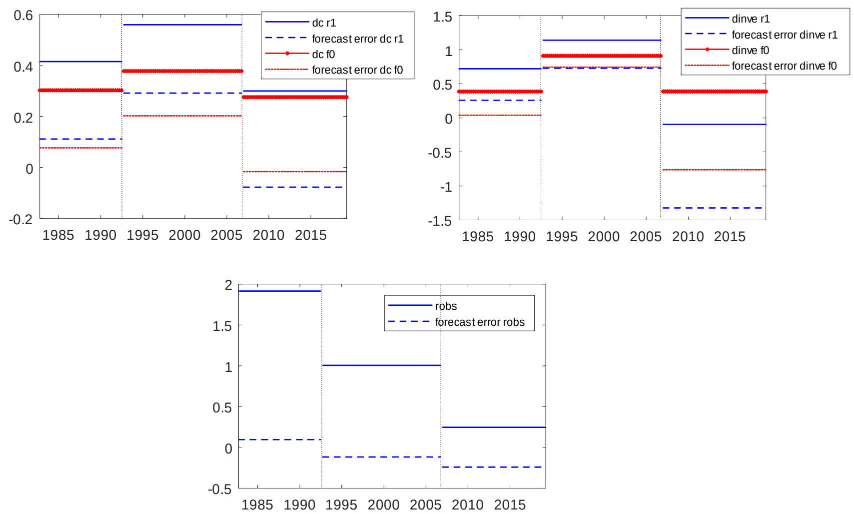

Figure 3: Shifts in the mean of observed macroeconomic variables and RE model 5-q forecast errors

2.4.2 Stability of the transmission mechanism

Another important drawback of the RE models with fixed coefficients is that they fail to produce any time-variation in the transmission mechanism of shocks.

Extensive empirical evidence based on laboratory experiments and SPF data analysis indicates that agents update their perceptions and forecasts in response to new information and pronounced structural changes (Manzan, 2021; Bauer et al., 2022). The intensity and efficiency of updating can vary over the business cycle affecting the sensitivity of expectations to shocks as shown in CG (2015). In addition, there is evidence of significant heterogeneity in the updating behavior of forecasters. The RE model is not able to reproduce changing and heterogeneous beliefs and to generate time-varying responses to shocks.

The stability of the RE model coefficients is also not in line with the evidence presented in Lindé-Smets-Wouters (2016), who document the non-Gaussian nature and the heteroscedasticity/garch dynamics in forecast errors and structural innovations of estimated DSGE models. Compared to the original SW-model that was estimated on final data, the use of real-time data substantially worsens all these test-statistics on the residuals. In particular, later revisions in the data have an important and systematic impact (see also Orphanides 2001; Jacobs-Van Norden 2011; Bognanni 2016). A large literature in finance and macro suggests that at least part of this time variation in model volatility and the transmission mechanism can be explained by changes in beliefs and/or sentiment (Milani 2017; Bauer et al., 2022).

2.4.3 Predictability of model-based forecasts

Previously, we illustrated the ability of the RE model to fit the SPF data very well and to properly exploit the useful content from the surveys. At the same time, we emphasized the underreaction problem documented in multiple studies on forecast data and also confirmed for our investment SPF. In this section, we illustrate that the RE model while observing and incorporating the timely information from the SPF, also inherits inefficiencies contained in the surveys due to the model-consistency of expectations imposed by the full rationality assumption. To illustrate this property, we apply the predictability test as in CG (2015) to the RE-model forecasts.[17] In particular, we run the predictability regression (1) based on the best-performing RE-model ($RE_{\{all\}}$) forecasts. Table 6 shows the results of the test for the RE-model predictions. The significantly positive value for b-coefficient indicates that the RE model augmented with survey data underestimates the investment realizations just like the SPF-nowcasts. This happens because the model is trying to fit the investment expectations observables, which are clearly not consistent with FIRE, and to reconcile them with the model-based investment predication, even to the detriment of fitting investment itself.

Table 6: Predictability test: RE-forecasts

	Inflation	Consumption	Investment	GDP
b-coefficient	0.04	0.18	**0.72**	0.14
st.error	0.22	0.14	0.17	0.11
p-value	0.854	0.233	0.000	0.191

The predictability problem detected in our RE-model forecasts can also arise due to a purely statistical reason, such as a small sample size. In addition, as discussed in the previous section, ignoring the time-varying long-term trends of the macroeconomic data can translate into a non-zero mean of the structural innovations leading to a certain degree of predictability in RE-model forecasts. In this regard, our results are in line with the evidence shown in Hajdini and Kurmann (2022), who illustrate that, in a model with Markov regime shifts, the predictability of forecasts can be consistent with the FIRE hypothesis.[18]

[17] Recent literature uses the predictability regression mostly to verify the validity of the FIRE hypothesis. In this exercise, the objective is rather to explore the properties of the forecast errors in a model that assumes model-consistent, rational expectations and utilizes the SPF data suffering from the under-reaction problem.

[18] They claim that the predictability of ex-post forecast errors cannot be considered as absolute evidence against FIRE. As a result, the predictability regression might not be a suitable technique to test alternative theories of expectation formation.

3 Models with non-rational expectations and adaptive learning

In this section, we introduce the departure from the complete rationality assumption to examine whether alternative expectations hypotheses allow for a more efficient integration of surveys in the macro models. We also want to evaluate to which extent models with more flexible belief specifications can address the limitations of the RE models outlined above.

We employ the AL approach to modelling the expectations as in Evans and Honkapohja (2001) and Milani (2007). It is assumed that agents know the structure but they are uncertain about the parameters of the model. To learn the parameters, they formulate linear econometric models based on their economic perceptions and re-estimate these models as soon as new information arrives. Following Slobodyan and Wouters (*SlW*) (2012), we assume that forecasting models are re-estimated every period on the basis of the Kalman filter learning algorithm.

SlW (2012) assume that agents' forecasts are based on very small forecasting models (perceived law of motion – PLM), in particular on a model in which the expected value of a forward-looking variable depends on a constant and two lags of this variable. With the standard set of observables, such beliefs specification efficiently captured the information available to agents and produced a good model fit.[19] In this paper, we show that such a simple structure of beliefs is no longer sufficient to explore properly the additional information content available in real surveys. In line with professional forecasters who process timely information available to them, our learning agents are allowed to incorporate the signals from surveys into their forecasting models. In particular, in our RB AL specification, we assume that, in addition to the lagged values of the forwards, agents also observe the innovations of i.i.d and persistent components of the structural shocks.[20]

In general form, agents use the following belief models (PLMs) to form expectations about the forward-looking variables:

$$y_t^f = \beta_{t-1}' X_{t-1} + u_t, \tag{7}$$

where X_{t-1}' is the information set which can include the constant term, lags of model's endogenous state variables and structural shocks; errors u_t, which depend on a linear combination of the true model inovations ϵ_t, are normally distributed with zero mean and variance-covariance matrix Σ. β_{t-1} is the vector of the belief coefficients updated with the Kalman Filter[21]. It is assumed that all the belief coefficients follow $AR(1)$ process with autocorrelation ρ. Agents

[19] AR(1) and AR(2) belief specification became relatively common in the learning literature and has been applied in a number of studies: Hommes (2014), Rychalovska (2016).

[20] Similar to *RSW* (2023a), who expand the inflation and wage PLMs to incorporate the markup shock components in addition to autoregressive terms.

[21] We call this belief updating mechanism a secondary Kalman filter. While the primary Kalman filter denotes the process used for the likelihood evaluation.

process the information from the surveys and other macro data, summarized in the information set X, in order to formulate and update their beliefs about the impact of the endogenous states as well as the persistent and transitory shock components on the predicted variables. For example, in some time periods, agents may perceive the persistent component of the shock as the more important driver of the real economic activity relative to the "true" impact of this disturbance. As a result, agents' beliefs can deviate from the model-consistent expectations and generate waves of excessive pessimism or optimism in expectations over the business cycle. Our approach based on time varying beliefs is therefore different from the setup of Milani, in which large deviations in expectations are mostly driven by the exogenous disturbance.

The predictions formed by (7) then replace the forward-looking variables in the linearized structural model representation to produce the Actual Law of Motion (ALM) given by:

$$y_t = \mu_t + T_t y_{t-1} + R_t \epsilon_t. \tag{8}$$

Therefore, the model transmission mechanism given by μ_t , T_t and R_t becomes time varying due to the evolution of agents' beliefs β_t. The true model innovations ϵ_t are normally distributed with zero mean and variance-covariance matrix Q. For more details on the specification of the AL mechanism and the Kalman filter updating see *SlW* (2012).

We consider three alternative specifications of the belief models.

3.1 Minimum State Variable beliefs with flexible constant updating

Minimum State Variable (MSV) beliefs imply that agents use a complete information set, which includes the same endogenous state variables and shocks that appear in the RE solution. The beliefs are initialized around the REE-solution, as in *SlW* (2012). The expectation-formation process therefore deviates only marginally from the RE setup.

We extend the standard approach and propose a somewhat more flexible MSV specification. In particular, we allow for the independent processes driving the time evolution of the constant and the rest of the belief coefficients corresponding to the state variables and shocks. With such a belief specification, we permit the constant to be updated more intensively to capture the potentially large shifts in the expected means of the forward-looking variables. Such a specification should improve the model's capacity to deal with the issue of the trend breaks and persistent deviations from the steady state values observed in the data. The PLM process takes the following form:

$$y_t^f = \beta_{t-1}^{const} + \beta_{t-1}' \begin{bmatrix} y_{t-1}^S \\ w_{t-1} \end{bmatrix} + u_t, \tag{9}$$

where y^S denotes the vector of state variables and w denotes the vector of shocks; β^{const} denotes the constant belief coefficient. As mentioned above, we introduce the separate priors on the persistence and the variance of β^{const} to allow for extra flexibility and independence in the evolution of this belief coefficient. As a result, the updating process for constants can become more responsive to the systematic forecast errors.

The estimation results of the DSGE model with MSV belief process (9) presented in Appendix A2 indicate an important role of the flexible constant. Given that other beliefs remain rather stable[22], the constant becomes the most important source of time-variation in the MSV setup. Table A2 indicates that this time-variation is substantial. In particular, autocorrelation of the constant is close to unity ($\rho^{AL}_{const} = 0.9937$ in the posterior mean). This implies that the evolution of the constant is close to the random walk process and is able to capture long-lasting trends of the variables. In addition, the parameter driving the volatility of the constant belief coefficient is equal to 0.75[23], which suggests more variation in the constant relative to other belief coefficients. Our later analysis indicates that the more flexible dynamics of the constant have important implications for developments in expectations and fitted realized variables. Table A2 demonstrates that the structural parameters of the MSV model with flexible constant remain very similar to the model based on the full rationality assumption.

3.2 Restricted beliefs model with flexible constant updating

A Restricted beliefs model with flexible constant (RB) implies that the PLM information set is limited relative to MSV. Our PLM specification is presented by the following general form:

$$y^f_t = \beta^{const}_{t-1} + \beta'_{t-1}\begin{bmatrix} y^f_{t-1}, y^f_{t-2}, y^f_{t-3} \\ \varepsilon^{w_{ar}}_{t-1}, \varepsilon^{w_{iid}}_{t-1} \end{bmatrix} + u_t, \tag{10}$$

where $\varepsilon^{w_{ar}}$ and $\varepsilon^{w_{iid}}$ denote the innovations of the persistent and transitory shock components. Thus, we assume that each forward-looking variable is predicted by the time varying constant with independent updating, three autoregressive terms as well as the innovations of the transitory and persistent components of the selected structural shocks. We explore the variance decomposition from the MSV model to obtain a broad idea about the importance of the structural shocks for the dynamics of real macro variables. We then formulate the forecasting models that incorporate the most important shocks for each variable.[24] Therefore, our forecasting models exhibit a sufficiently rich structure,

[22] We do not observe significant updating of the majority of the belief coefficients, except for the constant. This can be due to correlations across multiple variables in the PLM, which induces a multicollinearity problem in the updating process. Therefore, the problem is comparable to large TVP-models e.g. Koop et al. (2013), Chan et al. (2020).

[23] In the standard setup we fix this value at 1.

[24] In particular, we find that the risk premium shock naturally explains a significant portion of variation of the consumption expenditure and the price of capital but is also important for investment, hours and rental rate. For example, our consumption PLM includes 3 lags of

which resembles the MSV solution to a significant extent but remains more parsimonious in order to facilitate the estimation of the time varying transmission mechanism. Similar to the MSV setup, the initialisation of the beliefs in the RB PLM model is based on the parameters implied by the REE-dynamics.

The estimation results of the DSGE model with the RB belief process (9) are presented in Appendix A2. The results indicate that the RB model shows meaningful time variation of the belief coefficients (the ρ^{AL} parameter is rather high at 0.87). The constant is not particularly active in this specification (autocorrelation of the constant ρ^{AL}_{const} is equal to 0.37). It implies that the model does not exploit the constant to generate additional time variation. Regarding the estimated structural parameters and shock processes, we observe that the RB model has 3 times smaller estimated measurement error in the investment SPF equation. In addition, we observe a dramatic reduction in the standard error of the persistent investment technology shock. It implies that the RB model is more successful in fitting the realized investment as well as investment expectations data. The evidence regarding the direction of the change (relative to RE and MSV) in parameters of structural rigidities is rather mixed, which is in line with the AL literature.

3.3 Heterogeneous beliefs model with flexible constant updating

Heterogeneous beliefs (HB) model with flexible constant (HB) combines the MSV and RB specifications. More specifically, instead of assuming homogenous expectations with regular updating of the belief model (PLM), we also consider a setup in which agents have the possibility to use a combination of the alternative belief models based on their past forecasting performance:

$$E_t^{HB} y_{t+1}^f = \omega_t^{MSV} E_t^{MSV} y_{t+1}^f + \omega_t^{RB} E_t^{RB} y_{t+1}^f, \tag{11}$$

with the weight ω_t^i on each PLM $i = \{MSV, RB\}$ evolving as a function of the past belief forecast errors:

$$\omega_t^i = \frac{\exp\left(-\zeta m_t^i\right)}{\sum_i \exp\left(-\zeta m_t^i\right)} \quad with \quad m_t^i = \log\left(\prod_j Sigma_{j,t}^i\right), \tag{12}$$

where $Sigma_j^i$ denotes the exponentially smoothed empirical variance of forecast errors (for variable j, belief model i), which follows the process:

$$Sigma_{j,t}^i = \theta Sigma_{j,t}^i + (1-\theta)\left(\varepsilon_{j,t}^i\right)^2. \tag{13}$$

consumption as well as persistent and iid components of the risk premium shock. Investment dynamics is driven, to a significant extent, by investment specific technology shock. In addition, investment, which is a highly volatile and persistent variable, can be better described by more complex processes. Specifically, we also incorporate the lags of persistent and iid shock components into the investment forecasting rule. Government spending shock contributes to explaining the volatility of hours. Hence, the components of the government spending shock are included in the PLM for hours worked.

Further, $\varepsilon_{j,t}^{i}$ is the current actual forecast error, θ is the parameter measuring the memory length and ζ regulates the sensitivity of the PLM weight to the fitness measure m_t^i (see Brock and Hommes (1997) and De Grauwe (2012) for similar discrete choice model setups). The lower value of the parameter θ implies greater importance of the more recent forecast errors and smaller dependence on the past. Higher value of ζ corresponds to higher sensitivity of weights to difference in fitness. We estimate these two important coefficients, which govern the evolution of the weight, to allow the data to reveal the degree of heterogeneity in agents' expectations.

The detailed estimation results of the DSGE model with HB are presented in Appendix Table A3. We can see that the parameter vector combines the features of both MSV and RB models. The parameters governing the evolution of the weight are of particular interest. The weights are rather sensitive to the fitness measure ($\zeta = 4.77$), which is consistent with the volatile dynamics of the RB weight presented in Figure 10 in the Appendix. When it comes to the parameters that determine the intensity of the updating, the constant in MSV PLM remains extremely active, with the ρ_{const}^{AL} parameter estimated at the level close to unity[25]. The time variation in RB PLM is similar to the level estimated in the pure RB setup.

The model with combined beliefs and time-varying weight coefficient could be interpreted as agents changing their information set and degree of data processing over time. Thus, our HB specification is consistent with the concept of time variation and cycle-dependence of information rigidity described in CG (2015). In good and relatively stable times, when people have a good understanding of economic processes, they tend to become less attentive to new information and shift towards simpler forecasting rules. The dynamics of the weight on RB PLM, shown in Figure 10, are broadly consistent with this argument: the coefficient tends to increase and attain the highest values in between the recessions. Thus, for example, during the period of expansion between 2004-2008, the ω_t^{RB} stays close to 1. Later during the recession and afterwards, the weight on RB declines, indicating that people tend to process more information and switch to MSV. The average estimated weight for RB PLM is 0.34.[26]

The HB specification, which combines alternative belief models, allows for an additional flexibility in fitting the observed agents' expectations. At the same time, the setup implies that two different PLMs are restricted to operate under the same parameter vector. These parameters must be robust to produce competitive forecasts by each of the two specifications with potentially different transmission mechanisms of the shocks. As a result, the HB specification might not be able to fully explore the advantages of each belief specification.[27]

[25] We fix this parameter at 1 in MCMC to speed up covergence and facilitate identification of the parameters governing the weight updating process. For the same reason, we impose the equality between the ρ_{const}^{AL} and ρ^{AL} for RB PLM.

[26] This fraction is also very stable in simulation experiments (not reported in this paper), which indicate that RB gets more weight after repeated innovations in the same direction.

[27] Thus, the performance of the pure RB model and the RB PLM in a combined belief model may differ due to the differences in the parameter vector, at which these models are evaluated.

4 Model comparison results

Observing the SPF nowcast data adds discipline to the belief specification and the updating process. By estimating the models with different expectation formation mechanisms on the combined real-time and survey dataset, we allow the data to tell which alternative belief model provides a better description of the observables. In this section, we present the empirical analysis of the overall performance of the models with alternative beliefs. We compare the models' fit and forecasting performance. In addition, we evaluate the performance of the models on the extended sample which includes the Covid-19 period.

4.1 Comparison of the model fit

Table 7 presents the marginal likelihood for different models estimated on an augmented dataset of real-time data and SPF-nowcasts:

Table 7. Models with alternative beliefs: comparison of the model fit and predictive power

Model	1-q RMSFE				*Log(Marg. Lik)*
	Inflation	Consumption	Investment	GDP	
$RE_{\{all\}}$	0.21	0.40	1.46	0.33	-385.07
MSV	1.00	1.00	1.00	1.02	-381.38
RB	0.99	1.01	0.95*	1.02	-351.59
HB	1.01	1.01	0.89	0.99	-355.09
RB^S	1.00	1.01	0.95*	1.03	-359.96

Note: Models are evaluated over the sample period 1981Q2 - 2019Q2 using the first four observations as a presample. All models are estimated on the same dataset which includes 17 observable variables: real-time dataset and 4 SPF variables for inflation, and growth rates of consumption, investment and output. Model notations: $RE_{\{all\}}$ denotes the RE model with re-specified fundamental shocks; MSV denotes the MSV model with flexible constant; RB is the Restricted Belief model with flexible constant; HB is the model with Heterogeneous Beliefs and flexible constant; RB^S is the model with Restricted Beliefs, flexible constant and sentiment shocks. For RE-model forecasts, the RMSFEs are reported in levels. For models with non-rational beliefs, we report the ratios of the RMSFE produced by these models to the RMSFE of the model with RE. ***, **, and * denote the rejection of the null of the DM test, stating that the squared forecasts errors from non-RE and RE models do not differ significantly, at 1%, 5%, and 10% significance level, respectively. Therefore, values below 1 marked with star(s) indicate that forecasts produced by non-RE models are statistically better than the RE-model's predictions.

Table 7 illustrates that models with alternative belief specifications fit the data much better compared to the model that relies on the FIRE. RB and HB specifications, which assume a more severe deviation from the RE hypothesis and allow for both the limited information set and time varying transmission mechanism, outperform in this model comparison exercise. The model with MSV beliefs and flexible constant updating shows a moderate gain relative to the model with RE. Table 7 also illustrates that sentiment shocks do not add any explanatory power relative to the model with the RB specification.

The 1-q forecast comparison confirms the exceptionally good predictive power of the RE model. Additional gains in terms of the point forecasts produced by non-rational belief models are modest and mainly concentrated in investment.[28]

To gain further insights into the ability of different models to explain the specific episodes of the observed data, Figure 4 presents the (cumulative) period-by-period likelihood difference for three alternative belief models relative to the RE-model likelihood (all evaluated at the posterior mode) as well as the component of the likelihood difference function linked to forecast precision denoted as $'\Delta v'$ (dashed line). The second component of the likelihood function, which explains changes in the fit due to the contribution of the second moments, can be evaluated as the difference between the total likelihood change (solid line) and the $'\Delta v'$ term. [29]

Figure 4 illustrates that the gain from MSV beliefs with flexible constants relative to the fixed transmission mechanism implied by RE model is accumulating gradually over time, with the strongest improvement realized in the second half of the 1990s and in the 2000s, which corresponds to the periods of pronounced gradual trend changes in economic variables.[30] The likelihood decomposition results shown in Figure 4 indicate that such a superior performance of the model with MSV beliefs and flexible constant was mostly due to more precise point forecasts that reduce the systematic forecast errors. In particular, we observe the biggest improvement in the point forecast error likelihood component ($'\Delta v'$) between 1995 and 2008. At the same time, the forecast precision deteriorates around the GFC indicating that the mechanism with the flexible constant updating, which does good job in capturing the long-term gradual trend shifts, is not sufficient to capture the abrupt changes seen in the crisis period. The second moments bring a rather moderate gain, which is realized mainly in the second part of the sample.

The RB specification shows a sustained improvement over RE in the 1980s and the middle of the 1990s and performs particularly well during the periods of increased volatility and structural changes in the second part of the sample. In particular, the model is exceptionally successful during and after the GFC. Figure 4 highlights that RB and RE models show overall very similar point forecasting performance. The RB model is able to produce better forecasts in the 1980s and during the GFC but does worse in the middle of the 2000s. As a result, the improved fit of the RB model originates mainly from the realistic time-variation of the predictive density distribution. In particular, the gain from the second moments is accumulated gradually over time, with the most pronounced increase around the GFC. As shown in the Appendix[31], such an

[28] This result is consistent with our likelihood decomposition analysis, which is shown further in this section and which emphasizes the role of the second moments in the improved fit of the RB model.

[29] See a more detailed discussion on the likelihood decomposition in the Appendix. More specifically, dashed lines on the figure show the "forecast precision " term presented in the second line of equation (16).

[30] More specifically, rapid pace of developments in information technology fuelled a temporary surge in U.S. productivity growth, which slowed down before the Great Recession.

[31] See equation (16).

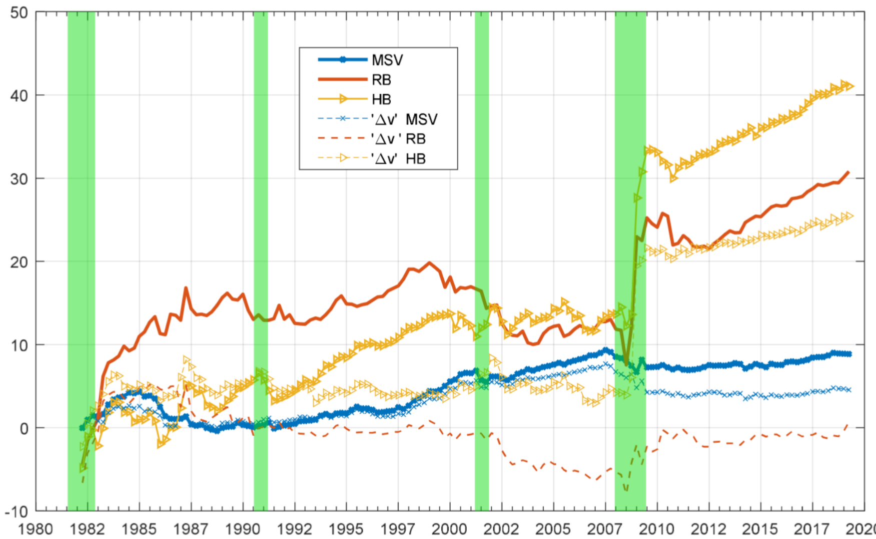

Figure 4: Cumulative difference in likelihood and likelihood decomposition

improvement may originate from the lower average volatility of the RB model as well as from the temporary increased volatility in times of particularly large forecast errors.[32]

Finally, the HB specification, which benefits from additional flexibility due to shifts between the alternative belief models, combines the advantages from the two setups. In particular, the HB model explores the potential of the model with flexible constant updating to capture shifting economic trends in the first half of the sample while at the same time preserving the ability of the RB model to remain efficient in the highly volatile environment such as the GFC.Figure 4 demonstrates that the HB specification dominates the RE model due to the combination of the systematically better forecast predictions and also due to the time variation of the density distribution. The most pronounced gain in the likelihood is realized around the GFC.

4.2 Comparison of the forecasting performance

In this section, we further explore the ability of the models with alternative expectation formation mechanisms to explain and predict the business cycle. In particular, we compare the longer-term forecasting performance of the models versus SPF on the whole sample and also focusing on the last decade covering the episode of the GFC. Another part of our analysis emphasizes the ability

[32]Examination of the evolution of the model volatility needed to verify the validity of this arguments is presented in section 5.2.1.

of the models to predict the specific states of the business cycle. Finally, we evaluate the performance of the models with alternative beliefs on an extended sample, which includes the Covid period.

4.2.1 Longer-term forecast comparison: SPF versus structural models

Table 8 presents the 5-q ahead forecast comparison.[33] We report the RMSFE in levels for the SPF forecasts and RE-model predictions. In addition, we present the accuracy of non-rational belief models as ratios of the corresponding RMSFEs to the value of the RMSFE computed for the RE model. The numbers smaller than 1 indicate that the alternative belief model performs better compared to RE.

The results indicate that the RE model remains very competitive relative to SPF predictions also for longer horizons. In particular, model does better than professional forecasters in predicting inflation dynamics. The forecasts for other variables are not statistically significantly different from the SPF surveys up to 5-quarters ahead.[34] This result implies that observing the data on SPF nowcasts is sufficient to generate the model predictions consistent with the longer-term projections of professional forecasters. We do not need to observe the longer-term SPF forecasts to obtain this result. Moreover, Table 8 illustrates that non-rational belief models tend to improve their predictive power (relative to the RE model) at longer horizons.[35]

The evidence from the out-of-sample forecast comparison exercise, which is conducted on the sample 2008q1-2019q2, generally confirms the conclusions based on in-sample analysis. In particular, Table 8 (panel B) indicates that on the period that includes the GFC, the RB model does better than RE in predicting investment as well as interest rate dynamics.[36] For investment, the model outperforms the professional forecasters. In addition, the RB specification shows very good predictive power for other real variables, such as consumption and output, producing forecasts which are either better or as good as survey data.[37] Notable improvement in the longer-term forecasting ability is also observed for two other AL models. As we describe further in section 5.1, such a superior performance at the longer-term horizon is related to the ability of alternative belief models with a flexible constant to capture the trending behaviour of the variables.

[33] See full tables, which include 1q to 5-q ahead forecasts, in the Appendix.

[34] See table A5 in the Appendix, which shows the DM test statistics for the equality between the SPF and the structural model forecasts.

[35] For comparison with short-term predictions, see Table 7 or more detailed Table A4 in the Appendix.

[36] Table A6 in the Appendix indicates that it is true at both short-term and long-term horizons.

[37] Certain worsening of out-of-sample predictions for inflation for longer horizons relative to in-sample forecasts is related to the fact that the RB model tends to overpredict the declining inflation. However, the model still does as well as the survey. In fact, given that our forecast sample includes the episode of severe economic distress, it is rather natural that out-of-sample model performance deteriorates somewhat relative to in-sample predictions.

Table 8: Long-term forecast comparison

	Realized data					SPF data			
	π_{r1}	dy_{r1}	dc_{r1}	$dinve_{r1}$	$robs$	π_{f0}	dy_{f0}	dc_{f0}	$dinve_{f0}$
	A. in-sample: 1981q4-2019q2								
RMSFE 5q									
SPF	0.29	0.49	0.50	2.06					
RE	0.25	0.48	0.50	2.06	0.38	0.15	0.34	0.30	1.15
MSV	1.00	0.99	0.97	0.99	1.00	0.91	0.98	0.95	0.96
RB	0.99	1.01	0.99	0.96**	0.99	1.02	0.99	0.98	0.89**
HB	1.06**	0.97	0.97	0.95**	0.99	1.04	0.96	0.98	0.90
	B. out–of-sample: 2008q1-2019q2								
RMSFE 5q									
SPF	0.25	0.56	0.42	2.59					
RE	0.24	0.56	0.42	2.74	0.39	0.12	0.42	0.29	1.53
MSV	0.99	1.02*	1.00	0.98***	0.98	1.04	0.95*	0.99	0.94***
RB	1.06	0.95*	1.02	0.90***	0.81***	1.33**	0.89***	1.04	0.80***
HB	1.03	1.01	1.08	0.90***	0.94***	1.17	0.99	1.13***	0.88**

Note: For the SPF predictions of the first release of variables as well as RE-model forecasts, the RMSFEs are reported in levels. For models with non-rational beliefs, we report the ratios of the RMSFE produced by these models to the RMSFE of the model with RE. ***, **, and * denote the rejection of the null of the DM test, stating that the squared forecasts errors from non-RE and RE models do not differ significantly, at 1%, 5%, and 10% significance level, respectively. Therefore, values below 1 marked with star(s) indicate that forecasts produced by non-RE models are statistically better than the RE-model's predictions. For out-of-sample forecasting exercise, the models are re-estimated every quarter.

4.2.2 Forecasting performance over the business cycle

Despite the fact that, overall, alternative belief models do not produce superior short-term forecasts, the previous section demonstrated convincing evidence of improved performance of AL models (and the RB model in particular) in predicting the most volatile model variable, namely investment. In this section, we would like to better understand this result by comparing the predictive power of the surveys, RB and RE models during the specific states of the business cycle. In particular, we split the sample into 3 sub-periods, which capture different stages of the business cycle. The "recession" dates are defined in accordance with the NBER definition. The post-crisis "recovery" sub-period includes the first 8 quarters after the crisis. Finally, the "stabilization" phase includes the period when the economy is not in a recession or post-crisis recovery. Tables 9 and 10 summarize the results of the forecast comparison exercise.

The results shown in Table 9 indicate that the RB model does exceptionally well in predicting the recessions driven by financial factors, which were at the core of the economic downturns in the beginning of the 1990s and in 2008. In particular, the RB model shows superior performance in predicting the contraction of the real variables, such as investment and output, during these most pronounced recessions present in our sample and also does exceptionally well in capturing the post-crisis recovery for these variables. Inflation and interest rate dynamics is also better predicted by the RB model in the majority of the episodes. These results are in line with the out-of-sample forecasting analysis shown in Table 10, which also emphasizes strong dominance of the RB model in capturing declining interest rate in the post-crisis period. Moreover, Table 10, which in addition compares the out-of-sample forecasting performance of the structural models and SPF, reveals that surveys do very well in predicting the recession of 2008 as well as the post-crisis period. During the crisis, the RB model does better than the survey for inflation and investment. Both RE and RB models do worse than survey predicting output and consumption. Both RE and RB models do better than the survey predicting the recovery of real variables – output, consumption and investment. RE predicts inflation recovery better than SPF while RB does as well as the survey. Therefore, our results illustrate that forecasting advantage of the RB model is concentrated mainly around the crisis and post-crisis periods.

To understand this property better, we examine whether there is a systematic difference in one direction between RB and RE model forecasts during specific states of the business cycle. We calculate the mean of the difference between the RB forecast and the RE forecast over the sub-periods and refer to this discrepancy as the forecast "bias". Table 11 indicates that the RB model is able to generate predictions with stronger cyclical properties compared to the RE model. In particular, negative forecast bias for real variables, which is typically observed in recessions, illustrates that the RB model is able to produce systematically more negative forecasts relative to the RE model during the periods of the crisis. The positive forecast gap observed in recovery periods indicates that the RB model predicts a stronger growth in the post-crisis periods. The ability of the RB specification to generate stronger business cycle makes the model more successful in reproducing the pronounced economic downturns and expansions, particulary in investment.[38] Persistently negative forecast bias for nominal variables, such as inflation and interest rate, implies that RB model predicts, on average, lower values of these variables. This property can be particularly helpful in capturing the declining inflation trends and low interest rate environment observed in the period after the GFC.

The ability of the RB model to outperform around the crisis episodes can be explained by a combination of the appropriate information set used in forecasting and by the systematic updating of beliefs regarding the state of the economy. The RB model implies that agents include only the most relevant variables in their forecasting functions. If the selected variables are informative enough about the economic conditions, relevant shocks may be transmitted rapidly and the quality of the predictions can be rather good. However, during the periods of major disturbances when the state of the economy is changing fast, a restricted information set may result in a significant deterioration of the predictions relative to the forecasts made by a fully rational agent. Therefore, it is essential to complement the restricted information set with the proper updating, which may compensate for the shortage of information in the periods of crises and structural changes.[39]

[38] On another hand, stronger cyclicality of the RB model may occasionally result in excessive amplification of the cycle leading to over-pessimistic or over-optimistic predictions of the macroeconomic dynamics. This is why the RB model does not systematically dominate the RE model in terms of forecasting.

[39] The robustness check exercise presented in the Appendix illustrates an important role of updating. When we shut down updating, forecasts and fit deteriorate significantly. Updat-

Table 9: RB and RE-model forecast comparison over the subperiods

	Realized data					SPF data			
	π_{r1}	dy_{r1}	dc_{r1}	$dinve_{r1}$	$robs$	π_{f0}	dy_{f0}	dc_{f0}	$dinve_{f0}$
Recessions									
1981q4-1982q4	0.95	1.16	1.06	0.50	0.97	1.23	1.18	0.97	2.09
1990q3-1991q2	0.96	0.59	0.99	0.78	0.99	1.08	1.07	1.09	1.08
2001q2-2001q4	1.01	1.13	1.09	0.95	0.99	1.08	0.94	1.09	1.03
2008q2-2009q2	0.95*	0.96	1.10	0.90*	0.95	1.03	1.12	1.03	0.83
Recoveries									
1983q1-1984q4	1.13*	0.87	0.86	0.85*	0.96	1.19	1.01	0.98	1.07
1991q3-1993q2	0.93	0.92	0.89**	0.86	1.02	0.98	1.04	0.92	1.00
2002q1-2003q4	0.95	1.00	1.00	0.93	0.94	0.92	1.20	1.16	0.98
2009q3-2011q2	0.99	0.96	0.93	0.83	0.59***	1.19	1.10	0.99	0.77
Stabilizations									
1985q1-1990q2	1.04	1.11	1.00	1.02	1.02	1.06	1.00	0.91	0.94
1993q3-2001q1	0.97***	1.03	0.93	1.02	0.95	0.88*	0.90	1.09	0.87**
2004q1-2008q1	1.05**	1.10	1.14	0.94	1.01	1.04	1.01	0.94	1.02
2011q3-2019q2	0.99	1.10**	1.19	0.98	0.85	1.05	1.07	0.92*	1.03

Note: The statistics are based on in-sample predictions. We report the ratios of the 1q RMSFE produced by RB flex model to the 1q RMSFE of the model with RE. ***, **, and * denote the rejection of the null of the DM test, stating that the squared forecast errors from non-RE and RE models do not differ significantly, at the 1%, 5%, and 10% significance level, respectively. Therefore, values below 1 marked with star(s) indicate that forecasts produced by the RB model are statistically better than the RE-model predictions.

Table 10. Out-of-sample forecast comparison over the sub-periods

	Realized data					SPF data			
	π_{r1}	dy_{r1}	dc_{r1}	$dinve_{r1}$	$robs$	π_{f0}	dy_{f0}	dc_{f0}	$dinve_{f0}$
	Recession: 2008Q2-2009Q2								
SPF	0.46	0.29	0.69	3.80					
RE	0.45	0.36	0.77	3.95	0.16	0.21	0.64	0.53	1.81
RB	0.96	0.95	1.09	0.94*	0.96	1.08	1.08	1.08	0.91
	Recovery: 2009Q3-2011Q2								
SPF	0.17	0.38	0.39	1.69					
RE	0.16	0.37	0.36	1.68	0.11	0.06	0.24	0.20	0.47
RB	1.06	0.96	0.94	0.84	0.65***	1.37	1.00	1.15	0.83
	Stabilization: 2011Q3-2019Q2								
SPF	0.17	0.25	0.24	1.02					
RE	0.18	0.27	0.24	1.14	0.04	0.07	0.21	0.16	0.39
RB	1.02	1.15**	1.20**	0.96	0.84**	1.18	0.90**	1.07	0.88

Note: The statistics in Table 10 are based on out-of-sample predictions. For the SPF and RE-model forecasts, the 1q RMSFEs are reported in levels. For models with non-rational beliefs, we report the ratios of the RMSFE produced by these models to the RMSFE of the model with RE. ***, **, and * denote the rejection of the null of the DM test, stating that the forecasts from non-RE and RE models do not differ significantly, at the 1%, 5%, and 10% significance level, respectively. Therefore, values below 1 marked with star(s) indicate that forecasts produced by non-RE models are statistically better than the RE-model predictions.

Table 11: Mean forecast bias over the subperiods

	Realized data					SPF data			
	π_{r1}	dy_{r1}	dc_{r1}	$dinve_{r1}$	$robs$	π_{f0}	dy_{f0}	dc_{f0}	$dinve_{f0}$
Recessions									
1981q4-1982q4	-0.15*	-0.10	-0.13	-0.98	-0.03	-0.05	-0.09	-0.13	-0.78
1990q3-1991q2	0.00	-0.18***	-0.16***	-0.39*	0.00	0.01	0.03	-0.01	-0.29
2001q2-2001q4	0.01	-0.09	-0.06	-0.16	0.00	0.00	-0.01	-0.04	-0.32**
2008q2-2009q2	-0.03**	-0.04	0.00	-0.17	0.00	0.00	-0.09	-0.13*	-0.53*
Recoveries									
1983q1-1984q4	-0.07**	0.14	0.09	0.47*	-0.02**	-0.03**	0.02	0.07*	0.35**
1991q3-1993q2	-0.03***	0.05	0.03	0.15	0.01	-0.01**	0.05	0.04	0.18**
2002q1-2003q4	-0.02***	0.13***	0.14***	0.26**	0.00	-0.02***	0.06	0.04	0.05
2009q3-2011q2	-0.05***	0.09	0.03	0.46*	-0.03***	-0.02***	-0.08**	-0.06**	0.25

ing is crucial particularly in crisis periods because agents need to process information more intensively.

Note: The statistics are based on in-sample predictions. We report the mean of the difference between the forecasts produced by the RB flex model and by the model with RE (note that this is equivalent to the mean of the difference in the forecast errors produced by two models). ***, **, and * denote the rejection of the null of the DM test, stating that the forecast errors from non-RE and RE models do not differ significantly, at the 1%, 5%, and 10% significance level, respectively. Therefore, the negative value marked with star(s) indicates that forecasts produced by the RB model are statistically smaller than the RE-model predictions.

The results presented in this section illustrate very good capacity of the model with non-rational expectations to compete with survey data and the RE model in predicting the macroeconomic dynamics at both the short- and long-term forecast horizon. Non-rational belief models can generate more pronounced boom-bust dynamics in investment and, therefore, seem to be appropriate to describe the important features of the actual real business cycle marked with pronounced economic downturns and expansions. The model therefore can be less prone to under-predictions of macro variables, which was shown to be one of the limitations of the RE model. Observing and efficiently utilizing the timely survey data via the re-specification of the main structural shocks could be particularly helpful during such episodes as the GFC, when models normally struggle to generate accurate predictions.

4.2.3 Models with non-rational beliefs and the Covid recession

The Covid shock differed strongly from the other crises in recent history. The disturbance had an unusual nature and complex economic consequences. The pandemic caused an unprecedented economic fallout and forced governments to implement extraordinary measures and stabilization policies. In this subsection, we evaluate the ability of our models with alternative belief specifications to perform during the Covid recession and the post-crisis period. In our exercise, we keep the baseline model structure while adapting the statistical properties of the shock processes in order to generate a scenario consistent with a Covid episode. With our framework, based on the standard structural shocks and not particularly amended to characterize the pandemic, we do not aim at providing a comprehensive interpretation of the Covid recession and explaining its main drivers. Instead we are interested in demonstrating to what extent our models with alternative beliefs and time-varying transmission mechanism could remain operational and competitive with models based on RE given a shock of extraordinary size and nature.

As mentioned above, we do not augment the baseline model with additional "Covid" shocks. To capture the magnitude of the crisis, we propose to modify the structure of the fundamental shock processes by introducing heteroskedasticity around the Covid recession period.[40] More specifically, in both RE and

[40] Fast growing literature explores different approaches to macroeconomic modeling of the COVID-19 pandemic. One strand describes the pandemic within the standard framework based on the established structural shocks. In particular, Fornaro and Wolf (2020) consider a negative productivity shock as a potential force driving the covid recession. Other studies focus on the dominant role of aggregate demand shocks (for example, shock to utility of consumption

non-rational belief models, we allow for the variance of the structural shocks to increase during the period 2020q1-2021q1. Technically speaking, we introduce 5 separate (for each Covid period) scaling factors, which multiply the elements of the variance-covariance matrix of fundamental shocks in the primary Kalman filter. In particular, we multiply the variance-covariance matrix Q of fundamental shocks ϵ_t from equation (8) by the scaling factor $\exp(\gamma_{t=2020Q1:2021Q1})$. In addition, for models with non-rational beliefs, we modify the secondary Kalman Filter, which is responsible for the belief adjustment[41]. Specifically, we introduce the theoretically-consistent heteroscedasticity intervention with the same scaling factors to allow for the increased volatility of PLM forecast errors and perceived uncertainty. Thus, variance-covariance matrix Σ of the PLM forecast errors u_t from equation (7) is multiplied by the scaling factors $\exp(\gamma_{t=2020Q1:2021Q1})$. [42] We evaluate our models on the extended sample (up to 2022q2), fixing the parameters at the level estimated on the main pre-covid sample (1981q2-2019q2) and estimating only the 5 scaling factors γ_t.[43]

Table 12. Model comparison over the extended sample

Model	1-q RMSFE				*Log(Posterior)*
	Inflation	Consumption	Investment	GDP	
$SPF - nowcast$	0.62	1.43	2.39	1.18	
$RE_{\{all\}}$	0.64	1.42	2.36	1.19	-361.57
MSV	0.99	1.00	0.98	1.00	-345.70
RB	0.98	1.08**	1.88***	1.31***	-387.52
HB	0.99	1.00	0.98*	1.00	-334.53

Note: Models are evaluated over the sample period 1981Q2 - 2022Q2 using the first four observations as a presample. The structural parameters are identified over the baseline sample 1981q2-2019q2. The RMSFE statistics are based on in-sample predictions (2020q1-2022q2). For the SPF predictions of the first release of variables as well as RE-model forecasts, the RMSFEs are reported in levels. For models with non-rational beliefs, we report the ratios of the RMSFE produced by these models to the RMSFE of the model with RE. ***, **, and * denote the rejection of the null of the DM test, stating that the forecasts from non-RE and RE models do not differ significantly, at the 1%, 5%, and 10% significance level, respectively. Therefore, values below 1 marked with star(s) indicate that forecasts produced by non-RE models are statistically better than the RE-model predictions.

as in Faria-e-Castro, 2021) or negative and persistent labor supply shock, mimicking the impact of shutdown and lockdown policies (as in Guerrieri et al., 2022). Another strand of the Covid-19 modeling literature introduces additional shocks and model features, which help to account for the pandemic effects and analyse realistic policy options. Cardani et al. (2022) augment the canonical DSGE model with "lockdown shocks" ("forced savings" and labour hoarding) and liquidity constrained firms to capture various demand and supply effects of the covid recession and analyse the containment and stabilization policies.

[41] See *SlW* (2012) for a detailed technical description of belief updating in a secondary Kalman filter.

[42] Our results (not shown in this paper) indicate that the correction for heteroscedasticity greatly improves the fit of both RE and non-rational belief models. Our results also illustrate a prominent role of the demand-side shocks, such as the risk premium and investment technology shocks, in explaining the pandemic period in our model. This result is in line with the literature that emphasizes the important role of demand factors in driving the pandemic.

[43] Table A10 in the Appendix presents the estimation results.

Table 12 presents the comparison of our estimation and forecasting results, which indicate that the RB model does not work optimally on the extended sample. We explain this result by the fact that the AL mechanism, based on a rather limited information set, cannot describe well the abrupt dynamics and the complexity of the macroeconomic effects during the Covid crisis. Given that the pandemic shock was not in line with the historical covariance structure, the existing PLM based on a subset of endogenous variable and structural shocks becomes more severely mis-specified relative to "normal" periods and the performance of the model deteriorates.[44] At the same time, our results indicate that, despite distorted correlation structure and problematic identification of the Covid shocks in the PLM, the belief updating process itself continues to work normally. Compared to the baseline scenario, a version with heteroscedasticity intervention displays overall smaller variation of beliefs. The introduced adjustment of the shocks' standard deviations "informs" agents about a temporary exogenous increase of economic volatility. As a result, our models with alternative beliefs show lower volatility of beliefs because agents do not associate the observed dynamics with changed macroeconomic fundamentals. Table 12 also illustrates that MSV and HB belief models that rely on a broader information set continue to perform well. The updating of the flexible constant in the MSV model is not particularly useful during the outbreak of the Covid crisis, but still has an important role in capturing more gradual shifts in the macroeconomic trends in the pre- and post-covid periods. The HB model also shows overall superior performance relative to the RE setup, with the MSV regime natuarlly dominating during the Covid and post-Covid periods.[45]

The results of the forecast comparison exercise indicate that the quality of the SPF predictions for consumption and output during the Covid period deteriorated significantly relative to the GFC period.[46] In particular, SPF agents were greatly surprised in the beginning of the pandemic (2020Q1) due to the unprecedented nature and novelty of the shock. In addition, agents were rather pessimistic and underpredicted significantly the fast recovery process. However, professional forecasters managed to utilize the information advantage and produce excellent predictions for 2020q2, which was the peak of the contraction in output and consumption. This greatly contributes to the improvement of the predictive ability of the models. As we can see, the RE model benefits from the

[44] More specifically, our analysis illustrates that, at the peak of the Covid crisis, the AL mechanism generates an excessive amplification of the real variables driven by the structural shocks that the model explores to describe the Covid period. Due to insufficient flexibility in the shocks structure of the baseline PLM, the same shock is forced to drive the dynamics of variables very differently affected by the Covid disturbance (for example, consumption and investment).

[45] Our results indicate that the HB model accumulates most of its gain in the pre-covid period, showing deterioration after 2020q1. One could explain it by the fact that the parameter vector of the HB model is restricted to be the same for both MSV and RB regimes. This prevents the model from achieving the result identical to the one of the pure MSV setup. The evolution of the weight coefficient in the extended model is very similar to the dynamics shown in Figure 12.

[46] See Table 10 showing the SPF performance during the GFC.

timely survey information and produces predictions similar to the SPF.[47] Table 12 also illustrates that the RB model does not produce efficient predictions for real variables due to the mis-specification of the Covid shocks in the PLM process as described above. However variables that not directly affected by the Covid disturbance, such as inflation and inflation SPF, are still predicted rather well relative to RE[48].

The results of this section indicate that the RB model can perform well during periods of distress that have a more persistent nature and a rather standard origin as, for example, the GFC. At the same time, the model with restricted PLM is not informative enough to describe the Covid recession. MSV and HB models, which combine the benefits of observing the complete information set and the time variation, appear to represent more robust frameworks suitable to describe various economic episodes including the Covid period.

5 Do alternative belief models relax the limitations of the RE-Model? The role of time variation

In this section, we discuss how alternative belief models can relax the important limitations of the models based on the full rationality assumption described in the previous sections.

5.1 Macroeconomic trend shifts

As described in section III, models with non-rational expectations allow for the non-zero, time-varying long-term mean of the model variables. More specifically, our MSV, RB and HB setups assume the flexible updating of the constant belief coefficient β_t^{const} in PLM equations (9)-(10). As a result, the ALM constant, represented by μ_t in equation (8), varies over the business cycle and can potentially improve the ability of the model to capture dynamic trends in the observed data.

Naturally, the time-variation of the flexible constant is more pronounced in the MSV and HB models, where this parameter plays the central role in the belief updating process.[49] In these specifications, flexible constants are the most active with the corresponding ρ_{const}^{AL} parameter being close to unity (see

[47] Another important role of SPF forecasts in the context of the Covid-19 crisis can be related to the usefulness of survey data for identification of transitory and persistent components of the risk premium and investment technology shocks. These shocks can be categorized as demand shocks, which, according to various studies, played a prominent role in the beginning of the pandemic. Therefore, we believe that survey data, which improves identification of the Covid shocks, can be efficiently used for modelling and analyzing the pandemic crisis.

[48] See Table A11 in the Appendix which presents more detailed forecast comparison.

[49] Given that the MSV PLM dominates in HB model, constants evolve in a similar way in these two setups.

table A2). As a result, constants show gradual updating with very slow or no-mean-reversion.

Table 13 demonstrates the ability of adaptive learning to capture persistent trends and utilize them for superior predictive performance. The ALM constants for consumption and investment, in particular for the MSV setup, successfully pick up rising macroeconomic trends in periods 1 and 2, which in turn determines the longer-term forecast dominance of the models with alternative beliefs over the RE model.[50] At the same time, we can see that the slow and gradual adjustment of the flexible constant in the MSV setup is not sufficient to beat the RE model in describing the sudden decline of the business cycle in period 3, which includes the GFC. The RB model, which contains a more elaborated belief-updating mechanism, in addition to the flexible constant updating, generates superior projections for the variables such as investment, investment expectations and interest rate, which showed the most notable shifts in the trends after 2008. Therefore, models with imperfectly-rational beliefs contain the endogenous internal propagation mechanism that can improve the ability of the model to reproduce the time varying nature of the data. Updating of the constant is an important element of this process.

Table 13: Evolution of the mean of ALM constant and long-horizon forecast bias across subperiods

Subperiods	1982q2-1992q3		1992q4-2006q4		2007q1-2019q2	
	MSV	RB	MSV	RB	MSV	RB
	ALM constant					
dc_{r1}	0.23	0.34	0.28	0.35	0.20	0.33
$dinve_{r1}$	0.88	0.64	1.26	0.67	0.26	0.60
	Forecast bias					
dc_{r1}	-0.07	-0.05	-0.12	-0.12	0.06	0.09
$dinve_{r1}$	-0.1	0.04	-0.17	-0.20	0.03	-0.18
dc_{f0}	-0.03	-0.05	-0.09	-0.11	0.03	0.02
$dinve_{f0}$	0.05	0.00	-0.22	-0.19	0.01	-0.26
$robs$	-0.03	-0.01	-0.02	0.01	-0.01	-0.12

Note: ALM constant is represented by μ_t in equation (8). Forecast bias measures the difference between the mean of the MSV or RB-model 5-q forecast error and the mean of the RE-model 5-q forecast error. Negative numbers imply that, on average, MSV or RB model generates more precise 5-q ahead predictions.

5.2 Time variation of the transmission mechanism

In addition to the updating of the constant, revisions of other belief coefficients, such as the autoregressive terms and structural shock components present in the PLMs, reflect the time varying perceptions of agents about macroeconomic

[50] In the RB model, the constant is not particulary active. However, such a relative stability is compensated by adjustment of other belief coefficients in the PLM (autoregressive terms and shock components). As a result, the RB model also shows improved longer-term forecasting performance.

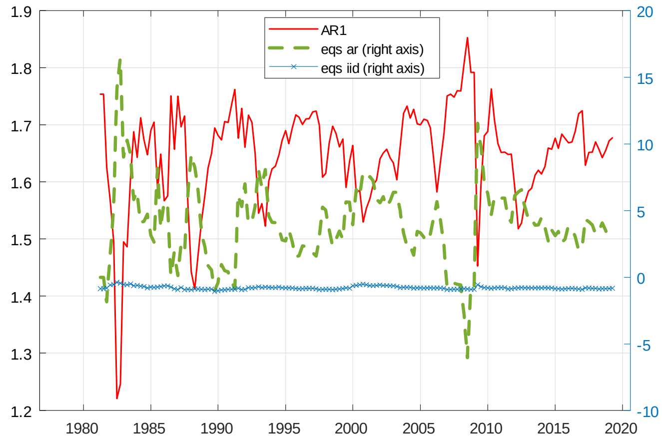

Figure 5: Investment belief coefficients in the RB model with a flexible constant

persistence. As a result, the responsiveness of the expectations and realized macroeconomic variables to shocks, driven by the evolution of agents' beliefs, will vary over the business cycle.

5.2.1 Time-varying IRFs

Figure 5 plots the evolution of the most important belief parameters β for investment PLM in the RB model: the first order autoregressive term and the innovations of the persistent and i.i.d. components of the investment technology shock are denoted as 'eqs ar' and 'eqs iid', respectively.

Figure 6 displays the time varying IRFs of investment to a persistent investment technology shock in RB and HB models. As discussed above, the evolution of beliefs determines the dynamic responses to shocks. In the HB model, the variation of the weight on RB PLM is an additional factor that also plays a role. As shown in Figure 6, in models with imperfectly rational beliefs, the responses can be amplified (even though they might be smaller on impact) and more persistent compared to the responses under RE (shown by a black solid line in the beginning of the sample). Therefore, the AL transmission mechanism can generate a stronger cumulative effect of shocks over the business cycle.

The overall effect of shocks on the endogenous variables is determined by the combined effect of the autoregressive coefficients and the beliefs about the contribution of the persistent and i.i.d. shock components. More specifically, an increase in the 'eqs par' or 'eqs iid' belief coefficient in the investment PLM implies that agents track the shock components more closely and expect that the shock may have a more pronounced immediate effect on investment. When this happens, the AR(1) belief coefficient tends to decline, accelerating the shock

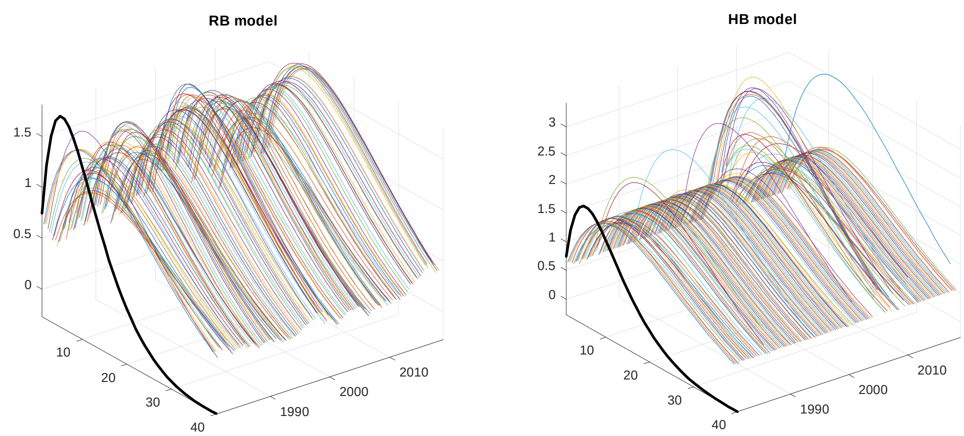

Figure 6: Time-varying Impulse Response Functions of investment to a persistent investment technology shock. Note: the black solid line shows the IRF under RE

propagation. On the other hand, an increase in the AR belief coefficients, and in particular AR(1) term that plays a dominant role, implies that the variable is seen as a relatively more persistent process. As a result, shocks may have somewhat delayed but longer-lasting effects and the IRFs will display more persistent dynamics during these periods, as shown in Figure 6.

5.2.2 Time-varying volatility under AL

As was emphasized in section 4.1, time varying volatility plays an important role in the performance of the non-rational belief models. Figure 7 presents the comparison of the implied volatilities generated by models with alternative belief specifications. We measure the implied volatility as ln|F|, where F is the variance-covariance matrix of one-step-ahead forecasts from the Kalman filter. The plot indicates that RB and HB models are on average less volatile compared to the RE and MSV models. The volatility generally follows a cyclical pattern and tends to increase around the periods of recessions, with the greatest surge observed during the GFC.[51] Note that the HB model generates the combined dynamics, which includes relatively stable periods and the moments of rather significant time variation in volatility. Such a pattern is related to the changing weight between the MSV and RB model in the HB specification. Naturally, more pronounced time variation in volatility is observed when the RB model receives a dominant weight.

[51] The examination of the variable-specific volatilities (not reported here) indicates that the overall model variation is driven to a significant extent by investment dynamics. This is not a surprising observation given that this time series has the highest volatility.

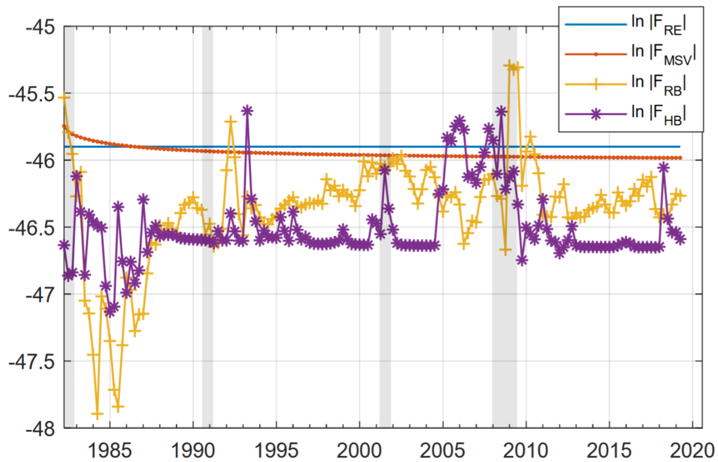

Figure 7: Implied volatility of forecast errors

Therefore, Figure 7 confirms the intuition discussed in section 4.1 that the likelihood improvement of the RB specification results mainly from the ability of the model to generate realistic cyclical evolution of volatility, which remains lower in normal times and rises during episodes of economic distress.

Furthermore, we would like to understand the causes of lower average RB model volatility, which could be explained by a different shock size and/or modified transmission mechanism. Table 14 presents the changes in the standard deviations of the main identified shocks (and their components) driving the real variables, across the three subperiods, which we also considered in the previous section.[52] The Table illustrates that the RE model explores more intensively the persistent components of the risk premium and especially the investment-specific technology shock to explain the investment dynamics, particularly during the period that includes the GFC. At the same time, the average volatility of these shock processes in the RB model does not increase on the last sub-period and remains generally lower in the whole sample. Also note that there is little difference in the volatility of the transitory component of these shocks in the RE and RB models. Therefore, we are able to illustrate an important property of the RB model: the modified transmission mechanism, based on imperfectly rational expectations and time varying beliefs, allows the model to better explain the persistent investment dynamics with smaller exogenous shocks, which in turn translates into lower average model volatility and an improved fit.

Table 14: Standard deviation of the persistent and transitory components of the main shocks: averages across the subperiods

[52] See also Table A2, which provides estimates of the standard deviation for innovations for the persistent and iid shock components.

Subperiods	1982q2-1992q3		1992q4-2006q4		2007q1-2019q2	
	RE	RB	RE	RB	RE	RB
	Persistent shock component					
Risk premium	0.058	0.061	0.056	0.044	0.061	0.043
Investment-specific technology	0.177	0.039	0.154	0.058	0.237	0.055
	Transitory shock component					
Risk premium	0.294	0.299	0.181	0.177	0.132	0.141
Investment-specific technology	0.668	0.638	0.444	0.441	0.519	0.543

5.3 Predictability of model-based forecasts

The previous section discussed the ability of the adaptive learning approach to overcome the well-known limitations of macro models based on the full rationality assumption. In this section, we focus on the aspect that is explicitly related to the presence of surveys in the dataset. More specifically, our earlier analysis revealed that professional forecasters inefficiently revise their investment predictions, underreacting to new information. Moreover, we have illustrated (see Table 6 in section 2.4.3) that model-consistent forecasts formed under the full rationality assumption inherit a similar problem. In this section, we show that the specifications with boundedly rational beliefs, which allow for the agents' expectation formation process to deviate from FIRE, may generate more efficient information processing within a model and, therefore, reduce the "predictability problem". We estimate the regression equation (1) to verify the degree of predictability b contained in the forecasts produced by RB and HB models. We distinguish between the predictions generated by the agents' forecasting rules (PLM) and the model-based forecasts (ALM).

Table 15. Predictability test: SPF, RE and alternative belief models

	Inflation		Consumption		Investment		GDP	
SPF-nowcast b-coefficient	-0.16		0.17		0.49***		0.17	
RE-model	ALM=PLM							
b-coefficient	0.04		0.18		0.72***		0.14	
RB flex	ALM	PLM	ALM	PLM	ALM	PLM	ALM	PLM
b-coefficient	0.24	0.17	0.12	0.40	0.17	0.49***	-0.14	0.11
HB flex	ALM	PLM	ALM	PLM	ALM	PLM	ALM	PLM
b-coefficient	-0.07	0.10	0.29	0.48	0.34	0.82***	0.00	0.15

Table 15 shows that the expectation formation mechanism assumed by non-rational belief models, which implies sluggish and incomplete adjustment to new information in the PLM, appears to be well suited to describe the specific features of the survey data, such as the predictability of investment forecast errors. In particular, the PLM investment forecasts produced by the RB model show a very similar degree of predictability compared to the level observed in the SPF data. The values of the b-coefficients estimated for the HB model are overall close to the values obtained for the RE model. This result is rather intuitive given that the HB model is dominated by the MSV transmission mechanism that resembles RE to a significant extent.

At the same time, we observe that models with imperfectly rational expectations can reduce systematic underprediction and, therefore, generate efficient ALM forecasts for all variables including investment (b-coefficient for investment in RB and HB models is not significantly different from zero). Such a result is consistent with the notable improvement of the precision of investment predictions documented in the previous sections. The transmission mechanism of the alternative belief models allows for the divergence between the model-based (ALM) and agents' forecasts (PLM). This reduces the impact of survey data on the model predictions. As a result, the ALM forecasts can better exploit valuable information about economic relationships described by the model structural equations as well as benefit from the beliefs updating process, in addition to the gains obtained from timely information contained in the surveys. In this way, the time varying and imperfectly rational ALM transmission mechanism may "correct" the systematic errors observed in the surveys, thus improving the overall efficiency of the model predictions.

6 Conclusion

The paper provides an efficient procedure for incorporating timely information from survey forecasts into DSGE-models with alternative expectation formation mechanisms. We emphasize the usefulness of survey data for forecasting and macroeconomic analysis and illustrate that observing surveys can improve the identification of shocks that play a dominant role in driving the real business cycle, such as the risk premium and investment technology shocks. Our strategy to integrate and efficiently utilize the timely information from surveys implies a re-specification of shocks into persistent and transitory components. Thus, due to the SPF, we are able to separate the sources of low and high frequency volatility and improve the predictive power of the model.

We demonstrate that models with RE and non-rational beliefs can fit surveys rather well. Therefore, we can successfully reconcile SPF forecasts of key real economic activity indicators (consumption, investment and output growth) and nominal variables (inflation) with model-based expectation measures and jointly observed realized macro data. At the same time, our results indicate that AL models produce overall superior model fit. In addition, our learning setup reveals the features of the belief specification consistent with survey evidence. In particular, we show that PLM, which can compete with the survey expectations, cannot be based on very small forecasting rules. It is essential to expand the information set used in the PLM and, more specifically, to enrich the belief specification with innovations of the structural shocks. Such a formulation of the belief model is consistent with the evidence of professional forecasters using rather extensive information set and trying to process new information.

We show that models with AL can overcome certain limitations of the RE models, such as the inability to capture evolving macroeconomic trends, stability of the transmission mechanism, and predictability of model-based forecasts inherited from surveys. In particular, we demonstrate that AL models can better capture trend shifts and utilize this property for improved long-term predictions. In addition, non-rational belief models can produce time-varying and amplified effects of shocks. This determines the ability of AL models to generate stronger cyclical variation of real variables and outperform both the SPF and RE-model forecasts in predicting investment dynamics. Finally, while surveys produce systematically inefficient forecasts, non-rational belief models are able to generate superior predictions due to the possibility to relax the RE constraint of internal consistency between the agents' and model forecasts. The departure of the model forecasts from the agents' predictions allows the former to better explore the economic relationships described by the model structural equations.

Therefore, we argue that both RE and AL frameworks are suitable to describe the expectations formation mechanism behind the surveys. At the same time, AL models seem to be more efficient in exploring the main benefits from surveys while limiting the costs caused by the expectation errors of professional forecasters.

7 References

Andrade, Philippe, and Hervé Le Bihan. 2013. "Inattentive Professional Forecasters." Journal of Monetary Economics 60 (8): 967–82.

Angeletos, George-Marios, Huo, Zhen, and Karthik A. Sastry. 2020. "Imperfect Macroeconomic Expectations: Evidence and Theory." NBER Chapters, in: NBER Macroeconomics Annual 35: 1-86.

Ang, Andrew, Bekaert, Geert, and Wei Min. 2007. "Do macro variables, asset markets, or surveys forecast inflation better?" Journal of Monetary Economics 54(4): 1163-1212.

Bauer, Michael D., Pflueger, Carolin, and Adi Sunderam. 2022. "Perceptions about Monetary Policy." NBER Working Papers 30480, National Bureau of Economic Research, Inc.

Bognanni, Mark, and John Zito. 2016. "New Normal or Real-Time Noise? Revisiting the Recent Data on Labor Productivity." Economic Commentary, Federal Reserve Bank of Cleveland, issue December.

Bordalo, Pedro, Nicola Gennaioli, Yueran Ma, and Andrei Shleifer. 2020. "Overreaction in Macroeconomic Expectations." American Economic Review 110 (9): 2748-82.

Brock, William, and Cars Hommes. 1997. "A rational route to randomness." Econometrica 65: 1059–1095.

Cardani, Roberta, Croitorov, Olga, Giovannini, Massimo, Pfeiffer, Philipp, Ratto, Marco, and Vogel, Lukas. 2022. "The euro area's pandemic recession: A DSGE-based interpretation" Journal of Economic Dynamics and Control 143(C).

Carvalho, Carlos, Eusepi, Stefano, Moench, Emanuel, and Bruce Preston. 2023. "Anchored Inflation Expectations." American Economic Journal: Macroeconomics, American Economic Association 15(1): 1-47.

Chan, Joshua C. C., Eisenstat, Eric, Chenghan, Hou, and Gary Koop. 2020. "Composite likelihood methods for large Bayesian VARs with stochastic volatility." Journal of Applied Econometrics 35(6): 692-711.

Clark Todd E., Krüger, Fabian, and Francesco Ravazzolo. 2017. "Using Entropic Tilting to Combine BVAR Forecasts With External Nowcasts." Journal of Business & Economic Statistics, Taylor & Francis Journals 35(3): 470-485.

Coibion, Olivier, and Yuriy Gorodnichenko. 2011. "Monetary Policy, Trend Inflation, and the Great Moderation: An Alternative Interpretation." American Economic Review 101 (1): 341-70.

Coibion, Olivier, and Yuriy Gorodnichenko. 2012. "What Can Survey Forecasts Tell Us about Information Rigidities?" Journal of Political Economy 120 (1): 116–59.

Coibion, Olivier, and Yuriy Gorodnichenko. 2015. "Information rigidity and the expectations formation process: A simple framework and new facts." American Economic Review 105(8): 2644–2678.

Coibion, Olivier, Gorodnichenko, Yuriy, and Rupal Kamdar. 2018. "The Formation of Expectations, Inflation, and the Phillips Curve." Journal of Economic Literature, American Economic Association 56(4): 1447-1491.

Coibion, Olivier, Gorodnichenko, Yuriy, and Tiziano Ropele. 2020. "Inflation Expectations and Firm Decisions: New Causal Evidence." The Quarterly Journal of Economics 135(1): 165-219.

Croushore, Dean, and Tom Stark. 2001. "A real-time data set for macroeconomists." Journal of Econometrics 105(1): 111-130.

Crump, Richard K., Eusepi, Stefano, Tambalotti, Andrea, and Topa, Giorgio. 2022. "Subjective intertemporal substitution." Journal of Monetary Economics 126(C): 118-133.

De Grauwe, Paul. 2012. "Booms and busts: New Keynesian and behavioural explanations" Chapters, in: Robert M. Solow & Jean-Philippe Touffut (ed.), What's Right with Macroeconomics?, chapter 6: 149-180.

Del Negro, Mark, and Frank Schorfheide. 2013. "DSGE Model-Based Forecasting." Handbook of Economic Forecasting 2a: 57–140.

Evans, George W. and Seppo Honkapohja. 2001. "Learning and Expectations in Macroeconomics." Princeton University Press, Princeton, NJ.

Eusepi, Stefano, and Bruce Preston. 2011. "Expectations, Learning, and Business Cycle Fluctuations." American Economic Review 101 (6): 2844-72.

Faria-e-Castro, Miguel. 2021. "Fiscal policy during a pandemic." Journal of Economic Dynamics and Control 125(C).

Fornaro, Luca, and Martin Wolf. 2020. "Covid-19 Coronavirus and Macroeconomic Policy." Working Papers 1168, Barcelona School of Economics.

Gabaix, Xavier. 2014. "A Sparsity-Based Model of Bounded Rationality." Quarterly Journal of Economics 129 (4): 1661–1710.

Gabaix, Xavier. 2019. "Behavioral Inattention." in Handbook of Behavioral Economics, ed. by B. D. Bernheim, S. DellaVigna, and D. Laibson, Elsevier, vol. 2.

Guerrieri, Veronica, Lorenzoni, Guido, Straub, Ludwig, and Iván Werning. 2022. "Macroeconomic Implications of COVID-19: Can Negative Supply Shocks Cause Demand Shortages?" American Economic Review 112(5): 1437-1474.

Hajdini, Ina, and Andre Kurmann. 2022. "Predictable Forecast Errors in Full-Information Rational Expectations Models with Regime Shifts." School of Economics Working Paper Series 2022-5, LeBow College of Business, Drexel University.

Hommes, Cars, and Mei Zhu. 2014. "Behavioral learning equilibria." Journal of Economic Theory 150: 778-814.

Hommes, Cars, Jump, Calvert, and Paul Levine. 2019. "Learning, heterogeneity, and complexity in the New Keynesian model." Journal of Economic Behavior and Organization 166(C): 446-470.

Jacobs, Jan P.A.M., and Simon Van Norden. 2011. "Modeling data revisions: Measurement error and dynamics of "true" values." Journal of Econometrics 161(2): 101-109.

Koop, Gary, and Dimitris Korobilis. 2013. "Large time-varying parameter VARs." Journal of Econometrics 177(2): 185-198.

Lindé, Jesper, Smets, Frank, and Raf Wouters. 2016. "Challenges for Central Banks´ Macro Models." Working Paper Series 323, Sveriges Riksbank (Central Bank of Sweden).

Malmendier, Ulrike, and Stefan Nagel. 2016. "Learning from Inflation Experiences." The Quarterly Journal of Economics 131(1): 53–87.

Mankiw, N. Gregory, and Ricardo Reis. 2002. "Sticky Information versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve." Quarterly Journal of Economics 117 (4): 1295–1328.

Mankiw, N. Gregory, Reis, Ricardo, and Justin Wolfers. 2004. "Disagreement about Inflation Expectations." NBER Chapters, in: NBER Macroeconomics Annual 18: 209-270.

Manzan, Sebastiano. 2021. "Are professional forecasters Bayesian?" Journal of Economic Dynamics and Control 123(C).

Milani, Fabio. 2007. "Expectations, Learning and Macroeconomic Persistence." Journal of Monetary Economics 54: 2065–2082.

Milani, Fabio. 2011. "Expectation Shocks and Learning as Drivers of the Business Cycle." The Economic Journal 121(552): 379–401.

Milani, Fabio. 2017. "Sentiment and the U.S. business cycle." Journal of Economic Dynamics and Control 82(C): 289-311.

Molavi, Pooya. 2019. "Macroeconomics with Learning and Misspecification: A General Theory and Applications." Meeting Papers 1584, Society for Economic Dynamics.

Orphanides, Athanasios. 2004. "Monetary Policy Rules, Macroeconomic Stability, and Inflation: A View from the Trenches." Journal of Money, Credit and Banking 36(2): 151-175.

Romer, Christina, D., and David H. Romer. 2004. "A New Measure of Monetary Shocks: Derivation and Implications." American Economic Review 94 (4): 1055-1084.

Rychalovska, Yuliya. 2016. "The implications of financial frictions and imperfect knowledge in the estimated DSGE model of the U.S. economy." Journal of Economic Dynamics and Control 73(C): 259-282.

Rychalovska, Yuliya, Slobodyan, Sergey, and Raf Wouters. 2023a. "Adaptive Learning and Survey Expectations of Inflation." Unpublished.

Rychalovska, Yuliya, Slobodyan, Sergey, and Raf Wouters. 2023b. "Imperfect expectations and macroeconomic effects of business cycle shocks." Unpublished

Sims, Christopher A. 2003. "Implications of Rational Inattention." Journal of Monetary Economics 50 (3): 665–90.

Slobodyan, Sergey, and Raf Wouters. 2012. "Learning in a Medium-Scale DSGE Model with Expectations Based on Small Forecasting Models." American Economic Journal: Macroeconomics 4(2): 65-101.

Smets, Frank, and Raf Wouters. 2007. "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach." American Economic Review 97(3): 586-606.

Smets, Frank, Warne, Anders, and Raf Wouters. 2014. "Professional forecasters and real-time forecasting with a DSGE model." International Journal of Forecasting 30(4): 981-995.

Tallman, Ellis W. and Zaman Saeed. 2020. "Combining survey long-run forecasts and nowcasts with BVAR forecasts using relative entropy." International Journal of Forecasting 36(2): 373-398.

Woodford, Michael. 2003. "Imperfect Common Knowledge and the Effects of Monetary Policy." In Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps, edited by Philippe Aghion et al., 25–58. Princeton, NJ: Princeton University Press.

Woodford, Michael. 2013. "Macroeconomic Analysis Without the Rational Expectations Hypothesis." Annual Review of Economics, Annual Reviews 5(1): 303-346.

8 Appendix (For Online Publication)

8.1 Estimated parameters

Table A1: Prior and posterior distributions for RE models.

		Prior distribution			RE $\{all\}$ Metropolis Chain			RE^S $\{all\}$ Metropolis Chain		
Shocks		type	mean	st.dev.	mean	5%	95%	mean	5%	95%
AR coefficients										
Stationary tech. shock	ρ_a	B	0.50	0.20	0.98	0.97	0.99	0.99	0.97	1.00
Risk premium shock	ρ_b	B	0.50	0.20	0.9	0.87	0.94	0.91	0.88	0.95
Invest. spec. tech. shock	ρ_i	B	0.50	0.20	0.84	0.78	0.91	0.76	0.44	0.94
Gov't cons. shock	ρ_g	B	0.50	0.20	0.98	0.97	0.99	0.98	0.97	0.99
Price markup shock	ρ_p	B	0.50	0.20	0.87	0.85	0.90	0.87	0.84	0.90
Wage markup shock	ρ_w	B	0.50	0.20	0.79	0.67	0.92	0.78	0.66	0.91
Response of g_t to ε_t^a	ρ_{ga}	B	0.50	0.20	0.33	0.23	0.42	0.32	0.22	0.42
MA risk premium shock	ρ_{bma}	B	0.50	0.20	-0.34	-0.45	-0.24	-0.37	-0.45	-0.28
Monetary policy shock	ρ_r	B	0.50	0.20	0.21	0.11	0.30	0.22	0.12	0.32
Sentiment shock π_f0	$\rho_{v\_\pi f0}$	B	0.50	0.20	-	-	-	0.30	0.07	0.52
Sentiment shock dc_f0	ρ_{v_dcf0}	B	0.50	0.20	-	-	-	0.31	0.08	0.52
Sentiment shock $dinv\_f0$	ρ_{v_dinvf0}	B	0.50	0.20	-	-	-	0.58	0.23	0.96
Standard deviations										
Stationary tech. shock	σ_a	G	0.20	0.15	0.46	0.41	0.50	0.46	0.41	0.51
I.i.d Risk premium shock	σ_b	G	0.20	0.15	0.22	0.19	0.24	0.22	0.20	0.24
I.i.d Invest. spec. tech. shock	σ_i	G	0.20	0.15	0.68	0.61	0.74	0.67	0.60	0.75
I.i.d Gov't cons. shock	σ_g	G	0.20	0.15	0.19	0.17	0.21	0.19	0.17	0.21
I.i.d Price markup shock	σ_p	G	0.20	0.15	0.19	0.17	0.20	0.19	0.17	0.21
I.i.d Wage markup shock	σ_w	G	0.20	0.15	0.19	0.16	0.21	0.19	0.17	0.21
Persistent price markup shock	ϑ_p	G	0.20	0.15	0.01	0.01	0.01	0.01	0.01	0.01
Persistent wage markup shock	ϑ_w	G	0.20	0.15	0.02	0.01	0.03	0.02	0.01	0.03
Persistent risk premium shock	ϑ_b	G	0.20	0.15	0.03	0.02	0.04	0.02	0.01	0.03
Persistent invest. tech. shock	ϑ_i	G	0.20	0.15	0.12	0.09	0.15	0.13	0.05	0.24
Persistent gov't cons. shock	ϑ_g	G	0.20	0.15	0.31	0.28	0.34	0.31	0.28	0.34
Mon. pol. shock	σ_r	G	0.20	0.15	0.12	0.10	0.13	0.12	0.10	0.13
m.e. π_f0	$\sigma_{e\_\pi f0}$	G	0.20	0.15	0.04	0.02	0.06	0.03	0.02	0.05
m.e. dy_f0	σ_{e_dyf0}	G	0.20	0.15	0.06	0.02	0.09	0.06	0.02	0.10
m.e. dc_f0	σ_{e_dcf0}	G	0.20	0.15	0.11	0.05	0.16	0.06	0.03	0.10
m.e. $dinv\_f0$	σ_{e_dinvf0}	G	0.20	0.15	0.08	0.03	0.14	0.08	0.03	0.13
Sentiment shock π_f0	$\sigma_{v\_\pi f0}$	G	0.20	0.15				0.02	0.01	0.03
Sentiment shock dc_f0	σ_{v_dcf0}	G	0.20	0.15				0.10	0.06	0.14
Sentiment shock $dinv\_f0$	σ_{v_dinvf0}	G	0.20	0.15				0.07	0.01	0.12
Structural Parameters										
Calvo prob. wages	ξ_w	B	0.50	0.10	0.89	0.84	0.93	0.88	0.84	0.92
Calvo prob. prices	ξ_p	B	0.50	0.10	0.94	0.92	0.95	0.93	0.92	0.95
Indexation wages	ι_w	B	0.50	0.15	0.39	0.21	0.57	0.39	0.20	0.56
Indexation prices	ι_p	B	0.50	0.15	0.07	0.03	0.11	0.07	0.03	0.10
Gross price markup	ϕ_p	N	1.25	0.12	1.23	1.12	1.34	1.25	1.14	1.37
Capital production share	α	N	0.30	0.05	0.13	0.11	0.14	0.13	0.11	0.14
Capital utilization cost	ψ	B	0.50	0.15	0.46	0.27	0.64	0.49	0.29	0.69
Investment adj. cost	φ	N	4.00	1.50	6.96	5.07	8.78	6.97	5.15	8.70
Habit formation	$\varkappa$	B	0.70	0.10	0.80	0.75	0.85	0.84	0.80	0.89
Inv elast of subst.cons.	σ_c	N	1.50	0.37	1.01	0.92	1.09	1.02	0.94	1.1
Labor supply elast.	σ_l	N	2.00	0.5	1.69	0.88	2.45	1.67	0.93	2.40
Log hours worked in S.S.	$\bar{l}$	N	0.00	2.00	0.09	0.07	0.12	0.10	0.08	0.12
Discount factor	$100(\beta^{-1}-1)$	G	0.25	0.10	0.17	0.07	0.26	0.18	0.07	0.28
Quarterly Growth in S.S.	$\bar{\gamma}$	N	0.40	0.10	0.33	0.28	0.38	0.32	0.25	0.38
Quarterly infl. rate. in S.S.	$\bar{\pi}$	G	0.62	0.10	0.69	0.58	0.8	0.69	0.58	0.79
Inflation response	r_π	N	1.50	0.25	1.30	1	1.54	1.30	1	1.55
Output gap response	r_y	N	0.12	0.05	0.09	0.06	0.11	0.08	0.06	0.11
Diff. output gap response	$r_{\Delta y}$	N	0.12	0.05	0.13	0.11	0.16	0.12	0.10	0.15
Interest rate smoothing	ρ_R	B	0.75	0.10	0.88	0.85	0.90	0.87	0.84	0.90
Trend $dinv\_r1$	$\bar{\gamma}_{dinvr1}$	N	0.40	0.10	0.29	0.20	0.38	0.28	0.20	0.37
Trend $dinv\_r2$	$\bar{\gamma}_{dinvr2}$	N	0.40	0.10	0.39	0.33	0.45	0.38	0.32	0.44
Trend dc_f0	$\bar{\gamma}_{dcf0}$	N	0.0	0.10	-0.08	-0.13	-0.03	-0.08	-0.13	-0.04
Trend dy_f0	$\bar{\gamma}_{dyf0}$	N	0.0	0.10	-0.28	-0.35	-0.20	-0.26	-0.35	-0.17
Trend $dinv\_f0$	$\bar{\gamma}_{dinvf0}$	N	0.0	0.10	0.12	0.00	0.24	0.12	-0.01	0.25
Log marginal likelihood					MCMC	-385.07		MCMC	-388.17	

Table A2: Prior and posterior distributions for non-RE models: MSV and RB

Shocks		Prior distribution			MSV Metropolis Chain			RB Metropolis Chain		
		type	mean	st.dev.	mean	5%	95%	mean	5%	95%
AR coefficients					.		.			
Stationary technology shock	ρ_a	B	0.50	0.20	0.97	0.96	0.98	0.99	0.99	0.99
Risk premium shock	ρ_b	B	0.50	0.20	0.90	0.86	0.94	0.92	0.90	0.94
Invest. spec. tech. shock	ρ_i	B	0.50	0.20	0.78	0.71	0.86	0.90	0.85	0.95
Gov't cons. shock	ρ_g	B	0.50	0.20	0.98	0.98	0.99	0.98	0.98	0.99
Price markup shock	ρ_p	B	0.50	0.20	0.85	0.80	0.90	0.64	0.42	0.88
Wage markup shock	ρ_w	B	0.50	0.20	0.55	0.40	0.70	0.52	0.28	0.75
Response of g_t to ε_t^a	ρ_{ga}	B	0.50	0.20	0.33	0.23	0.43	0.33	0.23	0.42
MA risk premium shock	ρ_{bma}	B	0.50	0.20	-0.35	-0.46	-0.24	-0.22	-0.27	-0.16
Monetary policy shock	ρ_r	B	0.50	0.20	0.20	0.11	0.28	0.16	0.08	0.24
Standard deviations										
Stationary technology shock	σ_a	G	0.20	0.15	0.45	0.41	0.50	0.44	0.40	0.48
I.i.d Risk premium shock	σ_{b_iid}	G	0.20	0.15	0.22	0.20	0.24	0.24	0.22	0.27
I.i.d Invest. spec. tech. shock	σ_{i_iid}	G	0.20	0.15	0.66	0.59	0.72	0.67	0.62	0.72
I.i.d Gov't cons. shock	σ_{g_iid}	G	0.20	0.15	0.19	0.17	0.21	0.20	0.18	0.22
I.i.d Price markup shock	σ_{p_iid}	G	0.20	0.15	0.18	0.17	0.20	0.18	0.16	0.20
I.i.d Wage markup shock	σ_{w_iid}	G	0.20	0.15	0.17	0.13	0.22	0.15	0.12	0.18
Persistent price markup shock	σ_{p_ar}	G	0.20	0.15	0.01	0.01	0.02	0.02	0.01	0.03
Persistent wage markup shock	σ_{w_ar}	G	0.20	0.15	0.06	0.02	0.09	0.04	0.01	0.06
Persistent risk premium shock	σ_{b_ar}	G	0.20	0.15	0.03	0.02	0.04	0.02	0.02	0.03
Persistent invest. tech. shock	σ_{i_ar}	G	0.20	0.15	0.15	0.11	0.19	0.05	0.03	0.06
Persistent gov't cons. shock	σ_{g_ar}	G	0.20	0.15	0.31	0.28	0.34	0.30	0.27	0.33
Monetary policy shock	σ_r	G	0.20	0.15	0.12	0.10	0.13	0.11	0.10	0.12
m.e. π_f0	$\sigma_{e\_\pi f0}$	G	0.20	0.15	0.05	0.04	0.06	0.07	0.05	0.08
m.e. dy_f0	σ_{e_dyf0}	G	0.20	0.15	0.06	0.02	0.09	0.06	0.02	0.09
m.e. dc_f0	σ_{e_dcf0}	G	0.20	0.15	0.12	0.07	0.16	0.15	0.13	0.17
m.e. $dinv\_f0$	σ_{e_dinvf0}	G	0.20	0.15	0.08	0.03	0.12	0.06	0.01	0.11
Structural Parameters										
Calvo prob. wages	ξ_w	B	0.50	0.10	0.91	0.89	0.94	0.79	0.74	0.84
Calvo prob. prices	ξ_p	B	0.50	0.10	0.94	0.93	0.95	0.90	0.87	0.92
Indexation wages	ι_w	B	0.50	0.15	0.41	0.24	0.59	0.42	0.25	0.60
Indexation prices	ι_p	B	0.50	0.15	0.08	0.04	0.12	0.09	0.05	0.14
Gross price markup	ϕ_p	N	1.25	0.12	1.22	1.12	1.32	1.26	1.14	1.37
Capital production share	α	N	0.30	0.05	0.12	0.11	0.14	0.13	0.11	0.15
Capital utilization cost	ψ	B	0.50	0.15	0.50	0.33	0.66	0.52	0.28	0.76
Investment adj. cost	φ	N	4.00	1.50	6.80	5.07	8.58	7.81	7.04	8.71
Habit formation	$\varkappa$	B	0.70	0.10	0.81	0.76	0.86	0.84	0.82	0.87
Inv elast of subst.cons.	σ_c	N	1.50	0.37	0.98	0.92	1.04	0.92	0.86	0.97
Labor supply elast.	σ_l	N	2.00	0.50	1.96	1.64	2.27	1.44	0.56	2.31
Log hours worked in S.S.	$\bar{l}$	N	0.00	2.00	0.07	0.05	0.10	0.09	0.06	0.12
Discount factor	$100(\beta^{-1}-1)$	G	0.25	0.10	0.17	0.07	0.26	0.24	0.11	0.36
Quarterly Growth in S.S.	$\bar{\gamma}$	N	0.40	0.10	0.36	0.33	0.39	0.34	0.29	0.38
Quarterly infl. rate. in S.S.	$\bar{\pi}$	G	0.62	0.10	0.89	0.74	1.04	0.70	0.60	0.81
Inflation response	r_π	N	1.50	0.25	1.39	1.06	1.70	1.28	1.01	1.49
Output gap response	r_y	N	0.12	0.05	0.10	0.07	0.13	0.08	0.05	0.10
Diff. output gap response	$r_{\Delta y}$	N	0.12	0.05	0.13	0.11	0.16	0.11	0.08	0.13
Interest rate smoothing	ρ_R	B	0.75	0.10	0.88	0.85	0.91	0.87	0.84	0.90
Trend $dinv\_r1$	$\bar{\gamma}_{dinvr1}$	N	0.40	0.10	0.29	0.21	0.38	0.29	0.21	0.38
Trend $dinv\_r2$	$\bar{\gamma}_{dinvr2}$	N	0.40	0.10	0.39	0.34	0.46	0.40	0.35	0.46
Trend dc_f0	$\bar{\gamma}_{dcf0}$	N	0.0	0.10	-0.11	-0.14	-0.07	-0.08	-0.13	-0.02
Trend dy_f0	$\bar{\gamma}_{dyf0}$	N	0.0	0.10	-0.32	-0.37	-0.26	-0.28	-0.35	-0.22
Trend $dinv\_f0$	$\bar{\gamma}_{dinvf0}$	N	0.0	0.10	0.14	0.06	0.25	0.13	0.02	0.24
Learning persistence	ρ^{AL}	U	0.00	1.00	-	-	-	0.87	0.85	0.89
Learning persistence constant	ρ^{AL}_{const}	U	0.00	1.00	0.99	0.99	1.00	0.34	0.00	0.69
Volatility const.belief coeff.	σ^{AL}_{const}	U	0.00	1.00	0.66	0.43	0.89	0.68	0.36	1.00
HB fitness parameter	ζ	N	5.0	1.00	-	-	-	-	-	-
Log marginal likelihood					MCMC	–381.38		MCMC	-351.59	

Table A3: Prior and posterior distributions for non-RE models: HB and RB with sentiment shocks

Shocks		Prior distribution			HB Metropolis Chain			RB^S Metropolis Chain		
		type	mean	st.dev.	mean	5%	95%	mean	5%	95%
AR coefficients										
Stationary technology shock	ρ_a	B	0.50	0.20	0.97	0.96	0.98	0.99	0.99	1.00
Risk premium shock	ρ_b	B	0.50	0.20	0.90	0.86	0.93	0.93	0.90	0.95
Invest. spec. tech. shock	ρ_i	B	0.50	0.20	0.91	0.85	0.97	0.58	0.13	0.92
Gov't cons. shock	ρ_g	B	0.50	0.20	0.98	0.98	0.99	0.98	0.97	0.99
Price markup shock	ρ_p	B	0.50	0.20	0.80	0.72	0.88	0.64	0.33	0.89
Wage markup shock	ρ_w	B	0.50	0.20	0.54	0.39	0.68	0.47	0.22	0.72
Response of g_t to ε_t^a	ρ_{ga}	B	0.50	0.20	0.32	0.23	0.42	0.33	0.23	0.42
MA risk premium shock	ρ_{bma}	B	0.50	0.20	-0.30	-0.39	-0.22	-0.21	-0.27	-0.15
Monetary policy shock	ρ_r	B	0.50	0.20	0.20	0.10	0.29	0.17	0.08	0.26
Sentiment shock π_f0	$\rho_{v\_\pi f0}$	B	0.50	0.20	-	-	-	0.24	0.07	0.40
Sentiment shock dc_f0	ρ_{v_dcf0}	B	0.50	0.20	-	-	-	0.32	0.01	0.68
Sentiment shock $dinv\_f0$	ρ_{v_dinvf0}	B	0.50	0.20	-	-	-	0.42	0.21	0.63
Standard deviations										
Stationary technology shock	σ_a	G	0.20	0.15	0.46	0.41	0.50	0.46	0.41	0.50
I.i.d Risk premium shock	σ_{b_iid}	G	0.20	0.15	0.22	0.20	0.25	0.25	0.22	0.27
I.i.d Invest. spec. tech. shock	σ_{i_iid}	G	0.20	0.15	0.66	0.60	0.73	0.66	0.60	0.73
I.i.d Gov't cons. shock	σ_{g_iid}	G	0.20	0.15	0.19	0.16	0.21	0.20	0.18	0.22
I.i.d Price markup shock	σ_{p_iid}	G	0.20	0.15	0.18	0.16	0.20	0.19	0.17	0.21
I.i.d Wage markup shock	σ_{w_iid}	G	0.20	0.15	0.17	0.13	0.21	0.15	0.11	0.18
Persistent price markup shock	σ_{p_ar}	G	0.20	0.15	0.02	0.01	0.03	0.02	0.01	0.02
Persistent wage markup shock	σ_{w_ar}	G	0.20	0.15	0.06	0.02	0.09	0.04	0.01	0.07
Persistent risk premium shock	σ_{b_ar}	G	0.20	0.15	0.03	0.02	0.03	0.02	0.02	0.03
Persistent invest. tech. shock	σ_{i_ar}	G	0.20	0.15	0.08	0.05	0.10	0.07	0.01	0.13
Persistent gov't cons. shock	σ_{g_ar}	G	0.20	0.15	0.31	0.27	0.34	0.30	0.27	0.33
Monetary policy shock	σ_r	G	0.20	0.15	0.11	0.10	0.12	0.11	0.10	0.12
m.e. π_f0	$\sigma_{e\_\pi f0}$	G	0.20	0.15	0.05	0.03	0.06	0.06	0.05	0.07
m.e. dy_f0	σ_{e_dyf0}	G	0.20	0.15	0.06	0.02	0.10	0.06	0.02	0.09
m.e. dc_f0	σ_{e_dcf0}	G	0.20	0.15	0.13	0.10	0.17	0.14	0.10	0.17
m.e. $dinv\_f0$	σ_{e_dinvf0}	G	0.20	0.15	0.09	0.03	0.14	0.07	0.01	0.13
Sentiment shock π_f0	$\sigma_{v\_\pi f0}$	G	0.20	0.15				0.03	0.01	0.05
Sentiment shock dc_f0	σ_{v_dcf0}	G	0.20	0.15				0.03	0.01	0.04
Sentiment shock $dinv\_f0$	σ_{v_dinvf0}	G	0.20	0.15				0.09	0.04	0.13
Structural Parameters										
Calvo prob. wages	ξ_w	B	0.50	0.10	0.89	0.86	0.92	0.82	0.76	0.87
Calvo prob. prices	ξ_p	B	0.50	0.10	0.92	0.90	0.95	0.91	0.88	0.94
Indexation wages	ι_w	B	0.50	0.15	0.37	0.20	0.54	0.41	0.23	0.60
Indexation prices	ι_p	B	0.50	0.15	0.08	0.04	0.12	0.08	0.04	0.12
Gross price markup	ϕ_p	N	1.25	0.12	1.21	1.11	1.31	1.23	1.12	1.34
Capital production share	α	N	0.30	0.05	0.12	0.11	0.14	0.13	0.11	0.14
Capital utilization cost	ψ	B	0.50	0.15	0.51	0.31	0.72	0.44	0.23	0.66
Investment adj. cost	φ	N	4.00	1.50	6.69	5.17	8.23	6.99	5.33	8.68
Habit formation	$\varkappa$	B	0.70	0.10	0.81	0.77	0.85	0.85	0.81	0.88
Inv elast of subst.cons.	σ_c	N	1.50	0.37	1.07	0.99	1.14	0.90	0.85	0.96
Labor supply elast.	σ_l	N	2.00	0.50	1.86	1.45	2.29	1.64	0.89	2.40
Log hours worked in S.S.	$\bar{l}$	N	0.00	2.00	0.09	0.07	0.11	0.10	0.07	0.13
Discount factor	$100(\beta^{-1}-1)$	G	0.25	0.10	0.16	0.07	0.26	0.24	0.10	0.37
Quarterly Growth in S.S.	$\bar{\gamma}$	N	0.40	0.10	0.37	0.34	0.39	0.32	0.27	0.36
Quarterly infl. rate. in S.S.	$\bar{\pi}$	G	0.62	0.10	0.83	0.72	0.94	0.70	0.61	0.80
Inflation response	r_π	N	1.50	0.25	1.34	1.05	1.60	1.25	1.00	1.46
Output gap response	r_y	N	0.12	0.05	0.08	0.06	0.11	0.08	0.05	0.10
Diff. output gap response	$r_{\Delta y}$	N	0.12	0.05	0.12	0.10	0.14	0.12	0.09	0.14
Interest rate smoothing	ρ_R	B	0.75	0.10	0.87	0.84	0.90	0.87	0.84	0.90
Trend $dinv\_r1$	$\bar{\gamma}_{dinvr1}$	N	0.40	0.10	0.30	0.21	0.40	0.30	0.21	0.39
Trend $dinv\_r2$	$\bar{\gamma}_{dinvr2}$	N	0.40	0.10	0.41	0.34	0.47	0.40	0.34	0.45
Trend dc_f0	$\bar{\gamma}_{dcf0}$	N	0.0	0.10	-0.10	-0.13	-0.07	-0.07	-0.12	-0.01
Trend dy_f0	$\bar{\gamma}_{dyf0}$	N	0.0	0.10	-0.31	-0.37	-0.26	-0.25	-0.33	-0.18
Trend $dinv\_f0$	$\bar{\gamma}_{dinvf0}$	N	0.0	0.10	0.18	0.12	0.25	0.09	-0.02	0.20
Learning persistence	ρ^{AL}	U	0.00	1.00	0.89	0.85	0.92	0.86	0.83	0.89
Volatility const. belief coeff.	σ^{AL}_{const}	U	0.00	1.00	0.43	0.29	0.58	0.68	0.35	1.00
Memory in forecast error variance	θ	U	0.00	1.00	0.30	0.23	0.38	-	-	-
HB fitness parameter	ζ	N	5.0	1.00	4.77	3.28	6.29	-	-	-
Log marginal likelihood					MCMC	-355.09		MCMC	-359.96	

8.2 Likelihood decomposition

To clarify the sources of the improvement in the fit, we decompose the likelihood function into several components, which enable us to distinguish the factors linked to the forecast precision and to the second moments of the compared models. In general terms, the log likelihood (at the posterior mode) could be given as:

$$\begin{aligned} \ln LIK_i \tilde{} \ln\{det \sum\nolimits_i^{-1/2} exp(-\frac{1}{2}\upsilon_i^T \sum\nolimits_i^{-1} \upsilon_i)\} &= \\ -\frac{1}{2}\{\ln\det \sum\nolimits_i + \upsilon_i^T \sum\nolimits_i^{-1} \upsilon_i\},\ i &\in \{RE, RB\} \end{aligned} \tag{14}$$

where υ denotes the vector of forecast errors and $\sum_i$ denotes the variance-covariance matrix of forecast errors for the model $'i'$. Ignoring the priors, the difference in the log posteriors of the RB and RE models, could be presented by the expression below.

$$\begin{aligned} \Delta \ln LIK &= \ln LIK_{RB} - \ln LIK_{RE} = \\ & \frac{1}{2}\{\ln\det \sum\nolimits_{RB}^{-1} - \ln\det \sum\nolimits_{RE}^{-1}\} - \frac{1}{2}\{\upsilon_{RB}^T \sum\nolimits_{RB}^{-1} \upsilon_{RB} - \upsilon_{RE}^T \sum\nolimits_{RE}^{-1} \upsilon_{RE}\} \end{aligned} \tag{15}$$

$$\begin{aligned} &= \frac{1}{2}\{\ln\det \sum\nolimits_{RB}^{-1} - \ln\det \sum\nolimits_{RE}^{-1}\} - \frac{1}{2}\upsilon_{RB}^T \Delta \sum\nolimits^{-1} \upsilon_{RB} \\ & -\frac{1}{2}(\upsilon_{RB} - \upsilon_{RE}) \sum\nolimits_{RE}^{-1} (\upsilon_{RB} + \upsilon_{RE}) \end{aligned} \tag{16}$$

Positive likelihood difference implies that the RB model overall outperforms the RE model. The term $\det \sum_{RB}^{-1}$ in (15) equals to the product of eigenvalues of the inverse variance-covariance matrix $\sum_{RB}^{-1}$ and in the diagonal case equals to the product of inverses of variances of individual variables. Thus, $\det \sum_{RB}^{-1}$ increases in the lower total variability allowed by the model. Therefore, the log likelihood difference will be increasing if the RB model explains the data with less volatility than the RE model does. We can further separate the effect of changes in the forecast errors and in the second moments by decomposing the second term in (15). The term $\upsilon_{RB}^T \Delta \sum^{-1} \upsilon_{RB}$ in (16) can be interpreted as a quadratic form in the differential variance-covariance matrix $\Delta \sum^{-1} = \sum_{RB}^{-1} - \sum_{RE}^{-1}$. This component captures an additional (and potentially different) effect of model volatility on likelihood. In particular, time-varying volatility in this case is acting as a factor which "scales" the size of the forecast errors. Lower total variability of RB model (positive $\Delta \sum^{-1}$ term) then suggests that the same errors will be evaluated at smaller variances, leading to a negative contribution to the likelihood difference. On the other hand, the ability of the RB model to "detect" the periods of increasing macroeconomic volatility and adjust the model variance accordingly may generate improvement in the likelihood because forecast errors, which also normally increase during such periods, will be evaluated at the higher variances. Therefore, we consider the first term (in curly brackets) and second term in (16) jointly (and denote them as $\Delta \sum$ in the Figure 5) in order to assess the overall effect of the model volatility on likelihood. The third term in (16) is the measure of the difference of the squared forecast errors (MSFE) normalized by the stable and time-invariant variance-covariance matrix implied by the RE model. Higher forecast precision of the RB model relative to the RE model, which implies $\upsilon_{RB} - \upsilon_{RE} < 0$, will produce a positive contribution to the likelihood difference function.

8.3 Tables on the forecasting performance

Table A4: 1q to 5q in-sample forecast comparison

	Realized data					SPF data			
	π_{r1}	dy_{r1}	dc_{r1}	$dinve_{r1}$	$robs$	π_{f0}	dy_{f0}	dc_{f0}	$dinve_{f0}$
RMSFE 1Q									
SPF	0.21	0.33	0.41	1.45					
RE	0.21	0.33	0.40	1.46	0.10	0.10	0.28	0.27	0.66
MSV	1.00	1.02	1.00	1.00	0.99	0.99	0.99	0.97**	1.00
RB	0.99	1.02	1.01	0.95*	0.98	1.03	1.02	1.03	0.94
HB	1.01	0.99	1.01	0.89	1.01	0.98	0.95*	0.99	0.79***
RMSFE 2Q									
SPF	0.24	0.41	0.47	1.75					
RE	0.23	0.44	0.48	1.79	0.19	0.12	0.31	0.29	0.89
MSV	0.99	1.00	0.98*	1.00	0.99	0.95*	1.00	0.98	0.98
RB	0.99	1.01	1.02	0.94	0.97**	1.01	1.06	1.04	0.94
HB	0.99	0.96*	0.99	0.92	0.99	0.97	1.00	1.03	0.87**
RMSFE 3Q									
SPF	0.25	0.46	0.49	1.94					
RE	0.24	0.46	0.48	1.95	0.26	0.14	0.33	0.29	1.03
MSV	0.99	1.01	0.98*	0.99	0.99	0.95	0.99	0.97	0.98
RB	0.99	1.01	1.01	0.99	0.98*	1.01	1.04	1.00	0.94*
HB	1.02	1.01	1.00	0.98	0.99	1.04	1.01	1.01	0.93
RMSFE 4Q									
SPF	0.28	0.48	0.51	2.00					
RE	0.25	0.48	0.49	2.02	0.32	0.15	0.34	0.30	1.11
MSV	1.00	1.01	0.98	0.99	0.99	0.92	0.98	0.96	0.96
RB	0.99	1.01	1.00	0.96**	0.99	1.01	1.00	0.98	0.91**
HB	1.06**	0.98	0.99	0.97	0.99	1.02	0.97	0.99	0.91
RMSFE 5Q									
SPF	0.29	0.49	0.50	2.06					
RE	0.25	0.48	0.50	2.06	0.38	0.15	0.34	0.30	1.15
MSV	1.00	0.99	0.97	0.99	1.00	0.91	0.98	0.95	0.96
RB	0.99	1.01	0.99	0.96**	0.99	1.02	0.99	0.98	0.89**
HB	1.06**	0.97	0.97	0.95**	0.99	1.04	0.96	0.98	0.90

Note: RMSFEs statistics are based on in-sample predictions. For the SPF predictions of the first release of variables as well as RE-model forecasts, the RMSFEs are reported in levels. For models with non-rational beliefs, we report the ratios of the RMSFE produced by these models to the RMSFE of the model with RE. ***, **, and * denote the rejection of the null of the DM test, stating that the squared forecasts errors from non-RE and RE models do not differ significantly, at the 1%, 5%, and 10% significance level, respectively. Therefore, values below 1 marked with star(s) indicate that forecasts produced by non-RE models are statistically better than the RE-model's predictions.

Table A5: DM test for equal accuracy of the model forecasts and SPF

	Realized data			
	π_{r1}	dy_{r1}	dc_{r1}	$dinve_{r1}$
RMSFE 1Q				
RE	0.41	0.06	-1.93*	0.99
MSV	0.37	0.78	-1.11	0.84
RB	0.16	0.50	-0.69	-1.88*
HB	0.30	-0.50	-0.66	-1.55
RMSFE 2Q				
RE	-0.83	1.22	0.54	0.83
MSV	-1.14	1.32	0.07	0.67
RB	-0.40	0.83	0.99	-1.24
HB	-1.15	0.09	0.41	-1.24
RMSFE 3Q				
RE	-1.70	-0.69	-0.20	-0.10
MSV	-1.93*	-0.61	-0.72	-0.40
RB	-1.87*	-0.68	0.18	-0.41
HB	-1.37	-0.66	-0.01	-0.79
RMSFE 4Q				
RE	-1.84*	-0.75	-0.56	-0.20
MSV	-1.96**	-0.75	-1.21	-0.25
RB	-1.68*	-0.39	-0.04	-0.72
HB	-1.20	-1.02	-0.52	-0.71
RMSFE 5Q				
RE	-1.80*	-0.51	0.16	-0.44
MSV	-1.91*	-0.59	-0.67	-0.67
RB	-1.89*	-0.29	0.14	-0.89
HB	-1.25	-0.82	-0.18	-1.20

Note: DM statistics are based on in-sample predictions. The positive numbers of the DM statistics indicate the positive sample mean loss differential, which is defined as the difference in the squared errors of the model forecast and SPF forecast. Therefore, the positive loss differential indicates better forecasting performance of SPF whereas the negative numbers imply that the models produce more precise predictions. ***, **, and * denote the rejection of the null of the DM test, stating that the squared forecasts errors from the model forecast and SPF predictions do not differ significantly, at the 1%, 5%, and 10% significance level, respectively.

Table A6: 1q to 5q out-of-sample forecast comparison

2008q1-2019q2	Realized data π_{r1}	dy_{r1}	dc_{r1}	$dinve_{r1}$	$robs$	SPF data π_{f0}	dy_{f0}	dc_{f0}	$dinve_{f0}$
RMSFE 1q									
SPF	0.22	0.28	0.34	1.68					
RE	0.22	0.29	0.35	1.77	0.09	0.09	0.30	0.24	0.73
MSV	0.99	1.10***	1.04**	1.03***	0.99	1.03	0.98	1.00	1.00
RB	1.00	1.07	1.10*	0.93**	0.87***	1.14**	1.00	1.08*	0.89
HB	0.97	1.14***	1.18***	0.99	0.96	0.98	1.13	1.19***	1.27
RMSFE 2q									
SPF	0.22	0.41	0.38	2.03					
RE	0.22	0.47	0.41	2.22	0.17	0.12	0.36	0.27	1.08
MSV	1.00	1.04***	1.00	1.01	0.97*	0.98	0.97	0.98	0.98**
RB	1.02	0.99	1.07	0.92**	0.84***	1.09*	0.99	1.06**	0.83**
HB	1.01	1.08	1.05	1.03	0.95	0.95	1.10	1.17***	1.06
RMSFE 3q									
SPF	0.23	0.50	0.38	2.38					
RE	0.23	0.53	0.40	2.50	0.25	0.12	0.40	0.28	1.32
MSV	0.98	1.04***	0.99	1.00	0.97	1.00	0.96	0.98	0.97***
RB	1.02	1.01	1.01	0.97	0.84***	1.12	0.95	1.05	0.86***
HB	0.99	1.06	1.07**	0.97	0.94**	1.05	1.00	1.12***	0.94
RMSFE 4q									
SPF	0.24	0.55	0.43	2.46					
RE	0.23	0.55	0.41	2.64	0.32	0.12	0.41	0.28	1.45
MSV	0.99	1.03***	1.02	0.99*	0.98	1.01	0.95*	0.99	0.95***
RB	1.04	0.98	1.04	0.93***	0.83***	1.21*	0.91**	1.04	0.83***
HB	1.06**	1.01	1.09	0.93**	0.94**	1.07	0.97	1.13***	0.88**
RMSFE 5q									
SPF	0.25	0.56	0.42	2.59					
RE	0.24	0.56	0.42	2.74	0.39	0.12	0.42	0.29	1.53
MSV	0.99	1.02*	1.00	0.98***	0.98	1.04	0.95*	0.99	0.94***
RB	1.06	0.95*	1.02	0.90***	0.81***	1.33**	0.89***	1.04	0.80***
HB	1.03	1.01	1.08	0.90***	0.94***	1.17	0.99	1.13***	0.88**

Table A7: Forecast comparison over the Covid and post-Covid period

	Realized data π_{r1}	dy_{r1}	dc_{r1}	$dinve_{r1}$	$robs$	SPF data π_{f0}	dy_{f0}	dc_{f0}	$dinve_{f0}$
RMSFE 1q									
SPF	0.62	1.18	1.43	2.39					
RE	0.64	1.19	1.42	2.36	0.50	0.23	4.09	4.46	4.62
MSV	0.99	1.00	1.00	0.98	0.98	0.98	1.00	1.00	1.00
RB	0.98	1.31	1.08	1.88	1.03	0.92	1.17	1.19	1.33
HB	0.99	1.00	1.00	0.98	0.98	1.09	1.06	1.02	0.91
RMSFE 2q									
SPF	0.86	3.75	4.04	4.01					
RE	0.83	4.77	5.25	5.41	0.66	0.32	3.62	3.90	4.64
MSV	1.00	1.00	1.00	1.01	0.99	1.00	1.00	1.00	1.01
RB	1.00	1.28	1.22	1.83	1.32	0.90	1.14	1.11	1.61
HB	1.04	1.02	1.06	0.90	1.02	1.11	1.00	1.02	0.88
RMSFE 3q									
SPF	0.92	4.01	4.57	3.68					
RE	0.89	4.13	4.51	4.81	0.76	0.38	3.55	3.82	4.59
MSV	1.01	1.00	1.00	1.01	1.00	1.02	1.00	1.00	1.01
RB	1.00	1.10	1.05	1.68	1.43	0.93	1.06	1.03	1.45
HB	1.04	1.00	1.01	0.85	1.02	1.12	1.00	1.01	0.87
RMSFE 4q									
SPF	0.96	4.02	4.52	3.75					
RE	0.95	4.11	4.55	4.38	0.81	0.43	3.55	3.83	4.48
MSV	1.02	1.00	1.00	1.01	1.01	1.04	1.00	1.00	1.00
RB	1.01	1.06	1.04	1.44	1.49	0.96	1.04	1.04	1.23
HB	1.05	0.99	1.01	0.86	1.00	1.11	0.99	1.01	0.90
RMSFE 5q									
SPF	0.96	4.01	4.50	3.76					
RE	1.00	4.07	4.53	4.11	0.83	0.48	3.55	3.83	4.45
MSV	1.02	1.00	1.00	1.00	1.01	1.04	1.00	1.00	0.99
RB	1.03	1.03	1.03	1.23	1.53	1.01	1.02	1.01	1.06
HB	1.04	1.00	1.01	0.92	0.98	1.10	0.99	1.00	0.93

Note: The statistics are based on in-sample predictions (2020q1-2022q2). For the SPF predictions of the first release of variables as well as RE-model forecasts, the RMSFEs are reported in levels. For models with non-rational beliefs, we report the ratios of the RMSFE produced by these models to the RMSFE of the model with RE. ***, **, and * denote the rejection of the null of the DM test, stating that the forecasts from non-RE and RE models do not differ significantly, at the 1%, 5%, and 10% significance level, respectively. Therefore, values below 1 marked with star(s) indicate that forecasts produced by non-RE models are statistically better than the RE-model predictions.

8.4 Robustness check: role of updating

In the baseline scenario, our AL agents, who explore only a subset of endogenous variables and shocks to formulate the forecasts, are allowed to regularly update their perceptions about the state of the economy as soon as more information becomes available. Their beliefs are summarized by the time-varying parameters β_t and β_t^{const} in equation (9). The time variation, which is an important feature of the AL process, can potentially allow detection of the trend shifts and structural breaks, changed macroeconomic persistence, modified impact of shocks and relationship between the variables, and increased uncertainty. In other words, updating is essential in periods of macroeconomic transformations because agents need to process information more intensively to stay up-to-date.

In this robustness check exercise, we would like to formally illustrate the importance of updating in our models with AL. More specifically, we keep the beliefs fixed (at their initial level consistent with the REE) during the estimation. As a result, we are able to disentangle the effects of the restricted information set and time-varying transmission mechanism. Figure A1 presents the cumulative likelihood difference for RB and HB models with and without updating. The positive trend indicates that the model with updating dominates the model with fixed transmission mechanism. The picture demonstrates that time-variation generates most of the improvement in the 1980s and after 2008. The prominent role of updating is observed during the GFC period, characterized by the dramatic increase in macroeconomic volatility and changed conventional dynamics of variables due to the novel policy measures.[53]

[53] Slobodyan and Wouters (2022) emphasize the ability of AL to handle successfully the periods of high inflation in the 1970s. We observe a similar property of AL during the GFC period. As shown in 4.1, AL can outperform the model with RE in the case of big shocks due to the ability of the former to generate higher implied model volatility, which improves the model fit.

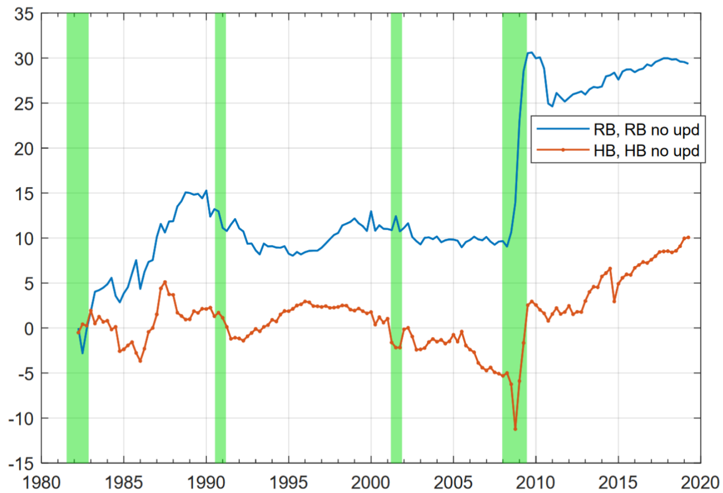

Figure A1: Cumulative difference in likelihood for RB and HB models with and without updating

Tables A8 and A9 present the comparison of the fit and forecasting performance of models with and without updating. The Tables illustrate that shutting down the updating leads to a deterioration of the model performance. The most pronounced decline in the forecast accuracy is observed for investment. Note that the model without updating still does best than the model with RE. These results again illustrate that the models with a restricted information set and time-varying transmission mechanism are better supported by data relative to models based on RE.

Table A8. Model fit comparison: role of updating

Model specification	*Log(Posterior)*	
	1981q2:2019q2	1981q2-2008q1
RE	-217.7	-137.45
RB	-176.42	-110.66
RB no updating	-201.29	-120.82
HB	-173.87	-115.15
HB no updating	-183.66	-118.52

Note: Models are evaluated over the sample periods 1981Q2 - 2019Q2 and 1981q2-2008q1 using the first four observations as a presample.

Table A9: Model forecast comparison: role of updating

	Realized data					SPF data			
	π_{r1}	dy_{r1}	dc_{r1}	$dinve_{r1}$	$robs$	π_{f0}	dy_{f0}	dc_{f0}	$dinve_{f0}$
RMSFE 1q									
RE	0.21	0.35	0.42	1.50	0.12	0.10	0.29	0.27	0.67
RB	0.21	0.35	0.42	1.42	0.11	0.10	0.30	0.28	0.63
RBno upd	1.00	1.04	1.02	1.08**	1.03**	1.02*	1.03	0.97	1.06
HB	0.21	0.34	0.42	1.34	0.11	0.09	0.28	0.27	0.54
HB no upd	0.96*	1.04	1.02	1.08**	1.03	0.94	1.03	1.01	1.13*

Note: RMSFEs statistics are based on in-sample predictions (1981q2-2019q2). For the RE, RB flex and HB flex model forecasts, the RMSFEs are reported in levels. For RB and HB models without updating, we report the ratios of the RMSFE produced by these models to the RMSFE of the corresponding models with updating. ***, **, and * denote the rejection of the null of the DM test, stating that the squared forecasts errors from models with and without updating do not differ significantly, at the 1%, 5%, and 10% significance level, respectively. Therefore, values above 1 marked with star(s) indicate that forecasts produced by models with updating are statistically better than the predictions generated by models without updating.

In addition, Table A3 shows that the role of updating is also confirmed for the RB model estimated on a shorter sample, which ends in 2008q1. In particular, we can see that in the sample, which does not include the GFC, about 50% of the improvement in likelihood is still due to updating. This confirms that the gain from time variation can be obtained not only in cases of big shocks but also in normal times.

8.5 Time-varying IRFs

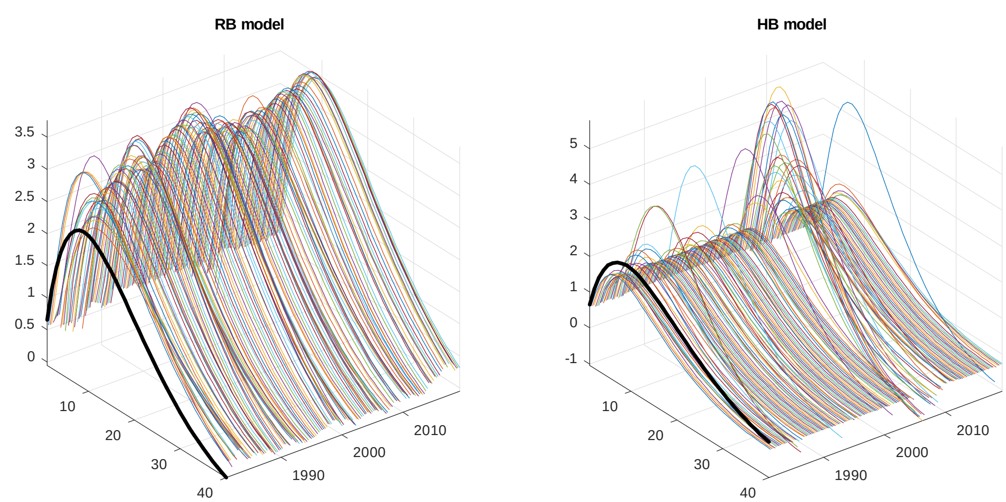

Figure A2: Time-varying Impulse Response Functions of investment to persistent risk premium shock. Note: black solid line shows the IRF under RE

8.6 Time varying volatility under AL: the main determinants

We perform several additional robustness check exercises in order to better understand the determinants of lower average RB model volatility. As Table A2 (model parameters) indicates, the estimated standard deviation of 'eqs par' shock is significantly lower under the RB model, with the evidence for other shocks being mixed. Therefore, the structural equations of the RB model seem to describe the investment dynamics with a smaller exogenous process. To evaluate the importance of the size of the shock for the model volatility, we conduct a counterfactual experiment combining the structural parameters from the RB model with the standard deviations of shocks estimated from the RE model. In another exercise, we focus exclusively on the investment specific technology shock and consider the combination of the RB structural parameters and shocks with the 'eqs par' shock, which is kept at the value consistent with RE. Figure A3, which plots the implied model volatility based on the combined parameter vectors, illustrates that the size of the shocks, and, more specifically, the magnitude of the investment technology shock, is an important determinant of the average volatility level. In particular, Figure A3 indicates that shocks consistent with the RE model make the RB set up more volatile, with the investment specific technology process playing a particularly important role in the increased volatility level.

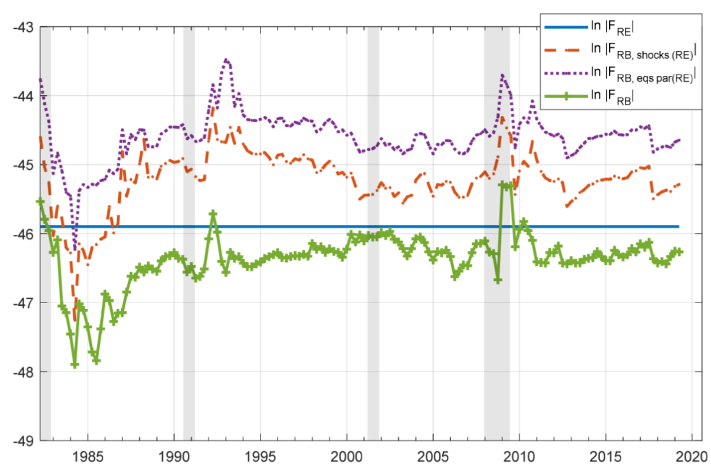

Figure A3: Implied volatility of forecast errors: the impact of the shock size

8.7 Heterogeneous beliefs model: evolution of the weight coefficient

Figure A4 presents the evolution of the weight on the RB specification, ω_t^{RB}.

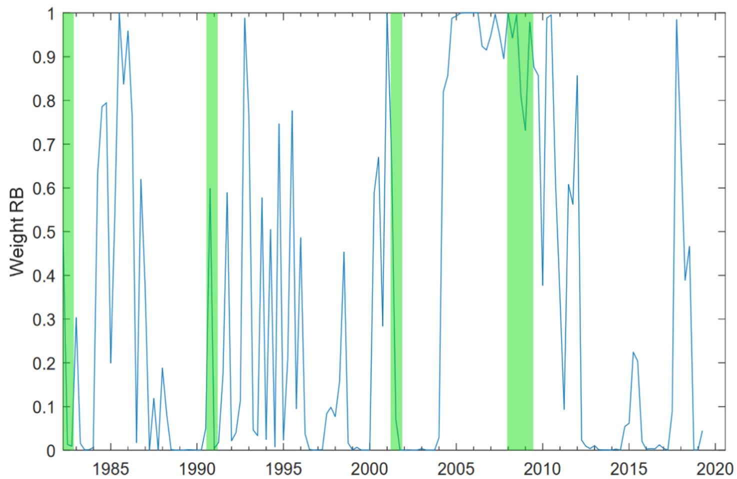

Figure A4: Weight on the RB specification in the HB model

8.8 Estimated scaling factor parameters and forecast comparison over the Covid period

Table A10. Model comparison over the extended sample

Model	*Optimal scaling factors* γ_t					*Log(Posterior)*
	2020q1	2020q2	2020q3	2020q4	2021q1	
$RE_{\{all\}}$	5.18	5.45	3.02	1.07	1.90	-361.57
MSV	5.20	5.47	2.98	1.13	1.98	-345.70
RB	5.91	7.67	6.96	4.26	3.74	-387.52
HB	5.76	6.09	3.27	1.36	2.32	-334.53

Note: Models are evaluated over the sample period 1981Q2 - 2022Q2 using the first four observations as a presample. The parameters are identified over the baseline sample 1981q2-2019q2.

Table A11: Forecast comparison over the Covid and post-Covid period

	Realized data					SPF data			
	π_{r1}	dy_{r1}	dc_{r1}	$dinve_{r1}$	$robs$	π_{f0}	dy_{f0}	dc_{f0}	$dinve_{f0}$
RMSFE 1q									
SPF	0.62	1.18	1.43	2.39					
RE	0.64	1.19	1.42	2.36	0.50	0.23	4.09	4.46	4.62
MSV	0.99	1.00	1.00	0.98	0.98*	0.98	1.00	1.00	1.00
RB	0.98	1.31***	1.08**	1.88***	1.03**	0.92*	1.17**	1.19**	1.33***
HB	0.99	1.00	1.00	0.98*	0.98*	1.09**	1.06*	1.02	0.91**
RMSFE 5q									
SPF	0.96	4.01	4.50	3.76					
RE	1.00	4.07	4.53	4.11	0.83	0.48	3.55	3.83	4.45
MSV	1.02	1.00	1.00	1.00	1.01	1.04	1.00	1.00	0.99
RB	1.03	1.03	1.03	1.23***	1.53***	1.01	1.02	1.01	1.06*
HB	1.04	1.00	1.01	0.92**	0.98*	1.10*	0.99	1.00	0.93**

Note: The statistics are based on in-sample predictions (2020q1-2022q2). For the SPF predictions of the first release of variables as well as RE-model forecasts, the RMSFEs are reported in levels. For models with non-rational beliefs, we report the ratios of the RMSFE produced by these models to the RMSFE of the model with RE. ***, **, and * denote the rejection of the null of the DM test, stating that the forecasts from non-RE and RE models do not differ significantly, at the 1%, 5%, and 10% significance level, respectively. Therefore, values below 1 marked with star(s) indicate that forecasts produced by non-RE models are statistically better than the RE-model predictions.