Belief Distortions in Risk Premia

Maximilian Boeck*

Università Bocconi

November 2023

Abstract

This paper studies belief distortions in risk premia and their reaction to financial shocks. Belief distortions are defined as ex-ante expectational errors between survey expectations and full-information rational expectations (FIRE) expectations of a machine efficient benchmark. Survey expectations of credit spreads deviate from FIRE and both under- and overreact to new information. The machine efficient benchmark exploits a high-dimensional real-time dataset with different machine learning estimators, predictively outperforming survey expectations at large. Belief distortions on financial markets identify periods of elevated optimism and pessimism on crash risk. I show that conditional on an adverse financial shock, belief distortions indicate that survey participants evaluate the future too optimistically before turning pessimistic.

Keywords: Business Cycles, Expectation Formation, Belief Distortions, Financial Shocks.

JEL Codes: C32, C52, E32, E44, E71, G41.

*Contact: Maximilian Boeck, Department of Economics, Università Bocconi. Via Roentgen 1, 20136 Milano, Italy. E-mail: maximilian.boeck@unibocconi.it. A previous version of this paper circulated under the title "Identification of Non-Rational Risk Shocks". For helpful comments and suggestions, I thank Pedro Bordalo, Maximilian Breitenlechner, Jesús Crespo Cuaresma, Martin Feldkircher, Sylvia Frühwirth-Schnatter, Emanuel Gasteiger, Raffaella Giacomini, Pia Heckl, Florian Huber, Ingrid Kubin, Lorenzo Mori, Gernot Müller, Katrin Rabitsch, Gerhard Rünstler, Alina Steshkova, Massimiliano Tancioni, Francesco Zanetti, Gregor Zens, Leopold Zessner-Spitzenberg, Thomas Zörner, as well as participants at the Forschungsseminar of the University of Salzburg, the 3rd Italian Workshop of Econometrics and Empirical Economics, the QuickTalks at King's College, the Spring Meeting of Young Economists, the 5th Vienna Workshop on High-Dimensional Time Series, the 4th Behavioral Macroeconomics Workshop, the 1st and 2nd Bergamo Workshop in Econometrics and Statistics, the Royal Economic Society 2023 Annual Conference, and the International Association of Applied Econometrics 2023 Annual Conference.

1. Introduction

Financial crises cause recessions and are costly. Models of financial frictions successfully illustrate the transmission of financial sector instabilities into a contraction of the broader economy. However, these models based on rational expectations fail to match the evidence on excessively low credit spreads in the build-up phase of financial crises. Prior to financial crises, bank equity is overvalued (Baron and Xiong, 2017), credit spreads are *too low for too long* (Krishnamurthy and Muir, 2017), and rapid credit expansions, asset price growth, and narrow credit spreads predict an ensuing crisis (Greenwood et al., 2022). Hence, markets do not seem to be aware of heightened risks.

This pattern is not unique to financial crises since excessive optimism precedes recessions consistently (Greenwood and Hanson, 2013; López-Salido, Stein and Zakrajšek, 2017). Excessive optimism or pessism are wrongly formed beliefs about the future or "the systematic misweighting of available information demonstrably pertinent to the accuracy of the belief" (Bianchi, Ludvigson and Ma, 2022). Agents can form wrong beliefs due to a variety of reasons, such as noisy or sticky information, rational inattention, overreaction to incoming news, or the use of simple extrapolative rules, among others. While belief formation has been studied extensively for macroeconomic quantities such as growth and inflation, belief formation on risk premia on financial markets has yet to be investigated. This paper examines the systematic neglect of crash risk by exploring belief distortions of credit spreads. How can one identify *ex-ante* periods of excessive optimism in the build-up phase of a financial crisis? How sizable are these belief distortions in risk premia? What is the dynamic reaction of belief distortions in risk premia conditional on a financial shock?

This paper uses time-series data to investigate belief distortions defined as the difference between survey expectations and full-information rational expectations (FIRE), or ex-ante expectational errors, in crash risk (Bianchi, Ludvigson and Ma, 2022). Specifically, I analyze the variation of belief distortions in credit spreads, which serve as a widely used proxy of risk premia and crash risk on financial markets. Belief distortions allow the interpretation that agents are temporarily overly optimistic or pessimistic. It is thus a measure of the sentiment in the economy from an *ex-ante* perspective. This allows me to investigate the dynamic reaction of belief distortions, and thus sentiments, in a vector autoregressive framework. Specifically, I identify a financial shock as unanticipated movement in the excess bond premium (EBP) similar to Gilchrist and Zakrajšek (2012). For a comparison, I also investigate the dynamic responses to survey forecast errors, which only allows an *ex-post* perspective.

Before belief distortions are constructed, I discuss and investigate the survey and machine efficient benchmark expectations in more detail. Survey expectations on credit spreads come from the Blue Chip Financial Indicators, in which a panel of financial executives is asked on their subjective risk expectations and offers thus a professionals' assessment of financial markets. I test whether survey expectations of credit spreads deviate from FIRE using the regression-based approach by Coibion and Gorodnichenko (2015). Credit spread expectations underreact at the aggregate and overreact at the individual level. These findings can be explained in a stylized model of private and public signals featuring diagnostic expectations (Bordalo, Gennaioli and Shleifer, 2018). Since market participants only observe a noisy and distorted signal of crash risk, agents form *wrong* or *undue* expectations about risk premia. The model highlights that belief distortions arise because agents do not process information rationally.

The second key ingrediant for belief distortions is to measure objective expectations of credit spreads that resemble FIRE. The approach taken in this paper is to construct a machine efficient benchmark based on machine learning (ML) techniques. This benchmark can be considered as a belief formation mechanism itself but is intended to use all available information efficiently in real-time. Practically, this boils down to a forecasting exercise using (non-)linear ML models, which involves Bayesian shrinkage priors, the Elastic Net, and Bayesian Addtive Regression Trees. Machine learning techniques are a remedy since they offer the flexibility and adaptability to study data-rich environments (Goulet Coulombe et al., 2022) and have also be proven to be useful in the prediction of asset risk premiums (Gu, Kelly and Xiu, 2020). The involved high-dimensional real-time dataset is intended to resemble the potential information set of a professional forecaster. Specifically, I use a quarterly macroeconomic and a monthly financial dataset. The machine efficient benchmark, or ML expectations, outperforms survey expectations with sizable predictive gains. Predictive power drops in crisis times but disproportionately stronger for survey expectations. This is a first indication that ML expectations pick up heightened risk in contrast to survey expectations, leading to sizable belief distortions. Although survey forecasters have in principle access to the same tools and information as the machine, expectations differ strongly. In an additional specification, I also control for the survey forecasters own forecast and still find predictive gains of the machine.

Equipped with these expectations, I construct belief distortions for Aaa and Baa credit spreads, which align well with key historical episodes. I append these belief distortons series to a set of macroeconomic variables as used in the contribution by Gilchrist and Zakrajšek (2012). This allows me to investigate the dynamic responses to a financial shock identified via timing restrictions on the EBP, whereas belief distortions are allowed to react contemporaneously to the shock. A positive one standard deviation financial shock elicits a jump in the EBP and causes an aggregate downturn. The surge in EBP is by about 20bp and real GDP contracts by about $0.2 - 0.3\%$. The reaction of belief distortion indicates that agents stay initially optimistic for one to two quarters before turning pessimistic in their evaluation of the future. A counterfactual exercise, in which the response of belief distortions to the financial shock is set to zero at all horizons, reveals that the transmission to the real economy is not affected but that credit spreads spike less strongly. Re-doing the analysis with survey forecast errors instead of belief distortions, allows me to inspect the expectation formation mechanism further. Here, we find the well-known pattern of *dynamic overshooting* (Angeletos, Huo and Sastry, 2021). Initially, agents underreact to the structural shock before overreaction takes place. These findings are robust to a number of specification choices.

The impulse response analysis points to two interesting facts. The dynamic responses of belief distortions has an *ex-ante* perspective, showing how agents' sentiment evolves in the future. Survey forecast errors, on the other hand, allow to examine the reaction of agents to news *ex-post*. So, the first exercise shows that agents evaluate the future too optimistically before turning pessimistic. The second exercise with survey forecast errors reveals that agents initially underreact but then overreact in their belief formation. It is reassuring that these responses are qualitatively similar, because this means that the ML expectations predict the realized crash risk reasonably well. This is also in accordance with the evidence in Krishnamurthy and Muir (2017) that credit spreads are generally too low prior to a financial crisis. Rational expectations models would need a long sequence of negative shocks or quite sizable ones, which we do not observe in reality. As Maxted (2023) shows theoretically, small financial shocks suffice in an environment of behavioral frictions. This aligns well with the outcomes of the counterfactual experiment, which reduces the impact response of credit spreads. I argue that agents do not realize the initial, smaller-sized shocks. Hence, they stay optimistic or underreact to this news. Only after the sentiment switches due to another shock (e.g., the bankruptcy of Lehman Brothers in the Great Financial Crisis), agents start to overreact to news and turn pessimistic.

The contribution of this paper is thus threefold. First, this paper shows that the belief formation process in credit spreads is not rational. Specifically, credit spreads under- and overreact to new information on the market. Second, I show that belief distortions on financial markets, measured through credit spreads, exist, are economically sizable, and in line with narrative evidence. Belief distortions are ex-ante expectational errors in beliefs and are constructed as the difference between machine efficient benchmark expectations and survey expectations. Third, I evaluate the responses of belief distortions and survey forecast errors conditional on a financial shock. This allows me to investigate the belief formation mechanism on financial markets in more detail.

The remainder of the paper proceeds as follows. The next section connects to the existing literature. In section 3, I investigate the belief formation on financial markets and explain the findings with a stylized model. In section 4, I construct the machine efficient benchmark and use these predictions to construct subsequently belief distortions. I also discuss the diagnostics of the constructed series. This is followed by the empirical exercise in section 5, where the dynamic responses to a financial shock are examined. Finally, section 6 concludes.

2. Related Literature

This paper contributes to three strands of the literature. First, it relates to the literature on the interaction between financial markets, the real economy, and the role of sentiments. While models based on financial frictions sucessfully explain the onset of a financial crisis and its broader transmission, these models require a long sequence of poor returns to trigger a shock strong enough to resemble a financial crisis and thus do not match the pre-crisis evidence. The starting point of this literature is the financial accelerator framework, which points to the important role of net worth and its inverse relationship to credit conditions (Bernanke, Gertler and Gilchrist, 1999). More recent approaches show that credit conditions are also affected by financial shocks on the enforcement constraint limiting the firm's ability to borrow (Jermann and Quadrini, 2012), by idiosyncratic uncertainty on investment payoffs leading to time-varying risk premia (Christiano, Motto and Rostagno, 2014), or by liquidity mismatches leading to run-like behavior associated with financial panics (Gertler, Kiyotaki and Prestipino, 2019). Akinci and Queralto (2022) explain financial crisis behavior in a non-linear model with endogenous equity issuance of financial intermediaries. Models taking sentiments into account can also succesfully match the pre-crisis evidence on low risk premia. Maxted (2023), for instance, incorporates behavioral frictions alongside financial frictions into a real business-cycle model.[1] The interaction of these frictions causes boom-bust cycles and thus belief-driven fluctuations. Similarly, the model by Krishnamurthy and Li (2020) allows sentiments along financial frictions to play a role in financial intermediation. They show that the model with only the frictional intermediation mechanism misses the pre-crisis behavior, while adding sentiments to the model resolves this conundrum. This paper offers a semi-

[1] See also the papers by Bianchi, Ilut and Saijo (forthcoming) and L'Huillier, Singh and Yoo (2021), which incorporate diagnostic expectations in leading business cycle frameworks.

structural approach and provides evidence that these mechanisms are active without relying on a fully-fledged structural model.

Second, the paper is related to a number of papers concerned with using survey data in semistructural macroeconomic models, in particular conditional on financial shocks. With respect to financial shocks, an array of papers provide different specifications and identification schemes. While Gilchrist and Zakrajšek (2012) introduce the EBP and provide an identification of unanticipated movements in this premium based on timing assumptions, Furlanetto, Ravazzolo and Sarferaz (2019) develop a sign-restriction approach to disentangle investment shocks (and other demand shocks) by imposing a co-movement in investment and stock prices. Barnichon, Matthes and Ziegenbein (2022), however, are interested in the non-linear transmission of financial shocks, while Boeck and Zörner (2019) investigate credit sentiment shocks in a non-linear framework. Caldara et al. (2016) utilizes a penalty function approach to disentangle uncertainty shocks and financial shocks. Regarding the use of survey data, an array of papers has investigated the macroeconomic consequences of undue expectations or over-optimism in growth (Enders, Kleemann and Müller, 2021; Benhima and Poilly, 2021; Beaudry and Willems, 2022). This paper investigates belief distortions conditional on a financial shock and extends the usual analysis in dimensions to explore belief formation in more detail.

Third, the paper contributes to the literature on imperfect information and why economic agents make systematic errors embedded in beliefs. These reasons include the presence of information frictions (Coibion and Gorodnichenko, 2015), the use of extrapolative expectations (e.g., De Long et al., 1990, Barberis, Shleifer and Vishny, 1998, or Barberis et al., 2015), the overweighting of personal experience (e.g., Malmendier and Nagel, 2011 and Malmendier and Nagel, 2016), the overreaction to incoming news (e.g., Bordalo, Gennaioli and Shleifer, 2018, Gennaioli and Shleifer, 2018, Bordalo et al., 2019, and Bordalo et al., 2020), the this-time-is-different thinking (Reinhart and Rogoff, 2009), or the use of simple heuristics to forecast (e.g., Anufriev and Hommes; Assenza et al., 2012; 2019). However, they all have in common that the presence of new information is given too much or too little weight. This happens because agents only have limited attention (neglecting the full information assumption) or new information is processed in a non-rational or behavioral way (neglecting the rational expectation assumption). Angeletos, Huo and Sastry (2021) reconcile those different streams and argue in favor of a pattern called *dynamic overshooting* – initial underreaction due to informational frictions is followed by overreaction due to extrapolation. This paper corroborates the findings of Bordalo et al. (2020) and Angeletos, Huo and Sastry (2021) but investigates the reaction to incoming news in credit spreads. Some initial evidence is already provided by Bordalo, Gennaioli and Shleifer (2018), which shows the predictability of forecast errors and revisions through the current credit spread but abstains from investigating the response to news through the regression framework of Coibion and Gorodnichenko (2015). The paper is closely related to the work of Bianchi, Ludvigson and Ma (2022) but differs in two important aspects. Methodologically, they provide a framework for belief distortions and use a regularized regression based approach to construct a machine efficient benchmark. This paper builds upon this approach but also utilizes non-linear modeling frameworks for the machine efficient benchmark. Empirically, they focus on belief distortions in inflation and output, while this paper focuses on credit spreads. Furthermore, this paper also investigates the dynamic interactions with key macroeconomic variables due to undue expectations.

3. Belief Formation on Financial Markets

This section provides empirical evidence on the belief formation process on financial markets and explains the evidence through a stylized model. Having established that financial market expectations do not process information rationally allows me to construct belief distortions in a next step. I start by testing the reaction of survey expectation to new information using the procedure in Coibion and Gorodnichenko (2015) but utilizing not only survey expectations data on the aggregate but also on the individual level (Bordalo et al., 2020). The stylized model of private and public signals with diagnostic expectations explains the findings and shows that belief distortions are related to future sentiment in response to news.

To characterize crash risk on financial markets, I am interested in the risk premium. Credit spreads are a natural choice for measuring risk premia. According to Elton et al. (2001), credit spreads differ across rating classes not only due to their risk premium, but also due to their expected default loss.[2] Additionally, a liquidity premium can arise in times of financial distress. To minimize the effect of the default premium, I use Moody's Aaa rated corporate bond yields.[3] For this yield, there is also evidence that the liquidity component did not rise during the subprime crisis (Dick-Nielsen, Feldhütter and Lando, 2012). I am also re-doing the analysis with Moody's Baa rated corporate bond yields to also gauge the effect of a higher liquidity premium. From these bond rates a long-term government yield of similar maturity is deducted to construct credit spreads.

[2] The third component of credit spreads, the tax premium, arises because interest payments on corporate bonds are differently taxed than those on government bonds, but this is disregarded in the analysis. Although they are an important influence in explaining credit spreads, due to their inability to explain differences in credit spreads they are not of concern in this setting.

[3] An interesting alternative is the excess bond premium (EBP) (Gilchrist and Zakrajšek, 2012), the residual of a micro-based approach to credit spreads freed from firm-specific information on default risk. Unfortunately, this is not suitable for the current framework due to unavailability of subjective expectations thereof.

Survey expectations on credit spreads come from the Blue Chip Financial Indicators. The survey is conducted on a monthly basis, asking around 40 panelists from major financial institutions about their expectations regarding several financial indicators. Data is taken from the end-of-quarter month survey in March, June, September, and December, which are released at the beginning of the month while the survey is conducted at the end of the previous month. Forecasts are available for the current quarter t and for quarters $t+1$ through $t+4$. The survey offers a cross-section of subjective forecasts of each panelists at each point in time. Specifically, I use the consensus (median) forecast to construct the Aaa and Baa spread. Data on the Aaa spread covers the period 1988Q1 to 2020Q1, while the time series is considerably shorter for the Baa spread spanning from 1999Q1 to 2020Q1. Data sources and the exact construction of the survey expectations of credit spreads are listed in Appendix A1.

Reaction to Incoming News

The empirical procedure of Coibion and Gorodnichenko (2015) and extended by Bordalo et al. (2020) works as follows. I denote the h-step ahead consensus forecast made at time t for the future value of y_{t+h} of a credit spread with $\mathbb{F}_t y_{t+h}$. The consensus forecast is constructed with $\mathbb{F}_t y_{t+h} = (1/I) \sum_i \mathbb{F}_{it} y_{t+h}$, where $\mathbb{F}_{it} y_{t+h}$ is the forecast of individual i and $I > 1$ is the number of forecasters. Forecast revisions at time t of individual i are defined as $FR_{it,h} = (\mathbb{F}_{it} y_{t+h} - \mathbb{F}_{it-1} y_{t+h})$ and $FR_{t,h} = (1/I) \sum_i FR_{it,h}$ follows likewise. Predictability of forecast errors is measured by estimating the following consensus regression

$$y_{t+h} - \mathbb{F}_t y_{t+h} = \beta_0^c + \beta_1^c \mathrm{FR}_{t,h} + \eta_{t+h}^c, \quad \eta_{t+h}^c \sim \mathcal{N}(0, \sigma_{c,\eta}^2). \tag{3.1}$$

If forecast errors are not predictable from forecast revisions, I cannot reject the null hypothesis of FIRE. This essentially reduces to testing whether $\beta_1 = 0$. Otherwise, overreaction (underreaction) is implied by a negative (positive) coefficient β_1. For instance, a positive coefficient $\beta_1 > 0$ together with a positive forecast revision, $\mathrm{FR}_{t,h} > 0$, implies that the consensus forecast is not optimistic enough. Bordalo et al. (2020) extend this analysis by also analyzing forecast error predictability at the individual level. They propose to estimate a pooled panel regression model,

$$y_{t+h} - \mathbb{F}_{it} y_{t+h} = \beta_0^p + \beta_1^p \mathrm{FR}_{it,h} + \eta_{t+h}^p, \quad \eta_{t+h}^p \sim \mathcal{N}(0, \sigma_{p,\eta}^2), \tag{3.2}$$

where the common coefficient β_1^p indicates whether the average forecaster under- or overreacts to their own information. Again, if $\beta_1^p = 0$ then FIRE cannot be rejected. Furthermore, they also suggest forecaster-by-

Table 1: Error-on-Revision Regressions.

	Consensus			Individual					
Variable	β_1^c (1)	SE (2)	Obs. (3)	β_1^p (4)	SE (5)	Obs. (6)	med(β_1^i) (7)	med(Obs.) (8)	I (9)
$h = 0$									
Aaa spread	0.00	0.15	129	−0.17	0.04	4,681	−0.15	24.0	124
Baa spread	0.12	0.17	85	0.04	0.04	2,691	−0.01	36.0	61
$h = 1$									
Aaa spread	0.12	0.11	128	−0.19	0.02	4,451	−0.21	26.0	114
Baa spread	0.28	0.06	84	0.02	0.02	2,553	0.02	37.0	57
$h = 2$									
Aaa spread	0.06	0.15	127	−0.21	0.02	4,309	−0.20	26.5	112
Baa spread	0.09	0.20	83	−0.06	0.02	2,487	−0.05	36.5	56
$h = 3$									
Aaa spread	0.08	0.19	126	−0.20	0.02	4,134	−0.24	26.5	108
Baa spread	−0.02	0.50	82	−0.17	0.03	2,403	−0.14	38.5	54

Notes: This table shows coefficients from forecast error on forecast revision regression. Column 1 to 6 show the coefficients of consensus time series regressions and individual level pooled panel regressions together with standard errors (SE) and number of observations (Obs.). Column 7-9 shows the median coefficients, median number of observations and number of forecasters (I) in forecaster-by-forecaster regressions. For consensus time series regressions and pooled panel regressions, standard errors are Newey-West with the automatic bandwidth selection procedure (Newey and West, 1994).

forecaster regressions,

$$y_{t+h} - \mathbb{F}_{it} y_{t+h} = \beta_0^i + \beta_1^i \mathrm{FR}_{it,h} + \eta_{t+h}^i, \quad \eta_{t+h}^i \sim \mathcal{N}(0, \sigma_{i,\eta}^2), \quad i = 1, \ldots, I. \tag{3.3}$$

This yields a distribution of individual coefficients β_1^i ($i = 1, \ldots, I$), where I focus on the median coefficient. Since this results in varying sample sizes for the estimation (due to the different lengths of different forecasters in the sample), I only keep forecasters with at least fifteen observations. Furthermore, I winsorize outliers.[4]

Results of the error-on-revision regressions are presented in Table 1. Looking at the coefficients from the consensus regression, $\beta_1^c > 0$ indicates underreaction with varying statistical power. On the contrary, coefficients from the pooled panel and individual level regression are consistently and precisely estimated negative, pointing to overreaction. These findings are similar to the one presented in Bordalo et al. (2020). This provides evidence to their explanation of individual overreaction, while the adjustment process is characterized by rigidity in the aggregate. Since the information is not yet received by their peers, the consensus forecast is slow to adapt. In particular, this form of rigidity only holds for both credit spreads when looking at shorter horizons and vanishes at longer ones. These findings also corroborate the findings of

[4] Here, I follow the approach taken by Bordalo et al. (2020). They exclude forecasts which are five interquartile ranges away from the median. In case there is no variation in the interquartile range, I apply the interquartile range of the previous period. This ensures consistency of the forecasts.

Angeletos, Huo and Sastry (2021) that *dynamic overshooting* is driving the dynamics. Taken at face value, this provides strong evidence that information is processed in a non-rational manner on financial markets. This is also consistent with the evidence presented in Bordalo, Gennaioli and Shleifer (2018) and Bordalo et al. (2019). While the former look also into credit spread expectations, they average across horizons and only use the consensus forecast. The latter provides evidence of stock market return expectations. Interestingly, these findings also square with experimental evidence on stock market return expectations. Kocher, Lucks and Schindler (2019) explain overpricing due to lack of traders' self-control transmitting into irrational exuberance in markets.

A Model of Private and Public Signals with Diagnostic Expectations

How can we explain both underreaction at the aggregate and overreaction at the individual level? This section presents a stylized model of private and public signals, similar as in Bianchi, Ludvigson and Ma (2022) but extended to include diagnostic expectations as in Bordalo et al. (2020). In the model, (professional) forecasters not only use publicly available information but also a private signal to form a prediction. The private signal contains a judgemental component. Suppose that forecasters observe a noisy, private signal s_{it} of the unknown state variable y_t,

$$s_{it} = y_t + \varepsilon_{it}, \qquad \varepsilon_{it} \sim \mathcal{N}(0, \sigma_\varepsilon^2), \tag{3.4}$$

where ε_{it} is i.i.d. distributed across forecasters and over time. This specification captures heterogeneity in information of individual forecasters. These individual forecasters also observe a vector of public signals $\boldsymbol{x}_t \sim \mathcal{N}(0, \sigma_x^2)$ with information up to time t.[5] Forecasters predict the current unknown state using the public information with $y_t = \boldsymbol{x}_t'\boldsymbol{\beta} + u_t$, where $\boldsymbol{\beta}$ is a set of parameters describing the statistical mapping from the $\boldsymbol{x}_t$ (public information) to $\boldsymbol{y}_t$ (unknown state). In the simplest case, the public signal evolves according to an AR(1) process, i.e., $\boldsymbol{x}_t = y_{t-1}$ and $\boldsymbol{\beta} = \rho$,

$$y_t = \rho y_{t-1} + u_t, \qquad u_t \sim \mathcal{N}(0, \sigma_u^2) \tag{3.5}$$

where the disturbance u_t is the unpredictable part of y_t and constitutes the structural shock. Furthermore, the disturbances u_t and ε_{it} are mutually uncorrelated. The AR(1) setting yields convenient closed-form solutions, although credit spreads may be better described by richer processes (e.g., using a VAR as in Coibion and

[5] Hence, it is pertinent that all information is in real-time up to time point t. This is the information set the forecaster has in principle access to.

Gorodnichenko, 2015 or with hump-shaped dynamics as in Fuster, Laibson and Mendel, 2010). Therefore, in the subsequent empirical analysis, I use higher-order autoregressive processes and include macro and financial factors formed from high-dimensional datasets. To sum up, forecasters do not directly observe the current value of the unknown state y_t but form a prediction by using a combination of both, the private and the public signal. Due to their private signal, judgement plays a nontrivial role. Ultimately, they are interested in a prediction of the unkown state $t + h$ periods ahead.

A rational forecaster updates her beliefs via Bayesian updating. The conditional expectations are summarized by a probability density $f(y_t \mid S_{it-1})$, where S_{it-1} denotes the entire history of signals observed by forecaster i. In each period, the forecaster observes a new signal s_{it} and updates her estimates of the current state using Bayes rule,

$$f(y_t \mid S_{it}) = \frac{f(y_t \mid S_{it-1}) \times f(s_{it} \mid y_t)}{f(s_{it})}. \tag{3.6}$$

Given that the shocks are Gaussian distributed, this yields the Kalman filter. A rational forecaster estimates the current state with $y_{it|t} = \int y f(y \mid S_{it}) dy$ and forecasts future values using the AR(1) structure, thus $y_{it+h|t} = \rho^h y_{it|t}$.

Belief formation of the agents is prone to behavioral fallacies. Bordalo, Gennaioli and Shleifer (2018) introduce the framework of diagnostic expectations, which accounts for judgement biases. In this belief formation mechanism, agents' beliefs are distorted by the representativeness heuristic. According to this heuristic, a certain attribute is judged to be excessively common in a population when that attribute is diagnostic for the population. It is diagnostic for the population, if this attribute occurs more frequently in the given population than in a relevant reference population. Following Bordalo et al. (2020), the overweighting of representative states is described by the distorted probability distribution

$$f^\theta(y_t \mid S_{it}) = f(y_t \mid S_{it}) \left(\frac{f(y_t \mid S_{it})}{f(y_t \mid S_{it-1} \cup \{y_{it|t-1}\})} \right)^\theta \frac{1}{Z_t}, \tag{3.7}$$

where Z_t is a normalization factor that ensures that the distribution integrates to one. The first term in the equation reflects to rational expectations done via Bayesian updating, as discussed before. The second term describes the belief component. This means that the state y_t is more representative or diagnostic at time t if the signal s_{it} received in this period raises the probability of that state relative to the case where the news equals the ex ante forecast, $s_{it} = y_{t|t-1}$. Furthermore, the parameter θ measures the severity of judging according to the representativeness heuristic. If $\theta = 0$, agents are fully rational. If $\theta > 0$, memory is limited.

Subsequently, the distorted distribution $f^{\theta}(y_t \mid s_{it})$ is inflated by the likelihood of representative states that come to mind quickly, while it deflates the likelihood of nonrepresentative states.

Hence, I make use of slightly adapted versions of proposition 1 and 2 from Bordalo et al. (2020) in the following. The propositions and proofs are in Appendix B. According to proposition 1, the distorted time-varying mean is

$$\mathbb{F}_{it} y_t = y_{it|t} + \theta K \left(s_{it} - y_{it|t-1} \right), \tag{3.8}$$

where $K = \frac{\Sigma}{\Sigma + \sigma_\varepsilon^2}$ denotes the Kalman filter and Σ is constant over time and refers to the steady state variance of the updating step in the Kalman filter. Here, $y_{it|t}$ refers to the rational update of the hidden state implied by the Kalman filter, while the term in the parenthesis refers to the belief component. Specifically, a negative relatization of ε_{it} leads to a lower value of s_{it} and a possibly larger news component if the rational prediction is quite different. The news component is the difference between the private signal and the rational forecast in $t-1$, which is diagnostic of the state. Hence, an individual forecaster is updating past rational beliefs $y_{it|t}$ with the news component. If $\theta > 0$, agents deviate from Bayesian updating and overreact to news. This means that the news component is overweighted when constructing expectations. Interestingly, also the signal-to-noise ratio plays a critical role here. If the private signal is extremely diffuse ($\sigma_\varepsilon^2 \to \infty$), the news component (and thus behavioral frictions) is used to a continuously decreasing extent.

Can this model explain the evidence presented earlier? Remember that the error-on-revision regressions in Table 1 display consistently positive coefficients at the consensus level, while the sign switches at the individual level. These coefficients point to underreaction in the aggregate and to overreaction of the individual forecaster. By using the second proposition of Bordalo et al. (2020), the sign of the regression coefficients at the consensus and individual level can be determined using the diagnostic Kalman filter. In particular, as long as signals are not too overweighted ($\theta \in (0, \sigma_\varepsilon^2/\Sigma)$) relative to the dispersion of signals, the coefficient at the consensus level is positive. Again, if the private signal is extreme diffuse ($\sigma_\varepsilon^2 \to \infty$), the extent of overreaction can be more pronounced and agents still underreact to *very* noisy information. Furthermore, the individual level coefficient is unambiguously negative. Then, individual forecasters overreact to incoming news while the diagnostic filter entails rigidity in consensus beliefs.

This shows that a model with noisy information and diagnostic expectations can square the evidence presented so far. In a next step, I explore belief distortions in this model setup. Assume that one is interested in the $t+h$-step ahead prediction of credit spreads. Then, agents use the subjective expectation formation mechanism $\mathbb{F}_t y_{it+h} = \rho^h \mathbb{F}_t y_{it|t}$. We still assume a simple AR(1) structure, while this can be extended to more complicated processes. Hence, the consensus diagnostic expectation at time t is given by

$$\mathbb{F}_t y_{t+h} = \rho^h \int \mathbb{F}_{it} y_t di = y_{t+h|t} + \theta K \left(y_{t+h|t} - y_{t+h|t-1} \right). \tag{3.9}$$

This equation shows the distorted consensus estimate for the current state. Aggregation implies that $y_{t+h|t}$ is the rational forecast of the hidden state, while $y_{t+h|t} - y_{t+h|t-1}$ is the news component. Note that the news component may be structural shock, such as a financial shock in the present case. In the aggregate, news is combined with the rational forecast of the hidden state according to the Kalman filter, which incorporates individual level information dispersion. Furthermore, the parameter θ governs the extent of overreaction due to the representativeness heuristic.

As previously introduced, belief distortions are defined as the difference between survey and FIRE forecasts (Bianchi, Ludvigson and Ma, 2022). Note that $y_{t+h|t}$ is the rational update of the hidden state implied by the Kalman filter. Note, though, that the term $\mathbb{E}_t y_{t+h}$ in the model equation truly represents the FIRE forecast and is not subject to any of the biases. I assume that the machine efficient benchmark recovers the FIRE forecast. Hence, belief distortions are defined as follows

$$Z_t = \mathbb{E}_t y_{t+h} - \mathbb{F}_t y_{t+h} = \underbrace{\mathbb{E}_t y_{t+h} - y_{t+h|t}}_{\text{bias}} - \underbrace{\theta K \left(y_{t+h|t} - y_{t+h|t-1} \right)}_{\text{reaction-to-news}}. \tag{3.10}$$

Belief distortions are thus a composite of two terms: a *bias* term and an *reaction-to-news* term. If the machine benchmark truly recovers the true, underlying model (i.e., FIRE forecast) the bias is zero. More importantly, the second term captures the news component, $y_{t+h|t} - y_{t+h|t-1}$, which denotes underreaction as long as the diagnosticity parameter θ is sufficiently small. Again, the signal-to-noise ratio plays a role when judging with representativeness. In case, agents observe the signal perfectly ($\sigma_\varepsilon^2 \to 0$), and are not prone to behavioral judgment ($\theta = 0$), the second term drops out. This motivates the dynamic analysis in which I examine the responses of belief distortions to news (i.e., a structural shock).

This results in two predictions regarding belief distortions: (i) on average belief distortions are zero (*no bias*), and (ii) belief distortions are potentially large in periods with a high amount of incoming information or structural shocks. These shocks are then amplified if agents judge according to representativeness ($\theta > 0$) and characterize undue expectations. The first prediction motivates the forecasting exercise, because the machine efficient benchmark aims to resemble FIRE forecasts, which drives a potential bias towards zero.

The second prediction, motivates the macroeconomic analysis: If belief formation is rational, the second term should not react at all and belief distortions are zero in the presence of structural shocks.

4. Measuring Belief Distortions

This section intends to measure belief distortions on financial markets and is mainly concerned with the construction of the machine efficient benchmark. If belief distortions arise systematically, this suggests that financial markets do not work efficiently due to the presence of financial anomalies.[6]

This is best illustrated by Figure 1, which plots both actual credit spreads (black) and survey expectations (blue) at the time of the financial crises. Both credit spread series start to increase gradually prior to the Great Financial Crisis (GFC). A large increase in credit spreads is a strong indication that the economy transitions into a financial crisis, which is consistent with the findings of Krishnamurthy and Muir (2017). Survey expectations follow the overall trend but adapt sluggishly and are still optimistic in the run-up to the crisis, neglecting tail risks. Credit spreads jump disproportionately after the Lehman collapse on September 15. Survey expectations, however, do not react alike, because economic agents stay optimistic despite the surge in risk premia (and thus underreact to news). If we replace actual credit spreads with its machine efficient prediction, the sluggish adjustment of survey expectations gives rise to positive belief distortions. The reverse is happening after the GFC, where survey expectations are still elevated while financial risk is already falling again. Belief distortions arise because of a non-rational belief formation mechanism, which keep agents optimistic and pessimistic for too long. As is clear from the stylized model, agents individually overreact to incoming news (e.g., the collapse of Lehman) due to financial shocks. In periods of belief mismatches, belief distortions in credit spreads arise. These observations, together with the earlier presented evidence on belief formation, make clear that agents produce systematic errors in beliefs.

Belief distortions, denoted as Z_t, are defined as follows

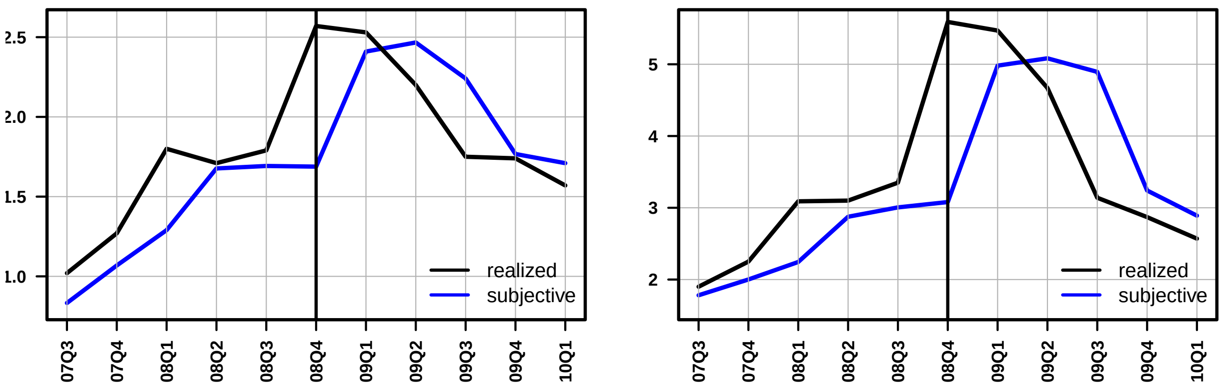
$$Z_t = \mathbb{E}_t y_{t+h} - \mathbb{F}_t y_{t+h}, \tag{4.1}$$

where y_{t+h} refers to the h-step ahead credit spread under consideration (either the Aaa or Baa credit spread), t indicates the time period. $\mathbb{F}_t$ refers to the survey expectations operator, while $\mathbb{E}_t$ refers to the FIRE operator. The resulting difference is denoted as a distortion in beliefs at time t. In the following, the paper creates a

[6] I follow here Brav and Heaton (2002, p. 575) in defining a *financial anomaly* as "a documented pattern of price behavior that is inconsistent with the predictions of traditional efficient markets, rational expectations asset pricing theory."

Figure 1: Credit Spreads and Subjective Expectations.

(a) Aaa Spread. **(b)** Baa Spread.



Notes: Credit spreads of Aaa and Baa rated bond yields to 10-year government bond yield and subjective survey expectations around the GFC. Black horizontal line denotes the 2008Q4, the height of the financial crisis. Data comes from FRED and the Blue Chip Financial Indicators.

machine efficient benchmark to construct FIRE forecasts. The paper develops an algorithm, which provides a new out-of-sample prediction each time new information is added to the model. Once a prediction is formed, belief distortions are constructed. Two steps are important to construct an machine efficient benchmark. First, I have to choose an appropriate loss function to select the best performing prediction out of a pool of prediction values. Second, the involved datasets have to be introduced, which resemble the potential information set of professional forecasters.

Creating a Machine Benchmark

This leads to the construction of the *machine benchmark* to identify possible distortions in beliefs. Let $\mathbb{E}_t y_{t+h}$ denote either a forecast of the Aaa or Baa credit spread at horizon $h \geq 1$ predicted at time t. It is thus imperative that the model of objective expectations for constructing $\mathbb{E}_t y_{t+h}$ be as rich as possible in information so that the measure does not miss pertinent information. Concurrently, the model has to be parsimonious to avoid spurious estimates.

To address these issues, I follow recent advances in the machine learning literature for macroeconomic forecasting (Goulet Coulombe et al., 2022). The recommendations are as follows: (i) caring about non-linearities, (ii) using a factor model for dimensionality reduction, and (iii) performing cross-validation. Hence, I proceed as follows. First, I use two very high-dimensional datasets: a real-time macroeconomic dataset and a monthly financial dataset.[7] From these datasets, I separately construct diffusion indices with a dynamic factor model. Second, I perform a horse-race of different specified linear and non-linear models and estimators to enhance predictive performance. Specifically, I use the following estimators: Bayesian linear regression with horseshoe (HS) shrinkage prior (Carvalho, Polson and Scott, 2010), the elastic net (EN, Zou and Hastie, 2005), and Bayesian Additive Regression Trees (BART, Chipman, George and McCulloch, 2010). Third, additional model uncertainty is taken into account by different specifications of the model.

Before going into detail about the forecasting exercise, I introduce both datasets, which are in real-time and constructed to resemble the *potential* information set of a professional forecaster. The real-time macroeconomic and financial is obtained from the Philadelphia Fed's Real-Time Dataset and the FRED. The resulting real-time macroeconomic indicator dataset contains observations on 45 variables. Then, 31 financial variables are added to the dataset, yielding a real-time dataset of 76 variables. The real-time macroeconomic and financial dataset provides information on output and income, consumption and investment, trade and government, money and prices, labor market and housing, and interest rates and spreads. Additionally, I also use monthly financial data to pick up information from financial markets on a higher frequency. The dataset consists of 147 monthly financial series measuring the behavior of a broad cross-section of asset returns through valuation ratios, such as the price-earnings ratio or the dividend-price ratio, dividends and prices, risk factors, and a broad class of different portfolios. This dataset is a version of the financial dataset used in Jurado, Ludvigson and Ng (2015) to construct the financial uncertainty index and is also used in the analysis of Bianchi, Ludvigson and Ma (2022) to create a machine benchmark to measure macroeconomic belief distortions. The complete list of variables of both datasets is given in the Appendix A3 and Appendix A4.

I also have to clarify how true out-of-sample forecasts are constructed. At each time point t, I have a total of $\tilde{T} = \tilde{T}_T + \tilde{T}_V$ observations available and want to construct a true out-of-sample one-step ahead prediction for $\tilde{T} + 1$. Furthermore, I use a set of different models $\mathcal{M} = \{\mathcal{M}_1, \ldots, \mathcal{M}_m\}$ of cardinality m. How do I decide on which model $\mathcal{M}_i \in \mathcal{M}$ to use? I simply choose the best-forecasting model according to a loss function in a pseudo out-of-sample model validation exercise. This provides me with an *ex-ante* criterion for model selection. For that, I run each model with a training sample of size $\tilde{T}_T$ and construct a one-step ahead prediction for the validation sample of size $\tilde{T}_V = 1$. The validation sample is always the most recent datapoint

[7] Before the computation of the factors, the monthly financial dataset is transformed to a quarterly dataset. In each quarter, the mid-of-quarter observation is used to align the information set of the machine as close as possible to the information set of the individual forecaster. Survey expectations are released on the first day of the end-of-quarter month while the surveys are conducted a couple of days earlier.

in the overall sample available at time t. As a loss function, I choose the root mean-squared error (RMSE). A symmetric loss function can be questioned as forecasters may have asymmetric loss functions (Capistrán and Timmermann, 2009). Since survey forecast errors are close to zero this suggests that economic agents do not systematically over- or underassess financial risk. Then, I re-run the best-fitting model to construct a true out-of-sample forecast.

In its most general form, I denote the forecasting model as

$$y_{t+h} = f(\mathcal{Z}_t) + \varepsilon_{t+h}, \qquad \varepsilon_{t+h} \sim \mathcal{N}(0, \sigma^2_{t+h}), \qquad h \geq 1 \tag{4.2}$$

where the $K \times 1$ vector $\mathcal{Z}_t$ denotes the information set of the model. I use three different estimators, two linear and a non-linear one. These look as follows

$$\begin{aligned}
\texttt{HS:} \qquad & f(\mathcal{Z}_t) = \mathcal{Z}_t' \boldsymbol{B}_h, \\
\texttt{EN:} \qquad & f(\mathcal{Z}_t) = \mathcal{Z}_t' \boldsymbol{B}_h, \\
\texttt{BART:} \qquad & f(\mathcal{Z}_t) = \sum_{s=1}^{S} g_s(\mathcal{Z}_t | \boldsymbol{\mathcal{T}}_s, \boldsymbol{\mu}_s),
\end{aligned} \tag{4.3}$$

where both linear estimators estimate the coefficients summarized in $\boldsymbol{B}_h$, while the BART model is composed of S distinct regression trees, each equipped with a single regression tree function g_s, the corresponding tree structure $\mathcal{T}_s$, and the parameters at the terminal nodes μ_s. In what follows, and in consistency with Chipman, George and McCulloch (2010), I cross-validate over $S = \{50, 250\}$ in all our empirical applications. In the baseline specification, $\mathcal{Z}_t = (1, y_t, \ldots, y_{t-j+1}, \boldsymbol{G}_t, \ldots, \boldsymbol{G}_{t-j+1})$ includes an intercept and up to p lags of both the endogenous variable and additional factor information in an r_G-dimensional vector $\boldsymbol{G}_t$. In particular, this vector contains factors from the two involved datasets $\boldsymbol{G}_t = (\boldsymbol{f}^M_{t,final}, \boldsymbol{f}^F_{t,final})$, where $\boldsymbol{f}^M_{t,final}$ constitutes factors from the real-time macro dataset and $\boldsymbol{f}^F_{t,final}$ denotes factors from the monthly financial dataset. The vector of factors also contains non-linear versions, respectively. Last, ε_{t+h} denote Gaussian innovations of the model. For the exact construction of the factors and the estimation procedures, Appendix C provides the details.

The Bayesian linear regression specification conveniently nests all model specification that are used in the forecasting exercise. In particular, it nests the random-walk (`RW`), a wide variety of autoregressive (`AR`) processes, and the full specification is an autoregressive distributed lag (`ADL`) model. When using the EN or BART, I use the full information set. In all specifications, I fix the number of factors $r_G = 10$, allowing eight linear factors and two non-linear versions to enter the specification for each dataset. Furthermore, I vary the amount of lags $p = \{1, 2\}$. The Bayesian linear regression may also feature a stochastic volatility specification in the second-moment. This yields in total $m = 13$ different models in the set of models $\mathcal{M}$.[8] The exact model specifications and the description of the different estimators is suspended to Appendix C.

In all cases, I compute one-step up to four-step ahead forecasts in a direct forecasting fashion. For both spreads, the sample size of the training and validation sample is kept constant at $\tilde{T} = 67$ quarters across the forecasting exercise. This yields an external evaluation sample from 1988Q1 to 2020Q1 of length $T_{aaa} = 129$ for the Aaa credit spread, while the true out-of-sample forecasts are recorded from 1999Q1 to 2020Q1 with a length of $T_{baa} = 85$ for the Baa credit spread. The external evaluation sample is determined by the availability of survey forecasts. Hence, the first training and validation sample starts in 1971Q1 (Aaa credit spread) and 1982Q1 (Baa credit spread). Bayesian estimations are based on 25,000 draws from the posterior distribution, where I discard the first 10,000 draws and use a thinning factor of 5. This leaves me with 5,000 draws from the respective posterior distribution. Furthermore, cross-validation is implemented for both machine learning techniques, the EN and BART. Specifically, cross-validation is not implemented in the usual K-fold fashion but in a *timeslice* fashion. Contrary to K-fold cross-validation, which draws random subsamples from the total sample to construct training and estimation sample, the *timeslice* validation procedure creates sample splits, which takes the inherent structure of time-series (dependency over time) of the observations into account and constructs sample splits going forward until the last observation in the (training) sample.

Additionally, I also explore an alternative, possibly more interesting, specification. In this specification, I directly control for the subjective forecaster's information set by controlling for the consensus survey forecast. Hence, this specification allows me to infer whether the machine really improves upon the survey expectation. Nevertheless, this comes against a cost. It drastically restricts the sample due to the rather short time series of survey expectations. Hence, the training and validation sample is only $\tilde{T}_{alt} = 40$ and start in 1995Q2 for the Aaa credit spread and only in 2010Q1 for the Baa credit spread.. The external evaluation sample thus reduces to $T_{aaa,alt} = 100$ and $T_{baa,alt} = 41$.[9] The specification of $\mathcal{Z}_t = (1, \mathbb{F}_t^{(k)} y_{t+h}, y_t, \ldots, y_{t-j+1}, \boldsymbol{G}_t, \ldots, \boldsymbol{G}_{t-j+1}, \boldsymbol{W}_t, \ldots, \boldsymbol{W}_{t-j+1})$ includes also the kth percentile of the survey forecast distribution. The superscript (k) denotes the coefficients corresponding to adding the k-th percentile of the survey forecast distribution to the specification. Furthermore, an r_W-dimensional vector $\boldsymbol{W}_t$

[8] The results also hold when the model space is increased significantly, by allowing for more lags and including different sets of covariates from the baseline model.

[9] While the sample length is definitely too short for the Baa credit spread expectations, I provide robustness with respect to the Aaa credit spread belief distortions of this specification in the macroeconomic model. See also the discussion in the sensitivity checks in subsection 5.6.

Table 2: Forecasting Evaluation.

	h=1		h=2		h=3		h=4	
	Aaa spread	Baa spread	Aaa spread	Baa spread	Aaa spread	Baa spread	Aaa spread	Baa spread
	Forecasting Performance							
$\text{RMSE}_{\mathbb{F}}$	0.230	0.449	0.296	0.615	0.343	0.720	0.386	0.788
$\text{RMSE}_{\mathbb{E}}$	0.193	0.364	0.278	0.535	0.309	0.772	0.369	0.852
min [RMSE]	0.192	0.370	0.287	0.605	0.322	0.632	0.343	0.696
	Ratios							
$\text{RMSE}_{\mathbb{E}}/\text{RMSE}_{\mathbb{F}}$	0.837	0.810	0.941	0.870	0.901	1.073	0.957	1.081
$\text{RMSE}_{\mathbb{E}}$/min [RMSE]	1.005	0.984	0.968	0.885	0.959	1.222	1.076	1.225
	Alternative Specification							
$\text{RMSE}_{\mathbb{F}}$	0.240	0.268	0.315	0.161	0.362	0.192	0.408	0.334
$\text{RMSE}_{\mathbb{E}}$	0.202	0.301	0.284	0.366	0.356	0.37	0.376	0.53
$\text{RMSE}_{\mathbb{E}}/\text{RMSE}_{\mathbb{F}}$	0.844	1.121	0.901	2.273	0.983	1.927	0.922	1.585
$\text{RMSE}_{\mathbb{E}}$/min [RMSE]	1.028	1.082	0.981	0.927	1.132	0.948	1.053	1.127
	Best-Fitting Models (pseudo)							
`RW`	17	20	35	22	41	29	41	28
`AR`	29	14	22	16	20	12	20	17
`ADL`	29	23	23	15	21	10	26	13
`EN`	19	16	23	13	23	17	26	13
`BART`	35	12	26	19	24	17	16	14

Notes: Forecasting evaluation via RMSEs of objective and subjective expectations. The table provides RMSE on forecasting performance, and RMSEs ratios. The alternative specification additionally controls for subjective expectations. It also provides which are the best-fitting models in the pseudo out-of-sample exercises, where the following models are considered: `RW` - random walk, `AR` - autoregressive model, `ADL` - autoregressive distributed lag model, `EN` - Elastic Net, `BART` - Bayesian Additive Regression Trees.

of additional information is added to the specification. This vector controls for past survey expectations and various distributional quantities of the survey forecast distribution. The details of the specification are to be found in Appendix C.

For the evaluation of the forecasts, I use RMSEs of the forecast errors of both, objective and subjective, expectations, as follows

$$\text{RMSE}_{\mathbb{F}} = (1/T_i) \sum_{t=1}^{T_i} \sqrt{(\mathbb{F}_t y_{t+h} - y_{t+h})^2}, \quad \forall h, \tag{4.4}$$

$$\text{RMSE}_{\mathbb{E}} = (1/T_i) \sum_{t=1}^{T_i} \sqrt{(\mathbb{E}_t y_{t+h} - y_{t+h})^2}, \quad \forall h. \tag{4.5}$$

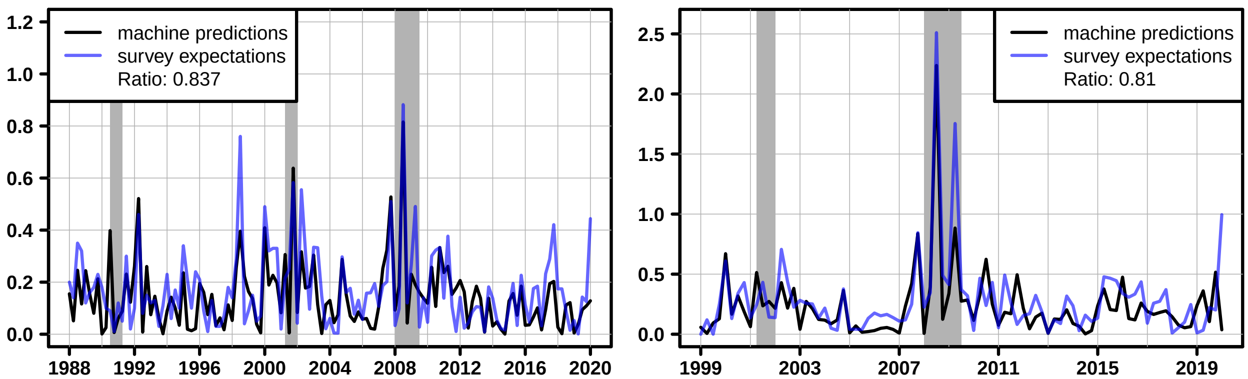
These statistics are presented in Table 2. The table presents RMSE of the survey and ML expectations and the minimum RMSE of the best-forecasting model (from an *ex-post* perspective) along with ratios of these RMSEs. The same outcomes are also provided for the alternative specification, which I discuss further below. Last, the table also shows the best-fitting models in the pseudo out-of-sample exercise, which is the *ex-ante* criterion for constructing objective expectations. The outcomes of this exercise reveal the following. First, it is reassuring that the machine works better than the survey forecasts, as can be directly seen from the ratio $\mathrm{RMSE}_{\mathbb{E}}/\mathrm{RMSE}_{\mathbb{F}}$. For instance, the ratio for the one-step ahead Aaa credit spread is 0.837 and for the Baa credit spread 0.810. Interestingly, the ratio is rather stable across both credit spreads. Only at longer horizons, the survey forecasts of the Baa credit spread are on par with the machine efficient benchmark. Second, the table also reveals that the *ex-ante* out-of-sample forecasting procedure performs well against the *ex-post* best-forecasting model. This is visible from the ratio $\mathrm{RMSE}_{\mathbb{E}}/\min[\mathrm{RMSE}]$, where $\min[\mathrm{RMSE}]$ is the minimum of the RMSEs of all considered models *ex-post*. The RMSEs of all models are shown in Table H1 in Appendix H, where the minimum is indicated with bold numbers. The table in the appendix reveals that autoregressive processes perform well for short forecasting horizons, while BART dominates longer-term forecasts. However, these RMSEs are computed with an informational advantage. Third, the table also reveals the number how often a particular model was the best-fitting model in the pseudo out-of-sample forecasts. This highlights the inherent model uncertainty in the estimation problem, because each model class is used quite frequently.

The results of the alternative specification are also presented in Table 2. Here, I use the $k = 50$th percentile of the distribution. A few interesting things have to be noted. Generally, $\mathrm{RMSE}_{\mathbb{F}}$ are slightly higher for the Aaa credit spread and much lower for the Baa credit spread. However, the machine benchmark offers RMSE in the ballpark as in the full specification. Hence, the ratio is below unity for the Aaa spread and above unity for the Baa spread. These changes are explained through the dynamics in the RMSE, which results from the rather short external evaluation sample for the Baa spread. This external evaluation sample also does not cover the GFC but only starts in 2010Q2. Therefore, this sample period covering the aftermath of the GFC until the present may not be ideal to measure belief distortions. Overall, evidence suggest that belief distortions still persist while controlling for the forecaster's information set.

In Figure 2, I want to further shed light on the forecasting performance. It shows a time-series plot of the RMSEs of both survey and ML expectations of the respective credit spread for the one-step ahead forecast. First, survey RMSEs lay on top of the machine RMSEs almost throughout the sample. This is further backed by the ratio of the RMSEs, which is clearly below unity. Second, forecasting performance decreases in crisis times, indicated by the NBER recession dates. This holds specifically for the GFC. Third, and most importantly, the survey RMSE spikes in crisis times further upwards than the machine RMSE. This is clear

Figure 2: Forecasting Evaluation.

(a) Aaa spread. **(b)** Baa spread.



Notes: RMSEs of out-of-sample forecasts of objective (black) and subjective (blue) expectations. Gray bars indicate the NBER recession dates.

evidence that survey participants forecast on average worse than the machine benchmark in a crisis, which gives rise to belief distortions.

Lastly, I also check whether the machine predictions do react to news by estimating error-on-revision regressions. Results in Table 3 present the results of this exercise. While I do find very small and insignificant coefficients for the Aaa spread along all horizons and the Baa spread for $h = 1$, coefficients for the Baa spread at longer horizons are larger in absolute value and statistically significant for $h = 2$. Hence, ML expectations do neither for the Aaa spread nor for the Baa spread under- or overreact at short horizons. At longer horizons, the picture is more diffuse for the Baa spread.

To sum up, creating a machine benchmark pays off. Generally, credit spreads are extremely forward-looking variables and commonly used as recession indicators. Hence, it comes as no surprise that predictive performance worsens in times of crisis. Furthermore, the strong mean-reverting behavior in credit spreads explains that higher horizon forecasts are more on par with the machine benchmark. In the short run, however, using the machine leads to considerable forecasting improvements against the survey forecasts. In this exercise, I only consider *point* forecasts to measure belief distortions. The objective of this exercise is to create a *machine benchmark* of an unbiased point forecast. Other aspects of forecasting, e.g., predictive accuracy, is of minor importance here. The loss function in Equation (4.4) and Equation (4.5) takes only point predictions into account.

Table 3: Error-on-Revision Regressions with Machine Predictions.

	h=1		*h=2*		*h=3*	
	β_1^z	SE	β_1^z	SE	β_1^z	SE
Variable	(1)	(2)	(3)	(4)	(5)	(6)
Aaa spread	−0.022	0.062	−0.034	0.114	−0.010	0.141
Baa spread	−0.018	0.081	−0.085	0.026	0.215	0.217

Notes: This table shows coefficients from forecast error on forecast revision regression with machine predictions. Standard errors are Newey-West with the automatic bandwidth selection procedure (Newey and West, 1994).

Construction of the Belief Distortion Series

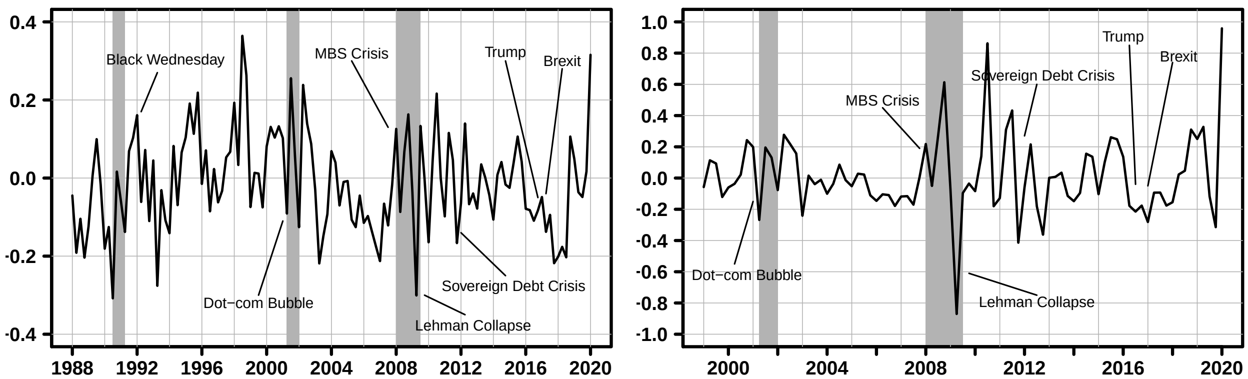
After the evaluation of the forecasting performance of both survey and ML expectations, I can construct a quarterly series of belief distortions as defined in Equation (4.1). For the construction of the belief distortion series, I use the difference between the ML and survey expectations. While I construct the difference between the survey and ML expectation at each horizon, I focus in the main analysis on the one-step ahead belief distortion shown in Figure 3.

The belief distortion series matches well with narrative evidence for key historical episodes. For instance, both series exhibit positive belief distortions on the onset of the GFC. In particular, I observe a positive spike at the beginning of the crisis in mid-2007 after the bursting of the mortgage-backed security bubble, and a negative spike at its climax after the bankruptcy of Lehman Brothers. As discussed earlier, in both instances economic agents individually overreact but underreact in the aggregate, resulting in a spike of the series. While economic agents were optimistic for too long (resulting in a lower expectation of the credit spread) in the beginning, the were also pessimistic for too long at the end of the crisis. Specifically, at the beginning of the crisis period sentiment stays elevated for quite some time before turning extremely negative. This fits into the story presented in Bordalo, Gennaioli and Shleifer (2018) that a currently high spread is indicative that future spreads are too high. The Lehman bankruptcy *per se* did not cause macroeconomic troubles (financial risk was already elevated before), but economic agents were uncertain about the signal about financial risk.[10] Higher dispersion in beliefs allows for possibly stronger ex-ante expectational errors which strongly affects belief formation. In comparison, the Dot-com bubble led to much smaller distortions in beliefs. Other key

[10] It is not shown here, but disagreement (defined as the standard deviation) among forecasters about credit spreads have risen strongly in the crisis.

Figure 3: Belief Distortions.

(a) Aaa Belief Distortions. **(b)** Baa Belief Distortions.



Notes: Belief distortions in the Aaa and Baa risk spreads. Gray bars indicate the NBER recession dates. Key historical episodes are delineated in the figure.

historical events, such as the Sovereign Debt Crisis in Europe, the election of Trump, or the referendum on Brexit have not led to a considerable deviation in belief distortions.

Besides the narrative assessment, I also perform some simple diagnostic checks of the validity of the series for measuring belief distortions. Results can be found in Appendix F. Both series fluctuate around zero, presenting no evidence of a bias in belief distortions. From a statistical standpoint, the belief distortion series show no evidence of autocorrelation or predictability. Regarding the latter, I perform a series of Granger-causality tests in Table F1, which test whether the belief distortion series can be forecasted by a set of macroeconomic variables. There is no evidence of any power in forecasting both belief distortion series. I also examine correlations to other structural shocks from the literature in Figure F2. Specifically, I compare the belief distortions series to financial shocks, uncertainty indicators, and macroeconomic shocks. To start with the latter, correlations to other macroeconomic shocks (monetary, fiscal, oil) are low and statistically not significant. On the contrary, correlations are positive and statistically significant for various uncertainty indicators and financial shocks. For instance, correlations go up to $\rho \approx 0.35$ for the Aaa belief distortions series, and are even higher to the Baa belief distortion series with $\rho \approx 0.60$. However, the numbers presented here are only unconditional correlations. This motivates the next step of the analysis, in which I examine the conditional response of belief distortions to a financial shock.

5. Dynamic Responses to a Financial Shock

In this section, I analyze the dynamic responses of belief distortions conditional on a financial shock. For the estimation and identification of financial shocks, I follow the approach of Gilchrist and Zakrajšek (2012). Their identification of financial shocks rests on the excess bond premium (EBP), which represents the cyclical changes between measured default risk and credit spreads. Specifically, identification is achieved via timing assumptions. In the baseline specification eight variables are included in the following order: (i) log real gross domestic product (GDP), (ii) log real consumption, (iii) real investment, (iv) prices measured with the log GDP deflator, (v) the EBP, (vi) the log S&P 500 as a stock market index, (vii) the effective (nominal) federal funds rate, and (viii) the belief distortion series. The identifying assumption implied by this recursive ordering is that innovations to the EBP affect economic activity and prices only with a lag, while the stock market, the policy rate, and belief distortions can react contemporaneously to the financial shock. For the current analysis, there are numerous appealing facts of this shock. First, it is widely accepted as unanticipated movements in financial frictions. Second, it drives a significant component of the business-cycle variation in macroeconomic activity and third, this shock is relevant for the belief formation on financial markets.

The sample is limited due to the length of the belief distortion series. The sample including the Aaa spread spans from 1988Q1 to 2020Q1, while the one including the Baa spread spans from 1999Q1 to 2020Q1. The reason for using quarterly data is due to the nature of the survey forecasts for computing the belief distortions. Examining both dynamic responses - Aaa and Baa belief distortions – allows me to investigate belief distortions of credit spreads with potentially higher default or liquidity premia. In a robustness check, I investigate whether the different samples influence the outcomes and show that this is not the case. The Bayesian VAR is then estimated over the respective sample period, using four lags of each endogenous variable and a constant as a deterministic term. A detailed overview on the data, the exact construction, and its sources can be found in Appendix A.

Econometric Approach

Let $\{\boldsymbol{y}_t\}_{t=1}^{T}$ denote an M-dimensional time series process. Consider the following reduced-form VAR(p) model

$$\boldsymbol{y}_t = \boldsymbol{c} + \boldsymbol{A}_1\boldsymbol{y}_{t-1} + \ldots + \boldsymbol{A}_p\boldsymbol{y}_{t-p} + \boldsymbol{u}_t, \quad \boldsymbol{u}_t \sim \mathcal{N}_M(\boldsymbol{0}, \boldsymbol{\Sigma}_t), \tag{5.1}$$

where p is the lag order, $\boldsymbol{c}$ is an $M \times 1$ vector of constants, $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_p$ are $M \times M$ coefficient matrices, and $\boldsymbol{u}_t$ denotes an $M \times 1$ vector of reduced-form Gaussian distributed innovations with a time-varying covariance matrix $\boldsymbol{\Sigma}_t$. For the details on the estimation of the VAR, see Appendix D.

For the identification of the VAR, I rely on timing restrictions. Hence, reduced-form innovations are related to the structural shocks following $\boldsymbol{u}_t = \boldsymbol{S}\boldsymbol{\varepsilon}_t$, where $\boldsymbol{\varepsilon}_t$ denotes the vector of structural shocks and $\boldsymbol{S}$ corresponds to the lower Cholesky factor of the covariance matrix, i.e., $\tilde{\boldsymbol{\Sigma}} = \boldsymbol{S}\boldsymbol{S}'$ holds. Here $\tilde{\boldsymbol{\Sigma}}$ denotes the time-varying mean of $\boldsymbol{\Sigma}_t$. Due to the ordering of the variables, the imposed timing restrictions are fulfilled and let $\boldsymbol{s}_5$ be the fifth column of $\boldsymbol{S}$, which yields the impact vector of an structural innovation to the EBP.

For the Bayesian estimation of the VAR, I rely on the Minnesota prior setup on the linear coefficients (Doan, Litterman and Sims, 1984; Litterman, 1986; Sims and Zha, 1998). In this framework, I assume a Gaussian prior distribution on the coefficients. The Minnesota prior specifies the prior belief that macroeconomic time series follow a priori a random walk. Additionally, it imposes the belief that higher-order lags are less important and thus closer linked to a value of zero. Practically, this means that the variance is smaller for coefficients on further lags. Finally, I impose that little is known about the deterministic term, so that the variance on these terms may be large. The stochastic volatility specification allows for changing variances over time, controlling for heteroskedasticity. This is of particular importance in the sample I consider here which is driven by unusually large shocks. Clark (2011) shows that adding stochastic volatility to a vector autoregression significantly improves its fit and forecasting performance. In Appendix D, I write down the prior specification in more detail.

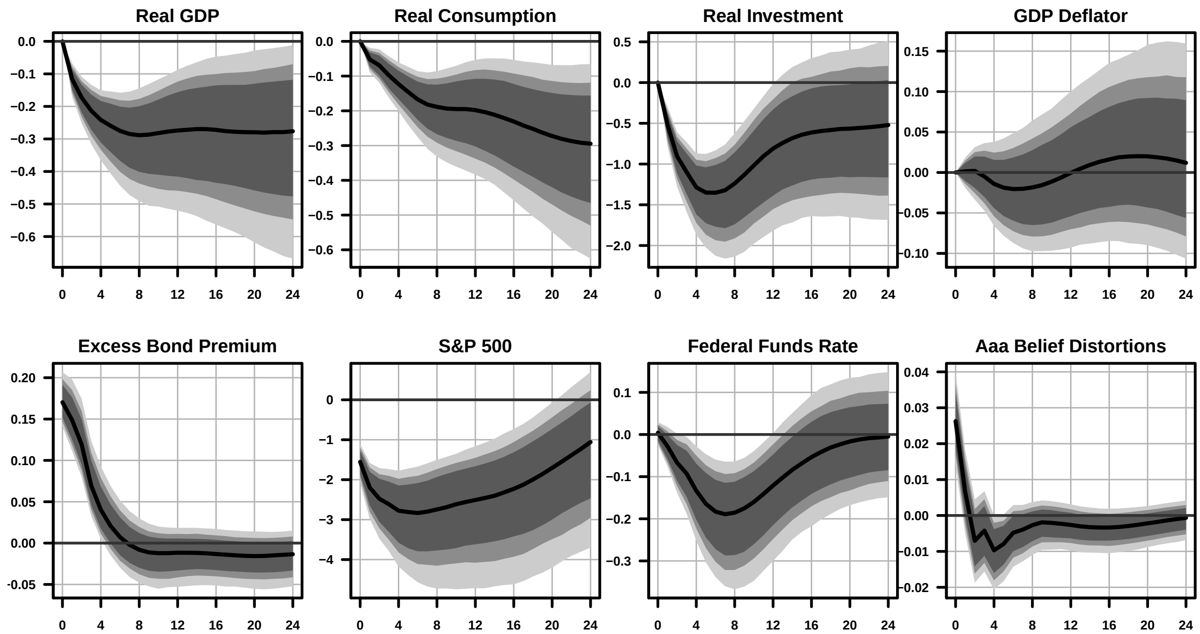
Impulse Response Analysis

I now present the results from the baseline VAR model. The VAR is estimated in (log-)levels, a constant is added as a deterministic term, and the lag order is set to $p = 4$. All models considered are based on 25,000 draws from the posterior distribution, where I discard the first 15,000 draws as burn-ins. Additionaly, I discard ex-post all non-stationary draws to ensure the stationarity of the VAR. In Appendix G, I report convergence diagnostics and the share of retained draws in each of the considered models.

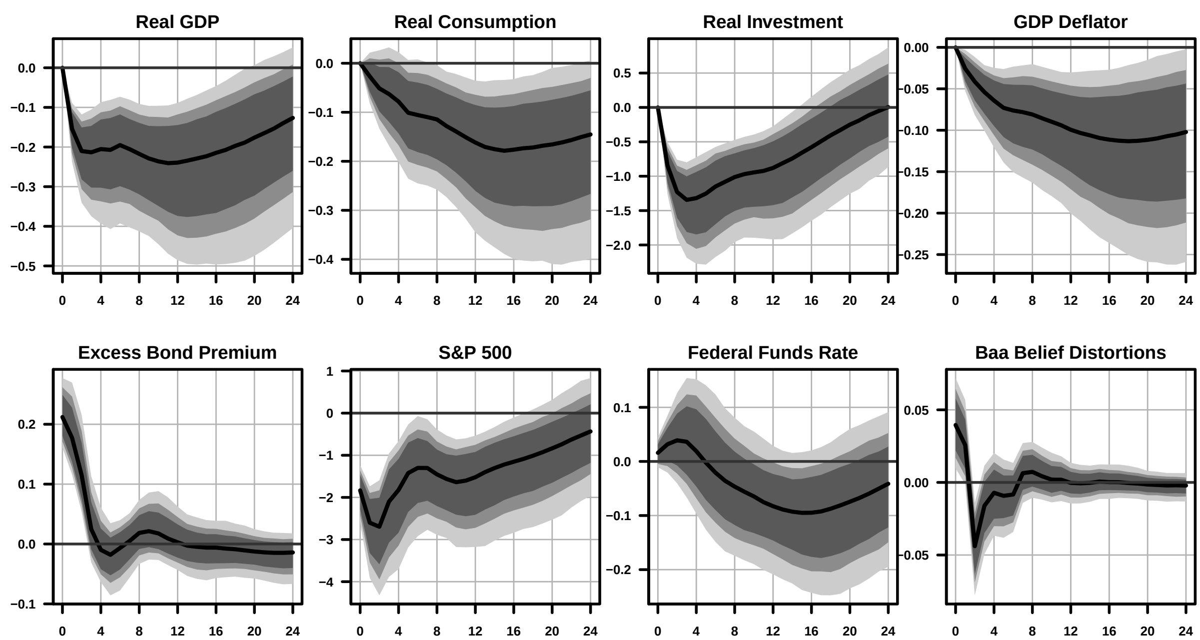
Figure 4 presents the impulse responses to a financial shock, normalized to a one standard deviation shock in the EBP. Real GDP, real consumption, real investment, the GDP deflator, and the stock market index are in logs and the responses can thus be interpreted as elasticities. The responses of the EBP, the federal funds rate, and belief distortions are in percentage points. The solid black lines are the posterior median and the gray

Figure 4: Impulse Response Functions to a Financial Shock (Belief Distortions).

(a) Model with Aaa Belief Distortion.



(b) Model with Baa Belief Distortion.



Notes: Impulse response functions of the baseline VAR. Black line denotes median response while gray shaded areas denote the 68/80/90 percent confidence intervals. The responses of real GDP, real consumption, real investment, GDP deflator, and the stock market index are scaled in percent, while the excess bond premium, the federal funds rate, and belief distortions are scaled in percentage points.

shaded areas are 68, 80, and 90 percent confidence bands. Impulse responses are computed for a horizon of 24 quarters.

The financial shock causes an increase in the EBP, standardized to one standard deviation. An unanticipated increase in the EBP elicits a surge of about 20 basis points (bp); 17bp in the model with Aaa belief distortions and 21bp in the model with Baa belief distortions. This causes a significant and long-lasting contraction of economic activity. Across both models, the (log-)level of real GDP bottoms out at about $0.2 - 0.3$ percent below trend. The consequences for real investments are far more detrimental, amounting to a -1.3% change at maximum. Real consumption, however, moves similarly to real activity. Prices react more muted and show distinct differences across both models, although point estimates show deflationary tendencies.[11] The rather muted response of prices aligns with the findings of Del Negro et al. (2020) who point to a flattening of the Phillips curve. Both models offer a pronounced decline in stock prices by about 1.5% on impact. The effects bottom out at -2.8% before gradually returning back to the steady state. Actions taken by the central bank show a more delayed but accomodating reaction via the federal funds rate.[12]

The main innovation here is to append the belief distortion series to the specification. The pattern of the impulse response across both models is similar. Belief distortions are initially positive for a short period (about 1-2 quarters) before turning negative about one to two years after the shock. Note that belief distortions are defined as the ML expectations minus the survey expectations. Hence, a positive belief distortion means that survey expectations are below ML expectations or in other words: survey participants evaluate the future more optimistically than the machine efficient benchmark (since a lower credit spread indicates easier financing conditions). These results indicate that agents stay initially optimistic in their evaluation of the future despite the adverse financial shock. Only after a couple of periods, their evaluation turns pessimistic. This highlights the important mechanism of beliefs conditional on a shock and supports the view of belief-driven economic fluctuations.

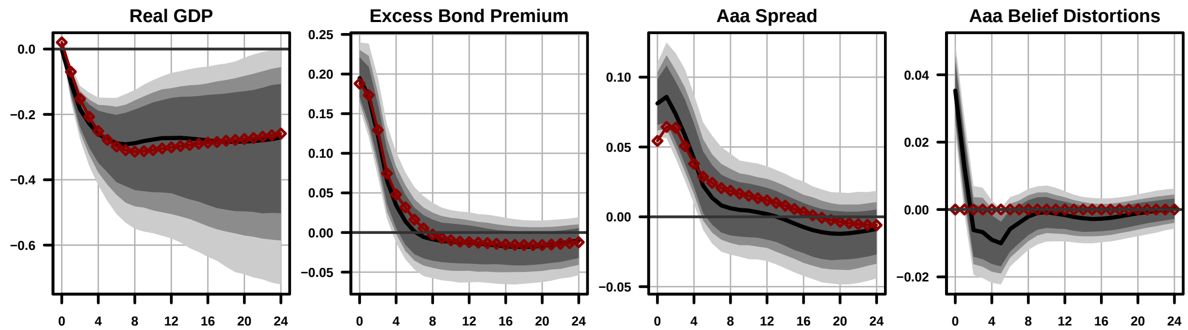
The macroeconomic effects are in line with the literature on financial shocks. Using almost the same specification, Gilchrist and Zakrajšek (2012) find that an increase in the EBP of about 20bp leads to a reduction in the level of real GDP of about 0.5 percentage points. Barnichon, Matthes and Ziegenbein (2022) find similar effects using the EBP and examining the nonlinear nature of financial disruptions. Similarly,

[11] Differences with respect to prices is due to the different sample lengths. In a robustness check, I show that prices react similarly in a model featuring Aaa belief distortions but with the same (shorter) sample length as in the model with Baa belief distortions. See also Figure I1.

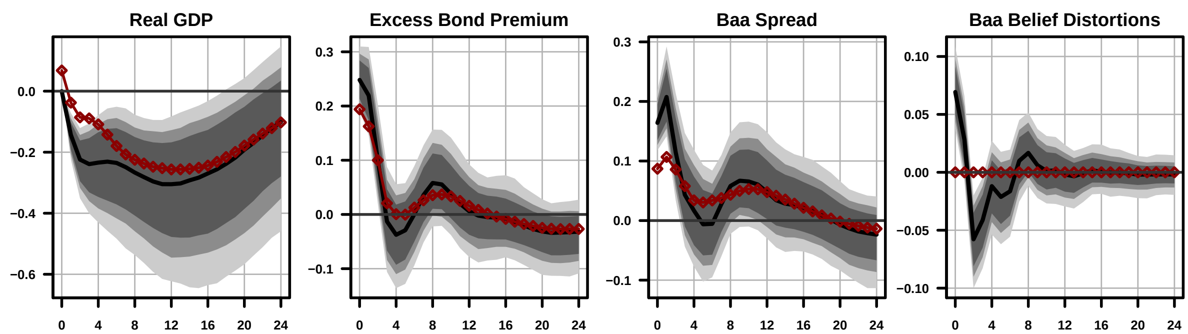
[12] Again, differences arise between the models. This are entirely due to the different sample lengths, see also the robustness check in Figure I1.

Figure 5: Counterfactual Impulse Response Functions to a Financial Shock (Belief Distortions).

(a) Model with Aaa Belief Distortion.



(b) Model with Baa Belief Distortion.



Notes: Impulse response functions of the baseline VAR. Black line denotes median response while gray shaded areas denote the 68/80/90 percent confidence intervals. The red line with diamonds denotes the counterfactual response. The responses of real GDP, real consumption, real investment, GDP deflator, and the stock market index are scaled in percent, while the excess bond premium, the federal funds rate, and belief distortions are scaled in percentage points.

Furlanetto, Ravazzolo and Sarferaz (2019) find a reaction of GDP of about 2 – 4% after a one-standard deviation financial shock identified with sign-restrictions. Furthermore, the work by López-Salido, Stein and Zakrajšek (2017) finds a 2% change of real GDP per capita when a unit change in the Baa spread happens. Observed standard deviations of the Aaa or Baa spread correspond to values in the range of 0.5 – 0.7%, implying a negative GDP response of 0.6 – 0.8% after an autonomous change in beliefs. Caggiano et al. (2021) disentangle financial uncertainty shocks from financial shocks and their shocks are also similar in magnitude. This also highlights the important role of financial uncertainty in the shock transmission.

Counterfactual Experiment

Since I have shown that belief distortions react dynamically to the financial shock, it is interesting to ask the following hypothetical question: What happens to the transmission of the financial shock if agents are not distorted in their belief formation mechanism? To answer this, I construct a counterfactual in which I shut off the response of belief distortions to the financial shock. To do so, I use the structural scenario analysis framework of Antolin-Diaz, Petrella and Rubio-Ramirez (2021). The impulse responses of belief distortions are constrained to be zero at all horizons. In order to satisfy these constraints, I allow for additional structural shocks. Specifically, because I do not identify any other shock I use the combination of all unidentified shocks to offset the response of the belief distortion series. The details on the exact implementation of the counterfactuals are provided in Appendix E. Additionally, I append the respective credit spread to the specification (after the EBP). This allows me to gauge the effects of optimism and pessimism measured through the reaction of belief distortions directly on the respective credit spread. Otherwise, the specification is unchanged. For the sake of brevity, I only report four out of nine impulse response functions along with its counterfactual response.

Results are presented in Figure 5. In the appendix, I check the plausibility of the counterfactuals in Figure E1 and report the full set of impulse responses in Figure E2. The exercise reveals some interesting facts. First, the response of the excess bond premium is not affected in the model with Aaa belief distortions and only slightly in the model with Baa belief distortions. Hence, the financial shock itself is not strongly affected by sentiments. Second, the response of GDP points to the fact that the transmission of the financial shock to the real economy is not affected by sentiments. Once a financial shocks is realized, it transmits to the real economy as suggested by the financial accelerator framework. In the model with the Baa spread, there is even a sizable effect for real GDP for a couple of quarters. Third, in both models the credit spread itself is reduced on impact in the counterfactual scenario. This effect is again stronger in the model with Baa belief distortions but also clearly visible in the model with Aaa belief distortions. It is also possible that there is an upward bias due to a higher default premium present in the model with Baa belief distortions/spreads. Still, this suggests that sentiments are at least partly driving the uptick in the respective credit spread. This corroborates the theoretical findings of Maxted (2023) who highlights the role of behavioral frictions alongside financial frictions. Exactly these wrongly formed expectations cause credit spreads to be too low in the run-up, and too high in the event of a financial crisis. Hence, belief distortions create an environment in which a crisis becomes more likely in the case of an adverse financial shock.

Quantitative Importance

As a next step, I analyze the quantitative importance of the financial shock. This analysis reveals how much of the variation in the variables in the VAR system is explained by the financial shock. Figure 6 presents the results. Initially, a financial shock explains a sizable share of the movements in the EBP before it stabilizes at around 40%. The financial shock also explains a large share (about 30%) of the reaction of the stock market. With respect to the real aggregates, the financial shock explains about 20% of real GDP, real consumption, and real investment in the long-run. Again, differences arise with respect to prices. In the model with Aaa belief distortions almost no variation in prices is explained, while the explained share in the model featuring Baa belief distortions sums up to 20%. The financial shock explains about 10% of the variation in the federal funds rate and the respective belief distortion series. Regarding belief distortions, this is not suprising since the series are characterized by ex-ante expectational *errors*, which should move only due to reaction to news. News arise due to a variety of reasons and from different sources of which an unanticipated tightening of the EBP is only one (useful) example.

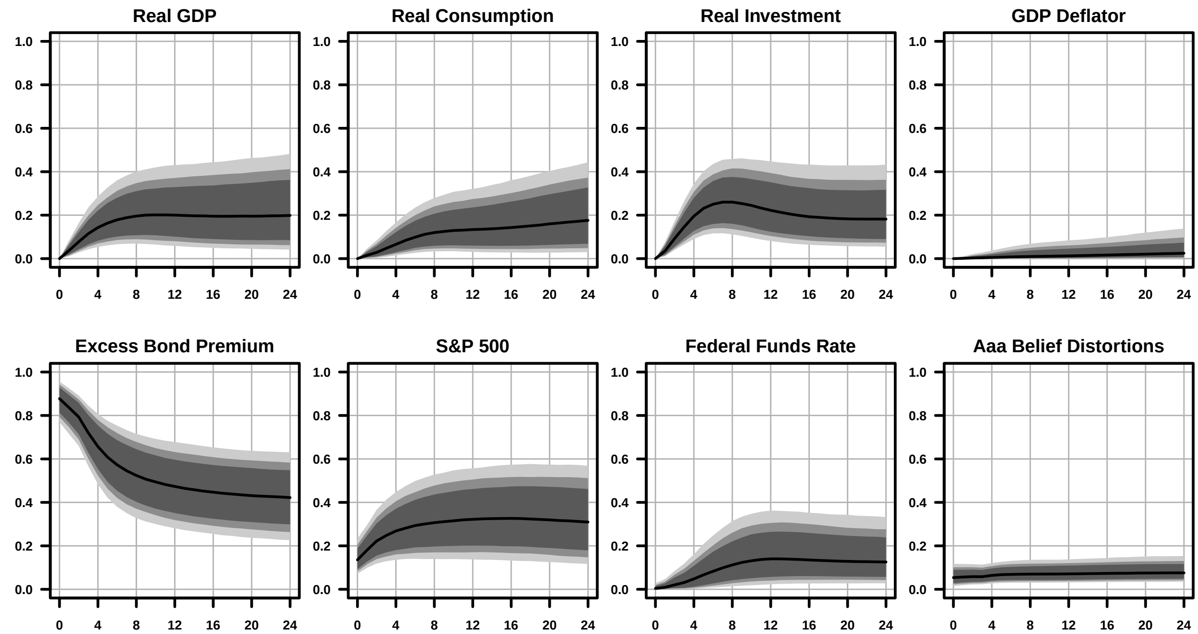
To compare the findings to the literature, usually 10-40% of the movement in real activity is explained by financial shocks. For instance, Gilchrist and Zakrajšek (2012) explain about 11% of the variation in output after a shock to the EBP. Furlanetto, Ravazzolo and Sarferaz (2019), however, explain up to 30% of the variation in GDP after their sign-identified financial shock. Brunnermeier et al. (2021) explain about 15% of the variation in industrial production to a financial shock. Lastly, Caldara et al. (2016) report 20-40% of explained variation of industrial production to a financial shock. Together with the impulse response analysis, this provides further evidence of a short-lived but statistically significant impact on belief distortions.

Reaction of Survey Forecast Errors

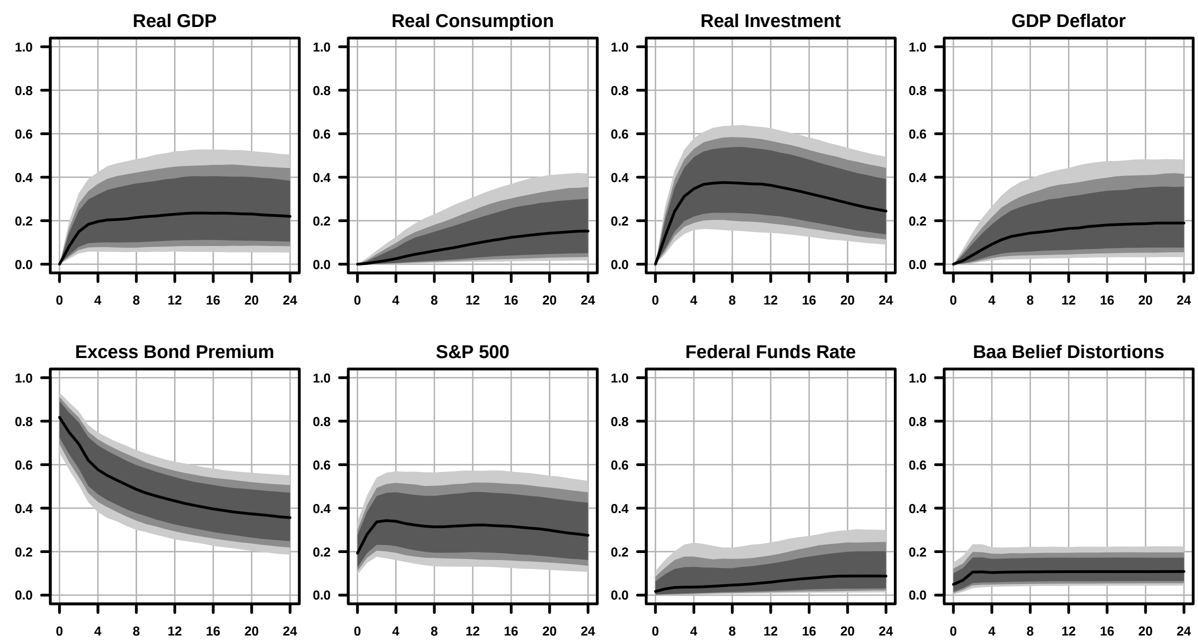
While I have already explored the dynamic responses of belief distortions to a financial shock, the reaction of survey forecast errors allows me to inspect the belief formation mechanism in more detail. The interpretation of the two exercises is different. Belief distortions show ex-ante expectational errors and their impulse responses show whether agents evaluate the future optimistically or pessimistically. Survey forecast errors, however, directly display the reaction to news (i.e., a structural financial shock) and indicate whether agents under- or overreact in their expectation formation to news (Kučinskas and Peters, 2022; Angeletos, Huo and Sastry, 2021). In contrast to the static regression approach in section 4, a dynamic analysis traces out the

Figure 6: Forecast Error Variance Decomposition.

(a) Shock to Aaa Belief Distortions.



(b) Shock to Baa Belief Distortions.



Notes: Forecast error variance decomposition of the variables in the system to a financial shock. Bold lines denote median response while gray shaded areas / dashed lines denote the 68/80/90 percent confidence intervals. The y-axis indicates the share of explained variance at a given impulse response horizon conditional on a financial shock.

reaction to news over time and offers a clearer picture. Specifically, agents underreact to the financial shock when forming expectations about credit spreads if agents perceive the impact of the shock to be smaller than it actually is. I use the same shock and identification as in the baseline model and append survey forecast errors instead of belief distortions to the specification.

The results of this exercise are presented in Figure 7. While the responses of the EBP, real aggregates, prices, stock market index, and short-term interest rates are basically the same as before (see Figure 4). Across the two models, the aforementioned and discussed differences in the response of prices and the federal funds rate arise.[13] Interestingly, the responses of Aaa/Baa survey forecast errors resemble the responses of the respective belief distortions. Across the two models, the pattern for both credit spread survey forecast errors is the same but the interpretation changes. The impulse responses of survey forecast errors indicate that initially agents underreact to the structural financial shock before the coefficients turn negative corresponding to overreaction. This pattern is called *dynamic overshooting* and has been found for other macroeconomic time series (Angeletos, Huo and Sastry, 2021).

Since the responses of belief distortions and survey forecast errors are quite similar, let me elaborate on the insights of this exercise.[14] The unconditional correlation of survey forecast errors and belief distortions is already substantial with $\rho^{uc}_{Aaa} = 0.56$ (s.e. 0.07) and $\rho^{uc}_{Baa} = 0.59$ (s.e. 0.09) and conditional on the shock the correlation is extremely high with $\rho^{c}_{Aaa} = 0.98$ (s.e. 0.04) and $\rho^{c}_{Baa} = 0.83$ (s.e. 0.12). Additionally, by examining the responses of ML forecast errors[15] conditional on the financial shock (see Figure I4), one can see that even the machine is underpredicting the surge in credit spreads. This points to the fact that credit spreads rise disproportionately to a financial shock. However, the interpretation is different. The evidence in belief distortions points to the fact that agents evaluate the future too optimistically before turning pessimistic. This aligns well with the evidence on reaction to news: initially agents underreact to the financial shock (news) but then overreact in their expectation formation behavior. Generally, this is reassuring that the ML expectations truly recover FIRE forecasts. Furthermore, this is in accordance with the evidence in Krishnamurthy and Muir (2017) that credit spreads are generally too low in the buildup of a financial crisis. I argue that agents neither realize the extent nor the extremeness of the financial shock, and thus underreact

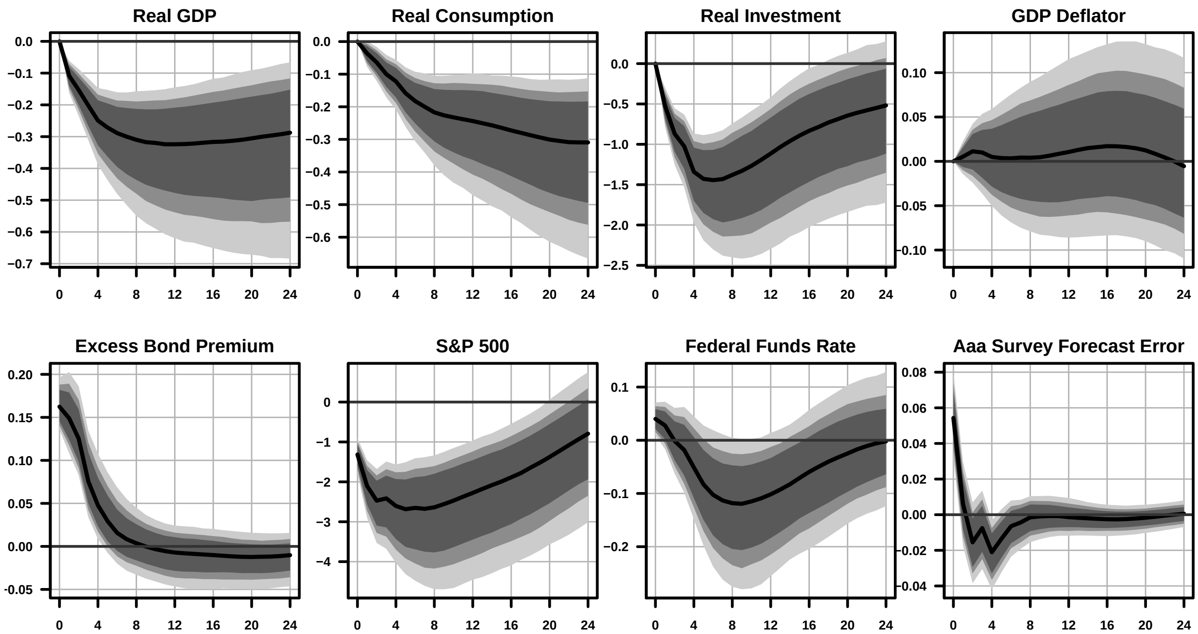
[13] Again, I have done the according robustness check and re-estimated the model featuring Aaa survey forecast errors with the shorter sample. Then, responses are qualitative and quantitative similar across both models, see also Figure I2.

[14] Recall the definitions of belief distortions and survey forecast errors. Belief distortions are defined as $Z_t = \mathbb{E}_t y_{t+h} - \mathbb{F}_t y_{t+h}$ and measure ex-ante expectational errors. Survey forecast errors are defined as $FE^{\text{survey}}_t = y_{t+h} - \mathbb{F}_t y_{t+h}$ and their impulse response functions measure reactions to news.

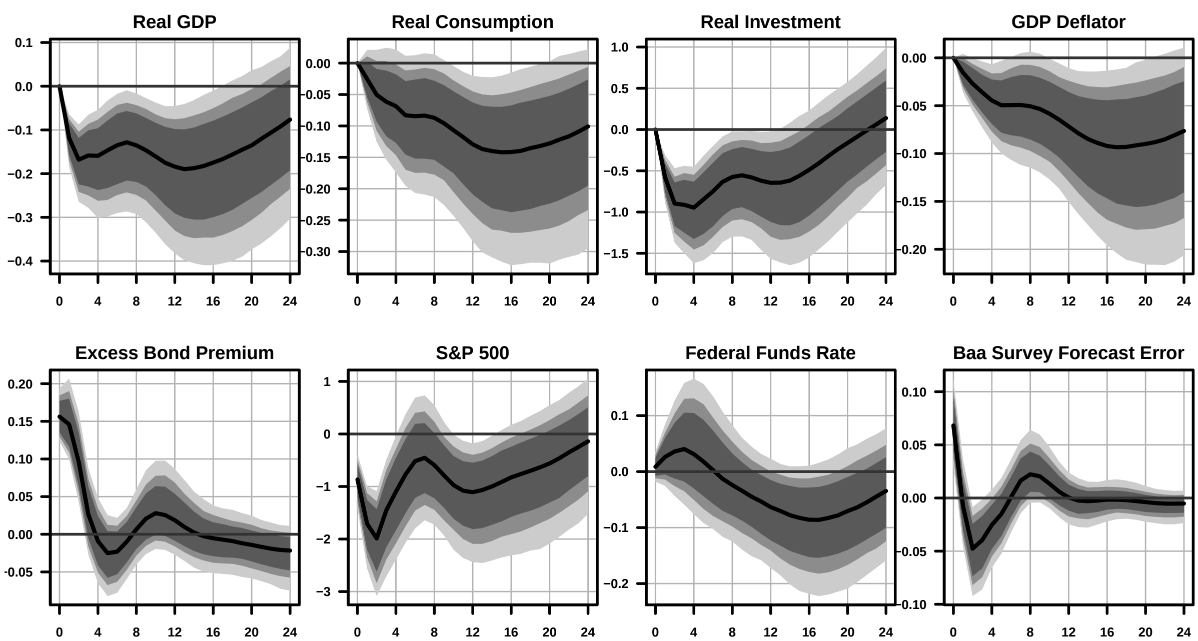
[15] ML forecast errors are defined as $FE^{\text{ML}}_t = y_{t+h} - \mathbb{E}_t y_{t+h}$ and their impulse response functions measure whether the machine efficient benchmark is distorted in their predictions conditional on the financial shock.

Figure 7: Impulse Response Functions to a Financial Shock (Survey Forecast Errors).

(a) Model with Aaa Belief Distortions.



(b) Model with Baa Belief Distortions.



Notes: Impulse response functions of the VAR with survey forecast errors. Black line denotes median response while gray shaded areas denote the 68/80/90 percent confidence intervals. The responses of real GDP, real consumption, real investment, GDP deflator, and the stock market index are scaled in percent, while the excess bond premium, the federal funds rate, and survey forecast errors are scaled in percentage points.

to the news and stay optimistic. In the GFC, this can be associated with the period in late 2007 / beginning of 2008 in which financial turmoil delineated. Only after the bankruptcy of Lehman Brothers in September 2008, credit spreads skyrocketed. This event is easily identified in hindsight but led to massive havoc on the markets. However, financial shocks have already hit the economy since late 2007 and the evidence shows the slow realization of crash risk until the point where agents start to overreact and turn pessimistic.

Sensitivity Analysis

I perform a comprehensive set of sensitivity checks. In particular, I discuss sensitivity with regards to the use of an alternative specification, higher-order belief distortions, and model missspecification issues. All corresponding figures can be found in Appendix I.

Alternative Specification. For the construction of belief distortions, this paper also explores an alternative specification, directly controlling for the median survey forecast of the respective credit spread. This has been discussed in section 4 and the results presented in Table 2. The sample span for the out-of-sample evaluation sample of the ML expectations significantly shorten, starting in 1995Q2 for the Aaa credit spread and 2010Q1 for the Baa credit spread, respectively. Since the Baa credit spread then does not cover the GFC any more and the evaluation sample is only 41 periods, I abstain from re-doing the dynamic analysis. However, I re-do the analysis for the Aaa belief distortions with this alternative specification. The results are presented in Figure I3. The qualitative and quantitative pattern in terms of IRFs and the FEVD is the same. Results are similar to the another robustness check with the the short sample (see Figure I1), which starts in 1999Q1. Belief distortions intitially are positive, translating into elevated sentiment before turning quickly pessimistic.

Higher-Order Belief Distortions. In the baseline VAR I use one-step ahead belief distortions. However, the dataset allows me to also examine higher-order belief distortions. Therefore, I also construct belief distortions for the two-, three-, and four-step ahead horizon. Additionally, I also use the mean over all four horizons as an additional check to take model uncertainty into account. Then, I re-estimate the baseline VAR and the VAR with survey forecast errors by exchanging the one-step ahead belief distortion series with higher-order belief distortions and mean belief distortions. Results of these exercises are available in Figure I5 and Figure I6 and show no stark differences to the baseline model.

Model Missspecification. In the baseline specification, we assume a specific structure of the VAR. Arguably, the most influential choices are prior distributions in a Bayesian setting and the stochastic volatility specification to control for heteroskedasticity. Although the Minnesota prior setup is well established, I re-estimate the model with a variant of the global-local shrinkage priors. Specifically, I use the horseshoe shrinkage prior already used earlier for the estimation of the machine efficient benchmark. The advantage of this prior is that the user does not have not to specify any hyperparameters but as any shrinkage prior, it trades off variance against bias. Results are presented in Figure I7 for the baseline model. The qualtitative pattern of the belief distortion series persist as in the baseline scenario although shrinkage is now entirely data-driven. The specification without stochastic volatility is presented in Figure I8 and again, results do not change much.

6. Concluding Remarks

In this paper, I investigate ex-ante expectational errors in belief about financial crash risk. For that purpose, I measure belief distortions in risk premia transmitted in credit spreads as the difference between a machine efficient benchmark resembling FIRE and survey expectations. These belief distortions arise due to undue or wrong expectations and are driven by underlying fundamental shocks. Since agents only observe a noisy and distorted signal about risk premia, these belief distortions are possibly driven by both under- and overreaction to news and characterize how optimistic or pessimistic agents evaluate the future.

Evidence suggests that economic agents overreact individually in their expectations about risk premia to incoming news but underreact in the aggregate due to noisy information. A stylized model of public and private signals with diagnostic expectations explains these findings. For the construction of the machine efficient benchmark, or ML expectations, a comprehensive set of models is estimated with linear and non-linear machine learning technique and an information set representing the potential information set of a professional forecaster in the survey. In each period, the best-performing model in terms of RMSEs is chosen in a pseudo out-of-sample forecasting exercise. This model is then used to construct a true out-of-sample forecast. Results indicate that the ML expectations outperform survey forecasts in terms of predictive accuracy. The predictive gains are sizable and in the range of a $15-20\%$ improvement. The constructed belief distortion series aligns well with key historical episodes and fluctuates around zero, showing no substantial bias in the survey forecasts.

In a next step, I evaluate the dynamic responses of belief distortions to a financial shock. The financial shock is identified with unanticipated innovations to the EBP identified with timinig assumptions. Real economic aggregates and prices do not react contemporaneously to innovations in the excess bond premium, while financial variables and belief distortions are allowed to react on impact. A positive one standard deviation financial shock elicits a jump in the EBP and causes an aggregate downturn. Belief distortions indicate that survey participants stay optimistic for a couple of periods before turning pessimistic in their evaluation of the future. In a counterfactual exercise the response of belief distortions to the financial shock is set to zero, essentially eliminating a channel via sentiments. While the responses of the real variables do not change much, the hike in the credit spread is attenuated. Re-doing the analysis with survey forecast errors instead of belief distortions, allows me to inspect the expectation formation mechanism further. The exercise reveals the pattern of dynamic overshooting, which means that agents initially underreact to news before overreacting to the financial shock. These findings are robust to a number of specification choices. Taken at face value, the evidence allows the interpretation that agents underestimate financial shocks in the beginning until financial turmoil is inevitable. The evidence presented here offers a forward-looking or *ex-ante* perspective on financial risks.

Declaration of Interest

The author declares to have no conflict of interest.

References

Akinci O. and Queralto A. (2022) "Credit Spreads, Financial Crises, and Macroprudential Policy". *American Economic Journal: Macroeconomics* Vol. 14(2), pp. 469–507.

Angeletos G.-M., Huo Z. and Sastry K A. (2021) "Imperfect Macroeconomic Expectations: Evidence and Theory". *NBER Macroeconomics Annual* Vol. 35(1), pp. 1–86.

Antolin-Diaz J., Petrella I. and Rubio-Ramirez J F. (2021) "Structural Scenario Analysis with SVARs". *Journal of Monetary Economics* Vol. 117, pp. 798–815.

Anufriev M. and Hommes C. (2012) "Evolutionary Selection of Individual Expectations and Aggregate Outcomes in Asset Pricing Experiments". *American Economic Journal: Microeconomics* Vol. 4(4), pp. 35–64.

Assenza T., Heemeijer P., Hommes C. and Massaro D. (2019) "Managing Self-Organization of Expectations Through Monetary Policy: A Macro Experiment". *Journal of Monetary Economics* Vol. 117, pp. 170–186.

Barberis N., Greenwood R., Jin L. and Shleifer A. (2015) "X-CAPM: An Extrapolative Capital Asset Pricing Model". *Journal of Financial Economics* Vol. 115(1), pp. 1–24.

Barberis N., Shleifer A. and Vishny R. (1998) "A Model of Investor Sentiment". *Journal of Financial Economics* Vol. 49(3), pp. 307–343.

Barnichon R., Matthes C. and Ziegenbein A. (2022) "Are the effects of financial market disruptions big or small?" *The Review of Economics and Statistics* Vol. 104(3), pp. 557–570.

Baron M. and Xiong W. (2017) "Credit Expansion and Neglected Crash Risk". *Quarterly Journal of Economics* Vol. 132(2), pp. 713–764.

Beaudry P. and Willems T. (2022) "On the Macroeconomic Consequences of Over-Optimism". *American Economic Journal: Macroeconomics* Vol. 14(1), pp. 38–59.

Benhima K. and Poilly C. (2021) "Does Demand Noise Matter? Identification and Implications". *Journal of Monetary Economics* Vol. 117, pp. 278–295.

Bernanke B S., Gertler M. and Gilchrist S. (1999) “The Financial Accelerator in a Quantitative Business Cycle Framework”. *Handbook of Macroeconomics* Vol. 1, pp. 1341–1393.

Bianchi F., Ilut C L. and Saijo H. (forthcoming) “Diagnostic Business Cycles”. *The Review of Economic Studies*.

Bianchi F., Ludvigson S C. and Ma S. (2022) “Belief Distortions and Macroeconomic Fluctuations”. *American Economic Review* Vol. 112(7), pp. 2269–2315.

Boeck M. and Zörner T O. (2019) *The Impact of Credit Market Sentiment Shocks – A TVAR Approach*. Working Paper 288. Vienna University of Economics and Business.

Bordalo P., Gennaioli N., Ma Y. and Shleifer A. (2020) “Overreaction in Macroeconomic Expectations”. *American Economic Review* Vol. 110(9), pp. 2748–82.

Bordalo P., Gennaioli N., Porta R L. and Shleifer A. (2019) “Diagnostic Expectations and Stock Returns”. *The Journal of Finance* Vol. 74(6), pp. 2839–2874.

Bordalo P., Gennaioli N. and Shleifer A. (2018) “Diagnostic Expectations and Credit Cycles”. *Journal of Finance* Vol. 73(1), pp. 199–227.

Brav A. and Heaton J B. (2002) “Competing Theories of Financial Anomalies”. *Review of Financial Studies* Vol. 15(2), pp. 575–606.

Brunnermeier M., Palia D., Sastry K A. and Sims C A. (2021) “Feedbacks: Financial Markets and Economic Activity”. *American Economic Review* Vol. 111(6), pp. 1845–79.

Caggiano G., Castelnuovo E., Delrio S. and Kima R. (2021) “Financial Uncertainty and Real Activity: The Good, the Bad, and the Ugly”. *European Economic Review* Vol. 136(103750).

Caldara D., Fuentes-Albero C., Gilchrist S. and Zakrajšek E. (2016) “The Macroeconomic Impact of Financial and Uncertainty Shocks”. *European Economic Review* Vol. 88, pp. 185–207.

Capistrán C. and Timmermann A. (2009) “Disagreement and Biases in Inflation Expectations”. *Journal of Money, Credit and Banking* Vol. 41(2-3), pp. 365–396.

Carvalho C M., Polson N G. and Scott J G. (2010) “The Horseshoe Estimator for Sparse Signals”. *Biometrika* Vol. 97(2), pp. 465–480.

Chipman H A., George E I. and McCulloch R E. (2010) "BART: Bayesian Additive Regression Trees". *Annals of Applied Statistics* Vol. 4(1), pp. 266 –298.

Christiano L J., Motto R. and Rostagno M. (2014) "Risk Shocks". *American Economic Review* Vol. 104(1), pp. 27–65.

Clark T E. (2011) "Real-Time Density Forecasts from Bayesian Vector Autoregressions with Stochastic Volatility". *Journal of Business & Economic Statistics* Vol. 29(3), pp. 327–341.

Coibion O. and Gorodnichenko Y. (2015) "Information Rigidity and the Expectations Formation Process: A Simple Framework and New Facts". *American Economic Review* Vol. 105(8), pp. 2644–78.

De Long J B., Shleifer A., Summers L H. and Waldmann R J. (1990) "Positive Feedback Investment Strategies and Destabilizing Rational Speculation". *Journal of Finance* Vol. 45(2), pp. 379–395.

Del Negro M., Lenza M., Primiceri G E. and Tambalotti A. (2020) "What's Up with the Phillips Curve?" *Brookings Paper of Economic Activity*, pp. 301 –373.

Dick-Nielsen J., Feldhütter P. and Lando D. (2012) "Corporate Bond Liquidity Before and After the Onset of the Subprime Crisis". *Journal of Financial Economics* Vol. 103(3), pp. 471–492.

Doan T., Litterman R. and Sims C. (1984) "Forecasting and Conditional Projection using Realistic Prior Distributions". *Econometric Reviews* Vol. 3(1), pp. 1–100.

Elton E J., Gruber M J., Agrawal D. and Mann C. (2001) "Explaining the Rate Spread on Corporate Bonds". *Journal of Finance* Vol. 56(1), pp. 247–277.

Enders Z., Kleemann M. and Müller G J. (2021) "Growth Expectations, Undue Optimism, and Short-Run Fluctuations". *Review of Economics and Statistics* Vol. 103(5), pp. 905–921.

Furlanetto F., Ravazzolo F. and Sarferaz S. (2019) "Identification of Financial Factors in Economic Fluctuations". *Economic Journal* Vol. 129(617), pp. 311–337.

Fuster A., Laibson D. and Mendel B. (2010) "Natural Expectations and Macroeconomic Fluctuations". *Journal of Economic Perspectives* Vol. 24(4), pp. 67–84.

Gennaioli N. and Shleifer A. (2018) *A Crisis of Beliefs: Investor Psychology and Financial Fragility*. Princeton University Press.

Gertler M., Kiyotaki N. and Prestipino A. (2019) "A Macroeconomic Model with Financial Panics". *Review of Economic Studies* Vol. 87(1), pp. 240–288.

Gilchrist S. and Zakrajšek E. (2012) "Credit Spreads and Business Cycle Fluctuations". *American Economic Review* Vol. 102(4), pp. 1692–1720.

Goulet Coulombe P., Leroux M., Stevanovic D. and Surprenant S. (2022) "How is Machine Learning Useful for Macroeconomic Forecasting?" *Journal of Applied Econometrics* Vol. 37(5), pp. 920–964.

Greenwood R. and Hanson S G. (2013) "Issuer Quality and Corporate Bond Returns". *Review of Financial Studies* Vol. 26(6), pp. 1483–1525.

Greenwood R., Hanson S G., Shleifer A. and Sørensen J A. (2022) "Predictable Financial Crisis". *Journal of Finance* Vol. 77(2), pp. 863–921.

Gu S., Kelly B. and Xiu D. (2020) "Empirical Asset Pricing via Machine Learning". *The Review of Financial Studies* Vol. 33, pp. 2223–2273.

Jermann U. and Quadrini V. (2012) "Macroeconomic Effects of Financial Shocks". *American Economic Review* Vol. 102(1), pp. 238–71.

Jurado K., Ludvigson S C. and Ng S. (2015) "Measuring Uncertainty". *American Economic Review* Vol. 105(3), pp. 1177–1216.

Kocher M G., Lucks K E. and Schindler D. (2019) "Unleashing Animal Spirits: Self-Control and Overpricing in Experimental Asset Markets". *Review of Financial Studies* Vol. 32(6), pp. 2149–2178.

Krishnamurthy A. and Li W. (2020) *Dissecting Mechanisms of Financial crises: Intermediation and Sentiment*. NBER Working Paper No. 27088.

Krishnamurthy A. and Muir T. (2017) *How Credit Cycles across a Financial Crisis*. NBER Working Paper 23850.

Kučinskas S. and Peters F S. (2022) "Measuring Under- and Overreaction in Expectation Formation". *Review of Economics and Statistics*, pp. 1–45.

L'Huillier J.-P., Singh S R. and Yoo D. (2021) *Incorporating Diagnostic Expectations into the New Keynesian Framework*. Available at SSRN 3910318.

Litterman R B. (1986) “Forecasting with Bayesian vector autoregressions—five years of experience”. *Journal of Business & Economic Statistics* Vol. 4(1), pp. 25–38.

López-Salido D., Stein J C. and Zakrajšek E. (2017) “Credit-Market Sentiment and the Business Cycle*”. *Quarterly Journal of Economics* Vol. 132(3), pp. 1373–1426.

Malmendier U. and Nagel S. (2011) “Depression Babies: Do Macroeconomic Experiences Affect Risk Taking?” *Quarterly Journal of Economics* Vol. 126(1), pp. 373–416.

Malmendier U. and Nagel S. (2016) “Learning From Inflation Experiences”. *Quarterly Journal of Economics* Vol. 131(1), pp. 53–87.

Maxted P. (2023) “A Macro-Finance Model with Sentiment”. *The Review of Economic Studies* Vol. 90 (5), pp. 2674–2712.

Newey W K. and West K D. (1994) “Automatic Lag Selection in Covariance Matrix Estimation”. *Review of Economic Studies* Vol. 61(4), pp. 631–653.

Reinhart C M. and Rogoff K S. (2009) *This Time is Different: Eight Centuries of Financial Folly*. Princeton University Press.

Sims C A. and Zha T. (1998) “Bayesian Methods for Dynamic Multivariate Models”. *International Economic Review*, pp. 949–968.

Zou H. and Hastie T. (2005) “Regularization and Variable Selection via the Elastic Net”. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* Vol. 67(2), pp. 301–320.

Appendix to ‘Belief Distortions and Risk Premia’

Maximilian Boeck*

Appendix Material

This appendix contains additional material not reported in the main text. First, it discusses the various data sources used in the paper for the survey data, macroeconomic data series, real-time macroeconomic data, and monthly financial data. Second, it contains the propositions and proofs. Third, it explains the involved estimation algorithms, i.e., dynamic factor estimation, Bayesian linear regression with the horseshoe (HS) shrinkage priors, Elastic Net (EN), and Bayesian Additive Regression Trees (BART). Furthermore, the exact specifications of the construction of the machine learning (ML) expectations and the macroeconomic model are laid out and discussed in detail. Fourth, the prior specification of the Bayesian vector autoregressive model with the Minnesota prior and stochastic volatility is discussed in detail. Fifth, it explains the exact construction of the structural scenario analysis. Sixth, diagnostics of the belief distortion series are discussed. Seventh, convergence diagnostics for all estimated models are presented. Eigth, additional results from the forecasting exercise are presented. Ninth and last, additional results from the VAR analysis are shown.

* Contact: Maximilian Boeck, Department of Economics, Università Bocconi. Via Roentgen 1, 20136 Milano, Lombardia, Italy. E-mail: maximilian.boeck@unibocconi.it.

A. Data

This results of this paper are obtained with data from various data sources. This section provides all necessary details on the data sources, transformations, and final construction. The paper utilizes four different categories of data: survey data, macroeconomic data, real-time macroeconomic data, and monthly financial data. In the following, I describe the data.

A1. Survey Data

In the paper, I use the Blue Chip expectation data from Blue Chip Financial Forecasts Wolters Kluwer (1984-2020).[2] The surveys are conducted each month by sending out surveys to forecasters in around 40 to 50 financial firms such as Bank of America, Goldman Sachs & Co., Swiss Re, Loomis, or J.P. Morgan Chase. The participants are surveyed around the 25th of each month and the results published a few days later on the 1st of the following month. The forecasters are asked to forecast the average of the level of U.S. interest rates over a particular calender quarter, i.e., the federal funds rate, prime bank rate, LIBOR rate, Treasury bill yields of the following maturities: 3-month, 6-month, 1-year, 2-year, 5-year, 10-year, 20-year, Aaa corporate bond yield, Baa corporate bond yield, and some mortgages rates. Forecasts are available for the current quarter t and for quarters $t+1$ through $t+4$. Panelists are not necessarily continuously in the survey and thus the composition of forecasters change throughout the sample. In total, I observe about 150 individual forecasters with varying sample lengths. From these cross-sections, I can construct quantities summarizing the distribution (i.e., mean, median, standard deviation, or skeweness). In this study, I utilize the forecasts of Aaa and Baa corporate bond yields and 10-year Treasury bill yields. In some instance, I use the 20-year Treasury bill rate if data is missing.

(i) Aaa credit spread: Let $\mathbb{F}_t^{(i)}\left[by_{t+h}^{aaa}\right]$ be forecster i's prediction of the level of Aaa credit spread by_{t+h}^{aaa} at horizon $h \geq 0$. Let $\mathbb{F}_t^{(i)}\left[y_{t+h}^{10}\right]$ be forecaster i's prediction of the level of the 10-year Treasury bill yield at horizon $h \geq 0$. Hence,

$$\mathbb{F}_t^{(i)}\left[cs_{t+h}^{aaa}\right] = \mathbb{F}_t^{(}i)\left[by_{t+h}^{aaa}\right] - \mathbb{F}_t^{(i)}\left[y_{t+h}^{10}\right], \tag{A.1}$$

constructs the survey expectation of the Aaa credit spread at time t for horizon h for individual i. The survey directly provides the consensus forecasts, which is the median of the observations. Hence,

$$\mathbb{F}_t\left[cs_{t+h}^{aaa}\right] = \text{median}_{i=1,\ldots,N_t}\left[\mathbb{F}_t^{(i)}\left[cs_{t+h}^{aaa}\right]\right]. \tag{A.2}$$

Data on the Aaa credit spread covers the period 1988Q1 to 2020Q1.

(ii) Baa credit spread: Let $\mathbb{F}_t^{(i)}\left[by_{t+h}^{baa}\right]$ be forecster i's prediction of the level of Baa credit spread by_{t+h}^{baa} at horizon $h \geq 0$. Let $\mathbb{F}_t^{(i)}\left[y_{t+h}^{10}\right]$ be forecaster i's prediction of the level of the 10-year Treasury bill yield at horizon $h \geq 0$. Hence,

$$\mathbb{F}_t^{(i)}\left[cs_{t+h}^{baa}\right] = \mathbb{F}_t^{(}i)\left[by_{t+h}^{baa}\right] - \mathbb{F}_t^{(i)}\left[y_{t+h}^{10}\right], \tag{A.3}$$

constructs the survey expectation of the Aaa credit spread at time t for horizon h for individual i. The survey directly provides the consensus forecasts, which is the median of the observations. Hence,

$$\mathbb{F}_t\left[cs_{t+h}^{baa}\right] = \text{median}_{i=1,\ldots,N_t}\left[\mathbb{F}_t^{(i)}\left[cs_{t+h}^{baa}\right]\right]. \tag{A.4}$$

Data on the Baa credit spread covers the period 1999Q1 to 2020Q1.

The surveys are conducted right before the publication of the newsletter. Each issue is always dated the 1st of the month and the actual survey conducted over a two-day period almost always between 24th and 28th

[2] The data were purchased and can be found under the following URL: `https://law-store.wolterskluwer.com/s/product/blue-chip-financial-forecast-print/01tG000000LuDUC`.

of the month. The major exception is the January issue when the survey is conducted a few days earlier to avoid conflict with the Christmas holiday. Therefore, I assume that the end of the last month (equivalently beginning of the current month) is when the forecast is made. For example, for the report in 2008 Sept, I assume that the forecast is made on September 1, 2008. To convert monthly forecasts to quarterly forecasts, I use the forecasts at the last month of each quarter as the quarterly forecasts. This is is to maximize the amount of information a forecaster may has in each quarter.

In total, the survey consists of about 150 distinguishable individuals with varying sample lengths due to changes in the composition of the forecaster in the survey. The data were manually checked for errors before using the data in the analysis. To detect outliers, I exclude all forecasts from the analysis which are five interquartile ranges away from the median. In case there is no variation in the interquartile range, I apply the interquartile range of the previous period to ensure the consistency of the forecasts. Obvious coding errors were corrected (e.g., in some instance the level of a yield was 119 instead of 11.9). Furthermore, one may worry that BC financial forecasts are distorted due to signaling reasons. However, forecasts for variables also entertained in the anonymous Philadelphia Fed Survey of Private Forecasters tend to be similar. The forecasts used in this study are only available in the Blue Chip professional forecasts.

A2. Macroeconomic Data

For the estimation of the macroeconomic model, we rely on macroeconomic data from the FRED database (Data, 2022) provided by the St. Louis' Federal Reserve. Data were downloaded using the `R`-package `fredr` (Boysel and Vaughan, 2019). All time series cover the time period 1970Q1 to 2021Q1. All series are seasonally adjusted, either by downloading the already adjusted series from FRED or by applying a quarterly X11 filter based on an AR(4) model to the unadjusted series. Some series in the database are observed only on a monthly basis and quarterly values are computed by obtaining quarterly averages. In Table A1 I provide a comprehensive overview of all involved variables and its exact definition and transformation before the estimations. The column *Tcode* shows the transformation I apply to a series: 1 – no transformation (levels); 2 – first difference; 4 – logarithms; 5 – first difference of logarithms; 6 – second difference in logarithms.

Table A1: Macroeconomic Data.

#	Mnemonic	Description	Tcode
Macroeconomic Data from FRED.			
1	GDPC1	Real Gross Domestic Product, 3 Decimal	1
2	PCESV	Personal Consumption Expenditure: Services	1
3	PCEDG	Personal Consumption Expenditure: Durable Goods	1
4	PCEND	Personal Consumption Expenditure: Nondurable Goods	1
5	GPDI	Gross Private Domestic Investment	1
6	INDPRO	Industrial Production Index	1
7	HOANBS	Nonfarm Business Sector: Hours of All Persons	1
8	UNRATE	Unemployment Rate	1
9	CIVPART	Labor Force Participation Rate	1
10	COMPRNFB	Nonfarm Business Sector: Real Compensation Per Hour	1
11	BUSLOANS	Commercial and Industrial Loans at All Commercial Banks	1
12	CONSUMER	Consumer (Individual) Loans at All Commercial Banks	1
13	LOANINV	Total Loans and Investments at All Commercial Banks	1
14	NASDAQCOM	NASDAQ Composite Index	1

Continued on next page

Table A1 – *Continued from previous page*

#	Mnemonic	Description	Tcode
15	FEDFUNDS	Effective Federal Funds Rate	1
16	GS1	1-year Treasury Constant Maturity Rate	1
17	GS10	10-year Treasury Constant Maturity Rate	1
18	AAA	Moody's Seasoned Aaa Corporate Bond Yield	1
19	BAA	Moody's Seasoned Baa Corporate Bond Yield	1
20	GDPDEF	Gross Domestic Product: Implicit Price Deflator	1
21	PPIACO	PPI: All Commodities	1
22	UMCSENT	University of Michigan: Index of Consumer Sentiment	1
23	BCE1Y	University of Michigan: Business Conditions Expected During the Next Year	1
24	BCE5Y	University of Michigan: Business Conditions Expected During the Next 5 Years	1
25	UE1Y	University of Michigan: Expected Change in Unemployment During the Next Year	1
26	SP500	S&P 500	1
27	CNP16OV	Population Level	1
Transformations.			
1	**Aaa spread**	$\texttt{AAA} - \texttt{GS10}$	
2	**Baa spread**	$\texttt{BAA} - \texttt{GS10}$	
3	**real GDP**	$100 \times \ln(\texttt{GDPC1})$	
4	**Real Consumption**	$100 \times \ln\left(\frac{\texttt{PCEND+PCESV}}{\texttt{GDPDEF}}\right)$	
5	**Real Investments**	$100 \times \ln\left(\frac{\texttt{GPDI+PCEDG}}{\texttt{GDPDEF}}\right)$	
6	**GDP Deflator**	$100 \times \ln(\texttt{GDPDEF})$	
7	**Excess Bond Premium**	$\texttt{EBP}$	
8	**S&P 500**	$100 \times \ln(\texttt{SP500})$	
9	**Federal Funds Rate**	$\texttt{FEDFUNDS}$	
10	**Short-term Interest Rate**	$\texttt{GS1}$	
11	**Bank Credit**	$100 \times \ln(\texttt{LOANINV})$	
12	**Business Loans**	$100 \times \ln(\texttt{BUSLOANS})$	
13	**Consumer Loans**	$100 \times \ln(\texttt{CONSUMER})$	
14	**Term Premium**	$\texttt{GS10} - \texttt{GS1}$	
15	**Hours**	$100 \times \ln\left(\frac{\texttt{HOANBS}}{2080}\right)$	
16	**Unemployment**	$\ln(\texttt{UNRATE})$	
17	**Labor Force Participation**	$100 \times \ln(\texttt{CIVPART})$	
18	**Consumer Prices**	$100 \times \ln(\texttt{CPIAUCSL})$	
19	**Producer Prices**	$100 \times \ln(\texttt{PPIACO})$	
20	**real Wage**	$100 \times \ln(\texttt{COMPRNFB})$	
21	**Consumer Sentiment**	$100 \times \ln(\texttt{UMCSENT})$	
22	**Business Expectations 1Y**	$100 \times \ln(\texttt{BCE1Y})$	
23	**Business Expectations 5Y**	$100 \times \ln(\texttt{BCE5Y})$	
24	**Unemployment Expectations**	$100 \times \ln(\texttt{UE1Y})$	
25	**NASDAQ**	$100 \times \ln(\texttt{NASDAQCOM})$	
26	**Industrial Production**	$100 \times \ln(\texttt{INDPRO})$	

A3. Real-Time Macroeconomic Data

For the construction of objective expectations, I only rely on data in real-time. This means I combine all data observed available at the time of a forecast data without including later revisions. This is done via the real-time dataset for macroeconomicsts by the Philadelphia Fed (Federal Reserve Philadelphia, 2022). In this dataset, I know the specific day that the data in each vintage are released. The series-specific documentation on the Philadelphia Fed's website provides details on the timing of the vintages for each series. For some series, exact release dates are known, and thus the vintages reflect data available at the time of the data release. For other variables, I only know that vintages contain data available in the middle of a month or quarter, but not the exact day. For another subset of variables with unknown release dates, I make the assumption that a forecaster at time t observes at most the time $t-1$ vintage of data. In addition to the macro variables with different vintages that we obtain from the Philadelphia Fed, we include a set of financial variables from the FRED database to the dataset (Data, 2022). This are available in real-time and are not revised in the future.

After combining all of the series that are known by the forecaster at each date, we convert monthly to quarterly by using the middle-of-quarter values. Again, this depends on the choice that we choose the survey forecasts at the end-of-quarter. Table A2 provides a complete list of real-time macro variables. Furthermore, the table includes the first available vintages for each variable that has multiple vintages. For most of the variables the last vintages corresponds with the end of the sample.[3] All series are seasonally adjusted, either by downloading the already adjusted series from FRED or by applying a quarterly X11 filter based on an AR(4) model to the unadjusted series. Some series in the database are observed only on a monthly basis and quarterly values are computed by obtaining quarterly averages. Concerning the data series used for computing factors, all variables are transformed to be approximately stationary. In particular, the column *Tcode* shows the transformation I apply to a series: 1 – no transformation (levels); 2 – first difference; 4 – logarithms; 5 – first difference of logarithms; 6 – second difference in logarithms.

Table A2: Real-Time Macroeconomic Data.

#	Mnemonic	Description	Tcode	
Output and Income				
1	ROUTPUT	Real Gross Domestic Product	5	1965Q4
2	NOUTPUT	Nominal Gross Domestic Product	5	1965Q4
3	IPT	Industrial Production Index: Total	5	1962M11
4	IPM	Industrial Production Index: Manufacturing	5	1962M11
5	CUT	Capacity Utilization Rate: Total	1	1983M7
6	CUM	Capacity Utilization Rate: Manufacturing	1	1979M8
7	WSD	Wages and Salary Disbursements	5	1965Q4
8	OLI	Other Labor Income	5	1965Q4
9	PROPI	Proprietor's Income	5	1965Q4
10	RENTI	Rental Income of Persons	2	1965Q4
11	DIV	Dividends	5	1965Q4
12	PINTI	Personal Interest Income	5	1965Q4
13	TRANR	Transfer Payments	5	1965Q4
14	SSCONTRIB	Personal Contribution for Social Insurance	5	1965Q4

Continued on next page

[3] EXEMPTIONS

Table A2 – *Continued from previous page*

#	Mnemonic	Description	Tcode	First Vintage
15	NPI	Nominal Personal Income	5	1965Q4
16	PTAX	Personal Tax & Nontax Payments	5	1965Q4
17	NDPI	Nominal Disposable Personal Income	5	1965Q4
18	PINTPAID	Interest Paid by Consumers	5	1965Q4
19	TRANPF	Personal Transfer Payments to Foreigners	5	1965Q4
20	NPSAV	Nominal Personal Saving	2	1965Q4
21	RATESAV	Personal Saving Rate, Constructed	2	1965Q4
Consumption and Investment				
22	RCON	Real Personal Consumption Expenditure: Total	5	1965Q4
23	RCONND	Real Personal Consumption Expenditure: Nondurable Goods	5	1965Q4
24	RCOND	Real Personal Consumption Expenditure: Durable Goods	5	1965Q4
25	RCONS	Real Personal Consumption Expenditure: Services	5	1965Q4
26	NCON	Nominal Personal Consumption Expenditure	5	1965Q4
27	RINVRESID	Real Gross Private Domestic Investment: Residential	5	1965Q4
28	RINVCHI	Real Gross Private Domestic Investment: Change in Private Inventories	2	1965Q4
Trade and Government				
29	RNX	Real Net Export of Goods and Services	2	1965Q4
30	REX	Real Exports of Goods and Services	5	1965Q4
31	RIMP	Real Import of Goods and Services	5	1965Q4
32	RG	Real Government Consumption & Gross Investment: Total	5	1965Q4
33	RGF	Real Government Consumption & Gross Investment: Federal	5	1965Q4
34	RGSL	Real Government Consumption & Gross Investment: State and Local	5	1965Q4
Money and Prices				
35	M1	M1 Money Stock	6	1965Q4
36	M2	M2 Money Stock	6	1971Q2
37	P	Price Index for GNP/GDP	6	1965Q4
38	PCON	Price Index for Personal Consumption Expenditure, Constructed	6	1965Q4
39	PIMP	Price Index for Imports of Goods and Services	6	1965Q4
Labor Market and Housing				
40	RUC	Unemployment Rate	2	1965Q4
41	EMPLOY	Nonfarm Payroll Employment	5	1964M12
42	H	Index of Aggregate Weekly Hours: Total	1	1971M9
43	HG	Index of Aggregate Weekly Hours: Goods Sector	1	1971M9
44	HS	Index of Aggregate Weekly Hours: Service Sector	1	1971M9
45	HSTARTS	Housing Starts	5	1968M2

Continued on next page

Table A2 – *Continued from previous page*

#	Mnemonic	Description	Tcode	First Vintage
Interest Rates and Spreads				
46	MPRIME	Bank Prime Loan Rate	1	-
47	FEDFUNDS	Effective Federal Funds Rate	1	-
48	TB3MS	3-Months Treasury Bill: Secondary Market Rate	1	-
49	TB6MS	6-Months Treasury Bill: Secondary Market Rate	1	-
50	GS1	1-Year Treasury Constant Maturity Rate	1	-
51	GS2	2-Year Treasury Constant Maturity Rate	1	-
52	GS3	3-Year Treasury Constant Maturity Rate	1	-
53	GS5	5-Year Treasury Constant Maturity Rate	1	-
54	GS10	10-Year Treasury Constant Maturity Rate	1	-
55	GS30	30-Year Treasury Constant Maturity Rate	1	-
56	AAA	Moody's Seasoned Aaa Corporate Bond Yield	1	-
57	BAA	Moody's Seasoned Baa Corporate Bond Yield	1	-
58	NASDAQCOM	NASDAQ Composite Index	5	-
59	EXSZUS	Switzerland / U.S. Foreign Exchange Rate	5	-
60	EXJPUS	Japan / U.S. Foreign Exchange Rate	5	-
61	EXUSUK	U.S. / U.K. Foreign Exchange Rate	5	-
62	EXCAUS	Canada / U.S. Foreign Exchange Rate	5	-
63	sTB3MS	TB3MS - FEDFUNDS	1	-
64	sTB6MS	TB6MS - FEDFUNDS	1	-
65	sGS1	GS1 - FEDFUNDS	1	-
66	sGS3	GS3 - FEDFUNDS	1	-
67	sGS5	GS5 - FEDFUNDS	1	-
68	sGS10	GS10 - FEDFUNDS	1	-
69	sMPRIME	MPRIME - FEDFUNDS	1	-
70	sAAA	AAA - FEDFUNDS	1	-
71	sBAA	BAA - FEDFUNDS	1	-
72	AAA10Y	AAA - GS10	1	-
73	BAA10Y	BAA - GS10	1	-
74	AAA30Y	AAA - GS30	1	-
75	BAA30Y	BAA - GS30	1	-
76	VXOCLS	CBOE S&P100 Volatility Index	1	-

A4. Monthly Financial Data

The 147 monthly financial series in this dataset are versions of the financial dataset used in Jurado, Ludvigson and Ng (2015), Ludvigson, Ma and Ng (2021), and Bianchi, Ludvigson and Ma (2022). It consists of a number of indicators measuring the behavior of a broad cross-section of asset returns, as well as some aggregate financial indicators not included in the macro dataset. These data include valuation ratios such as the dividend-price ratio or the price-earnings ratio, dividends and prices, and a group of variables we call "risk factors," since they have been used in cross-sectional or time-series studies to uncover variation in the market risk-premium. These risk-facors include the three Fama and French (1993) risk factors, namely the excess return on the market MKT_t, the "small-minus-big" (SMB_t) and "high-minus-low" (HML_t) portfolio returns, the momentum factor UMD_t, and the small stock value spread $R15 - R11$. Furthermore and following Fama and French (1992), returns on 100 portfolios are sorted into 10 size and 10 book-market categories and included to the dataset. The data source for prices, dividends, and the dividend-price ratio is the Center for Research in Security Prices (Center for Research in Security Prices, 2022). The price-earnings ratio is from Robert Shiller's website (Shiller, 2022). The portfolios and risk factors are available on Kenneth French's Darthmouth website (French, 2022). The raw data used to form factors are always transformed to achieve stationarity. In addition, when forming forecasting factors from the large macro and financial datasets, the raw data (which are in different units) are standardized before performing PCA. More details follow in Appendix C.

All returns are expressed in logs (i.e., the log of the gross return), are displayed in percent (i.e., multiplied by 100), and are annualized by multiplying by 12, i.e., if x is the original return, I transform $1200 \ln(1+x/100)$. The data series used in this dataset are listed below by data source. Additional details on data transformations are given below the table.

I convert monthly data to quarterly by using either the middle-of-quarter values. The decision to use middle-of-quarter values depends on the survey deadline of a particular forecast date. In the analysis, I use the survey forecasts of the last month of the quarter which are conducted at the end of the second month. Hence, it is conceivable that the forecasters would have information about the second month of quarter t. Therefore, I use the middle-of-quarter values.

I indicate data transformations as follows: the column *Tcode* shows the transformation I apply to a series: 1 – no transformation (levels); 2 – first difference; 4 – logarithms; 5 – first difference of logarithms; 6 – second difference in logarithms.

Table A3: List of Monthly Financial Dataset Variables.

#	Mnemonic	Description	Tcode
Prices and Dividends			
1	D_log(DIV)	$\Delta \log D_t^*$ see additional details below	5
2	D_log(P)	$\Delta \log P_t$ see additional details below	5
3	D_log(DIVre)	$\Delta \log D_t^{re,*}$ see additional details below	5
4	D_log(Pre)	$\Delta \log P_t^{re}$ see additional details below	5
5	d-p	$log(D_t^* - log(P_t)$ see additional details below	4
6	p-e	Price/earnings ratio	1
Equity Risk Factors			
7	R15-R11	(Small,High) minus (Small,Low) sorted on (size,book-to-market)	1
8	Mkt-RF	Market excess returns	1
9	SMB	Small Minus Big, sorted on size	1
10	HML	High Minus Low, sorted on book-to-market	1
11	UMD	Up Minus Down, sorted on momentum	1
Industries			
12	Agric	Agric industry portfolio	1
13	Food	Food industry portfolio	1
14	Beer	Beer industry portfolio	1

Continued on next page

Table A3 – *Continued from previous page*

#	Mnemonic	Description	Tcode
15	Smoke	Smoke industry portfolio	1
16	Toys	Toys industry portfolio	1
17	Fun	Fun industry portfolio	1
18	Books	Books industry portfolio	1
19	Hshld	Hshld industry portfolio	1
20	Clths	Clths industry portfolio	1
21	MedEq	MedEq industry portfolio	1
22	Drugs	Drugs industry portfolio	1
23	Chems	Chems industry portfolio	1
24	Rubbr	Rubbr industry portfolio	1
25	Txtls	Txtls industry portfolio	1
26	BldMt	BldMt industry portfolio	1
27	Cnstr	Cnstr industry portfolio	1
28	Steel	Steel industry portfolio	1
29	Mach	Mach industry portfolio	1
30	ElcEq	ElcEq industry portfolio	1
31	Autos	Autos industry portfolio	1
32	Aero	Aero industry portfolio	1
33	Ships	Ships industry portfolio	1
34	Mines	Mines industry portfolio	1
35	Coal	Coal industry portfolio	1
36	Oil	Oil industry portfolio	1
37	Util	Util industry portfolio	1
38	Telcm	Telcm industry portfolio	1
39	PerSv	PerSv industry portfolio	1
40	BusSv	BusSv industry portfolio	1
41	Hardw	Hardw industry portfolio	1
42	Chips	Chips industry portfolio	1
43	LabEq	LabEq industry portfolio	1
44	Paper	Paper industry portfolio	1
45	Boxes	Boxes industry portfolio	1
46	Trans	Trans industry portfolio	1
47	Whlsl	Whlsl industry portfolio	1
48	Rtail	Rtail industry portfolio	1
49	Meals	Meals industry portfolio	1
50	Banks	Banks industry portfolio	1
51	Insur	Insur industry portfolio	1
52	RlEst	RlEst industry portfolio	1
53	Fin	Fin industry portfolio	1
54	Oth	Oth industry portfolio	1
Size/BM			
55	1_2	(1,2) portfolio sorted on (size, book-to-market)	1
56	1_4	(1,4) portfolio sorted on (size, book-to-market)	1

Continued on next page

Table A3 – *Continued from previous page*

#	Mnemonic	Description	Tcode
57	1_5	(1,5) portfolio sorted on (size, book-to-market)	1
58	1_6	(1,6) portfolio sorted on (size, book-to-market)	1
59	1_7	(1,7) portfolio sorted on (size, book-to-market)	1
60	1_8	(1,8) portfolio sorted on (size, book-to-market)	1
61	1_9	(1,9) portfolio sorted on (size, book-to-market)	1
62	1_high	(1,high) portfolio sorted on (size, book-to-market)	1
63	2_low	(2,low) portfolio sorted on (size, book-to-market)	1
64	2_2	(2,2) portfolio sorted on (size, book-to-market)	1
65	2_3	(2,3) portfolio sorted on (size, book-to-market)	1
66	2_4	(2,4) portfolio sorted on (size, book-to-market)	1
67	2_5	(2,5) portfolio sorted on (size, book-to-market)	1
68	2_6	(2,6) portfolio sorted on (size, book-to-market)	1
69	2_7	(2,7) portfolio sorted on (size, book-to-market)	1
70	2_8	(2,8) portfolio sorted on (size, book-to-market)	1
71	2_9	(2,9) portfolio sorted on (size, book-to-market)	1
72	2_high	(2,high) portfolio sorted on (size, book-to-market)	1
73	3_low	(3,low) portfolio sorted on (size, book-to-market)	1
74	3_2	(3,2) portfolio sorted on (size, book-to-market)	1
75	3_3	(3,3) portfolio sorted on (size, book-to-market)	1
76	3_4	(3,4) portfolio sorted on (size, book-to-market)	1
77	3_5	(3,5) portfolio sorted on (size, book-to-market)	1
78	3_6	(3,6) portfolio sorted on (size, book-to-market)	1
79	3_7	(3,7) portfolio sorted on (size, book-to-market)	1
80	3_8	(3,8) portfolio sorted on (size, book-to-market)	1
81	3_9	(3,9) portfolio sorted on (size, book-to-market)	1
82	3_high	(3,high) portfolio sorted on (size, book-to-market)	1
83	4_low	(4,low) portfolio sorted on (size, book-to-market)	1
84	4_2	(4,2) portfolio sorted on (size, book-to-market)	1
85	4_3	(4,3) portfolio sorted on (size, book-to-market)	1
86	4_4	(4,4) portfolio sorted on (size, book-to-market)	1
87	4_5	(4,5) portfolio sorted on (size, book-to-market)	1
88	4_6	(4,6) portfolio sorted on (size, book-to-market)	1
89	4_7	(4,7) portfolio sorted on (size, book-to-market)	1
90	4_8	(4,8) portfolio sorted on (size, book-to-market)	1
91	4_9	(4,9) portfolio sorted on (size, book-to-market)	1
92	4_high	(4,high) portfolio sorted on (size, book-to-market)	1
93	1_low	(5,low) portfolio sorted on (size, book-to-market)	1
94	5_2	(5,2) portfolio sorted on (size, book-to-market)	1
95	5_3	(5,3) portfolio sorted on (size, book-to-market)	1
96	5_4	(5,4) portfolio sorted on (size, book-to-market)	1
97	5_5	(5,5) portfolio sorted on (size, book-to-market)	1
98	5_6	(5,6) portfolio sorted on (size, book-to-market)	1
99	5_7	(5,7) portfolio sorted on (size, book-to-market)	1
100	5_8	(5,8) portfolio sorted on (size, book-to-market)	1

Continued on next page

Table A3 – *Continued from previous page*

#	Mnemonic	Description	Tcode
101	5_9	(5,9) portfolio sorted on (size, book-to-market)	1
102	5_high	(5,high) portfolio sorted on (size, book-to-market)	1
103	6_low	(6,low) portfolio sorted on (size, book-to-market)	1
104	6_2	(6,2) portfolio sorted on (size, book-to-market)	1
105	6_3	(6,3) portfolio sorted on (size, book-to-market)	1
106	6_4	(6,4) portfolio sorted on (size, book-to-market)	1
107	6_5	(6,5) portfolio sorted on (size, book-to-market)	1
108	6_6	(6,6) portfolio sorted on (size, book-to-market)	1
109	6_7	(6,7) portfolio sorted on (size, book-to-market)	1
110	6_8	(6,8) portfolio sorted on (size, book-to-market)	1
111	6_9	(6,9) portfolio sorted on (size, book-to-market)	1
112	6_high	(6,high) portfolio sorted on (size, book-to-market)	1
113	7_low	(1,low) portfolio sorted on (size, book-to-market)	1
114	7_2	(7,2) portfolio sorted on (size, book-to-market)	1
115	7_3	(7,3) portfolio sorted on (size, book-to-market)	1
116	7_4	(7,4) portfolio sorted on (size, book-to-market)	1
117	7_5	(7,5) portfolio sorted on (size, book-to-market)	1
118	7_6	(7,6) portfolio sorted on (size, book-to-market)	1
119	7_7	(7,7) portfolio sorted on (size, book-to-market)	1
120	7_8	(7,8) portfolio sorted on (size, book-to-market)	1
121	7_9	(7,9) portfolio sorted on (size, book-to-market)	1
122	8_low	(8,low) portfolio sorted on (size, book-to-market)	1
123	8_2	(8,2) portfolio sorted on (size, book-to-market)	1
124	8_3	(8,3) portfolio sorted on (size, book-to-market)	1
125	8_4	(8,4) portfolio sorted on (size, book-to-market)	1
126	8_5	(8,5) portfolio sorted on (size, book-to-market)	1
127	8_6	(8,6) portfolio sorted on (size, book-to-market)	1
128	8_7	(8,7) portfolio sorted on (size, book-to-market)	1
129	8_8	(8,8) portfolio sorted on (size, book-to-market)	1
130	8_9	(8,9) portfolio sorted on (size, book-to-market)	1
131	8_high	(8,high) portfolio sorted on (size, book-to-market)	1
132	9_low	(9,low) portfolio sorted on (size, book-to-market)	1
133	9_2	(9,2) portfolio sorted on (size, book-to-market)	1
134	9_3	(9,3) portfolio sorted on (size, book-to-market)	1
135	9_4	(9,4) portfolio sorted on (size, book-to-market)	1
136	9_5	(9,5) portfolio sorted on (size, book-to-market)	1
137	9_6	(9,6) portfolio sorted on (size, book-to-market)	1
138	9_7	(9,7) portfolio sorted on (size, book-to-market)	1
139	9_8	(9,8) portfolio sorted on (size, book-to-market)	1
140	9_high	(9,high) portfolio sorted on (size, book-to-market)	1
141	10_low	(10,low) portfolio sorted on (size, book-to-market)	1
142	10_2	(10,2) portfolio sorted on (size, book-to-market)	1
143	10_3	(10,3) portfolio sorted on (size, book-to-market)	1
144	10_4	(10,4) portfolio sorted on (size, book-to-market)	1

Continued on next page

Table A3 – *Continued from previous page*

#	Mnemonic	Description	Tcode
145	10_5	(10,5) portfolio sorted on (size, book-to-market)	1
146	10_6	(10,6) portfolio sorted on (size, book-to-market)	1
147	10_7	(10,7) portfolio sorted on (size, book-to-market)	1

CRSP Data Details. Value-weighted price and dividend data are obtained from CRSP. From the Annual Update data, I obtain monthly value-weighted returns series vwretd (with dividends) and vwretx (excluding dividends).[4] These series have the interpretation

$$VWRETD_t = \frac{P_{t+1} + D_{t+1}}{P_t} \tag{A.5}$$

$$VWRETX_t = \frac{P_{t+1}}{P_t} \tag{A.6}$$

From these series, a normalized price series P_t can be constructed using the recusion

$$P_0 = 1, \tag{A.7}$$

$$P_t = P_{t-1} * VWRETX_t. \tag{A.8}$$

A dividend series can then be constructed using

$$D_t = P_{t-1}(VWRETD_t - VWRETX_t). \tag{A.9}$$

In order to remove seasonality of dividend payments from the data, instead of D_t we use the series

$$D_t^* = \frac{1}{12} \sum_{j=0}^{11} D_{t-j}, \tag{A.10}$$

i.e., the moving average over the entire year. For the price and dividend series under "reinvestment," we calculate the price under reinvestment, P_t^{re}, as the normalized value of the market portfolio under reinvestment of dividends, using the recursion

$$P_0^{re} = 1, \tag{A.11}$$

$$P_t^{re} = P_{t-1}^{re} * VWRETX_t. \tag{A.12}$$

Similarly, we can define dividends under reinvestment, D_t^{re}, as the total dividend payments on this portfolio (the number of "shares" of which have increased over time) using

$$D_t^{re} = P_{t-1}^{re}(VWRETD_t - VWRETX_t). \tag{A.13}$$

As before, we can remove seasonality by using

$$D_t^{re,*} = \frac{1}{12} \sum_{j=0}^{11} D_{t-j}^{re}. \tag{A.14}$$

Five data series are constructed from the CRSP data as follows:

[4] Note that both returns are defined as gross returns, while CRSP may provide you with net returns.

(i) `D_log(DIV)`: $\Delta \log D_t^*$,

(ii) `D_log(P)`: $\Delta \log P_t$,

(iii) `D_log(DIVre)`: $\Delta \log D_t^{re,*}$,

(iv) `D_log(Pre)`: $\Delta \log P_t^{re}$,

(v) `d-p`: $\log(D_t^*) - \log(P_t)$.

Kenneth French Data Details. The following data are obtained from the data library of Kenneth French's Dartmouth website (French, 2022):

(i) Fama/French Factors: From this dataset, I obtain the data series `Mkt-RF`, `SMB`, `HML`.

(ii) 25 portfolios formed on size and book-to-market (5×5): From this dataset, I obtain the series `R15-R11`, which is the spread between the (small, high) book-to-market and (small,low) book-to-market portfolio.

(iii) Momentum factor: From this dataset, I obtain the series `UMD`, which is equal to the momentum factor.

(iv) 49 industry portfolios: From this dataset, I use all value-weighted series excluding any series that have missing observations from January 1960 on. The omitted series are: Soda, Hlth, FabPr, Guns, Gold, Softw.

(v) 100 portfolios formed in size and book-to-market: From this dataset, I use all value-weighted series excluding any series that have missing observations from January 1960 on. This yields variables with the name X_Y where X stands for the index of the size variable (1, 2, . . . , 10) and Y stands for the index of the book-to-market variable (Low,2, 3, . . . , 9,High). The omitted series are 1_low, 1_3, 7_high, 9_9, 10_8, 10_9, and 10_high.

B. Propositions and Proofs

This section provides the proofs discussed in the paper. Proposition 1 and 2 are essentially from Bordalo et al. (2020) but are stated for the sake of completeness.

PROPOSITION 1: *The distorted density $f^{\theta}(y_t \mid s_{it})$ is normal. For $\rho > 0$ and in the steady state, it is characterized by a time-varying mean $\mathbb{F}^{\theta}_{it} y_t = y^{\theta}_{it|t}$ and a constant variance $\frac{\Sigma\sigma^2_{\varepsilon}}{\Sigma+\sigma^2_{\varepsilon}}$, where*

$$\mathbb{F}^{\theta}_{it} y_t = y^{\theta}_{it|t} = y_{it|t} + \theta K \left(s_{it} - y_{it|t-1}\right), \qquad K = \frac{\Sigma}{\Sigma + \sigma^2_{\varepsilon}}, \tag{B.1}$$

$$\Sigma = \frac{1}{2}(\sigma^2_u - (1-\rho^2)\sigma^2_{\varepsilon} + \sqrt{\left[(1-\rho^2)\sigma^2_{\varepsilon} - \sigma^2_u\right] + 4\sigma^2_{\varepsilon}\sigma^2_u}. \tag{B.2}$$

Proof of Proposition 1. The data generating process is $y_t = \rho y_{t-1} + u_t$, where $u_t \sim \mathcal{N}(0, \sigma^2_u)$ i.i.d. over time. Forecaster i observes a noisy signal $s_{it} = y_t + \varepsilon_{it}$, where $\varepsilon_{it} \sim \mathcal{N}(0, \sigma^2_{\varepsilon})$ is i.i.d. over time and space. S_{it} denotes the full history of the privately observed signals. Rational expectations are obtained iteratively by applying Bayes rule

$$f(y_t \mid S_{it}) = \frac{f(y_t \mid S_{it-1}) \times f(s_{it} \mid y_t)}{f(s_{it})}.$$

Defining these densities as follows

$$f(y_t \mid S_{it-1}) = (2\pi)^{-1/2}(\Sigma_{t|t-1})^{-1/2} \exp\left\{-\frac{1}{2\Sigma_{t|t-1}}(y_t - y_{it|t-1})^2\right\},$$

$$f(s_{it} \mid y_t) = (2\pi)^{-1/2}(\sigma^2_{\varepsilon})^{-1/2} \exp\left\{-\frac{1}{2\sigma^2_{\varepsilon}}(s_{it} - y_t)^2\right\},$$

$$f(s_{it}) = (2\pi)^{-1/2}(\Sigma_{t|t-1} + \sigma^2_{\varepsilon})^{-1/2} \exp\left\{-\frac{1}{2(\Sigma_{t|t-1} + \sigma^2_{\varepsilon})}(s_{it} - y_{it|t-1})^2\right\},$$

and combining them according to Bayes rule, yields

$$f(y_t \mid s_{it}) = (2\pi)^{-1/2}\left(\frac{\Sigma_{t|t-1}\sigma^2_{\varepsilon}}{\Sigma_{t|t-1} + \sigma^2_{\varepsilon}}\right)^{-1/2} \times \exp\left\{-\frac{1}{2}Q\right\},$$

where Q is denoted as follows

$$\begin{aligned}
Q &= \frac{(y_t - y_{it|t-1})^2}{\Sigma_{t|t-1}} + \frac{(s_{it} - y_t)^2}{\sigma^2_{\varepsilon}} - \frac{(s_{it} - y_{it|t-1})^2}{\Sigma_{t|t-1} + \sigma^2_{\varepsilon}} \\
&= \frac{(y_t - y_{it|t-1})^2}{\Sigma_{t|t-1}} + \frac{(s_{it} - y_{it|t-1} - y_t + y_{it|t-1})^2}{\sigma^2_{\varepsilon}} - \frac{(s_{it} - y_{it|t-1})^2}{\Sigma_{t|t-1} + \sigma^2_{\varepsilon}} \\
&= \frac{(y_t - y_{it|t-1})^2}{\Sigma_{t|t-1}} + \frac{(s_{it} - y_{it|t-1})^2 - 2(s_{it} - y_{it|t-1})y_t + 2(s_{it} - y_{it|t-1})y_{it|t-1} + y_t^2 - 2y_t y_{it|t-1} + (y_{it|t-1})^2}{\sigma^2_{\varepsilon}} \\
&\quad - \frac{(s_{it} - y_{it|t-1})^2}{\Sigma_{t|t-1} + \sigma^2_{\varepsilon}} \\
&= \frac{(y_t - y_{it|t-1})^2}{\Sigma_{t|t-1}} + \frac{(y_t - y_{it|t-1})^2}{\sigma^2_{\varepsilon}} - \frac{2(s_{it} - y_{it|t-1})(y_t - y_{it|t-1})}{\sigma^2_{\varepsilon}} + \frac{(s_{it} - y_{it|t-1})^2}{\sigma^2_{\varepsilon}} - \frac{(s_{it} - y_{it|t-1})^2}{\Sigma_{t|t-1} + \sigma^2_{\varepsilon}}
\end{aligned}$$

$$
\begin{aligned}
&= \left(\frac{1}{\Sigma_{t|t-1}} + \frac{1}{\sigma_\varepsilon^2}\right)(y_t - y_{it|t-1})^2 - \frac{2(s_{it} - y_{it|t-1})(y_t - y_{it|t-1})}{\sigma_\varepsilon^2} + \left(\frac{1}{\sigma_\varepsilon^2} - \frac{1}{\Sigma_{t|t-1} + \sigma_\varepsilon^2}\right)(s_{it} - y_{it|t-1})^2 \\
&= \frac{\sigma_\varepsilon^2 + \Sigma_{t|t-1}}{\sigma_\varepsilon^2 \Sigma_{t|t-1}}(y_t - y_{it|t-1})^2 - \frac{2(s_{it} - y_{it|t-1})(y_t - y_{it|t-1})}{\sigma_\varepsilon^2} + \frac{\Sigma_{t|t-1}}{\sigma_\varepsilon^2(\Sigma_{t|t-1} + \sigma_\varepsilon^2)}(s_{it} - y_{it|t-1})^2 \\
&= \frac{\Sigma_{t|t-1} + \sigma_\varepsilon^2}{\Sigma_{t|t-1}\sigma_\varepsilon^2}\left[(y_t - y_{it|t-1})^2 - 2\frac{\Sigma_{t|t-1}}{\Sigma_{t|t-1} + \sigma_\varepsilon^2}(y_t - y_{it|t-1})(s_{it} - y_{it|t-1}) + \frac{\Sigma_{t|t-1}}{(\Sigma_{t|t-1} + \sigma_\varepsilon^2)^2}(s_{it} - y_{it|t-1})^2\right] \\
&= \frac{\Sigma_{t|t-1} + \sigma_\varepsilon^2}{\Sigma_{t|t-1}\sigma_\varepsilon^2}\left(y_t - y_{it|t-1} - \frac{\Sigma_{t|t-1}}{\Sigma_{t|t-1} + \sigma_\varepsilon^2}(s_{it} - y_{it|t-1})\right)^2,
\end{aligned}
$$

yields the distribution $f(y_t \mid s_{it}) \sim \mathcal{N}\left(y_{it|t}, \Sigma_{t|t}\right)$ with

$$
\begin{aligned}
&y_{it|t} = y_{it|t-1} + K(s_{it} - y_{it|t-1}), \qquad K = \frac{\Sigma_{t|t-1}}{\Sigma_{t|t-1} + \sigma_\varepsilon^2}, \\
&\Sigma_{t|t} = \frac{\Sigma_{t|t-1}\sigma_\varepsilon^2}{\Sigma_{t|t-1} + \sigma_\varepsilon^2}.
\end{aligned}
$$

Here, K refers to the Kalman gain. $\Sigma_{t|t-1}$ is the variance of the prior $f(y_t | S_{it-1})$. The variance of $f(y_{t+1} | S_{it})$ is:

$$
\Sigma_{t+1|t} = \text{var}(\rho y_t + u_{t+1}) = \rho^2 \frac{\Sigma_{t|t-1}\sigma_\varepsilon^2}{\Sigma_{t|t-1} + \sigma_\varepsilon^2} + \sigma_u^2.
$$

Hence, the steady state variance, where $\Sigma = \Sigma_{t+1|t} = \Sigma_{t|t-1}$ is equal to

$$
\begin{aligned}
\Sigma &= \rho^2 \frac{\Sigma\sigma_\varepsilon^2}{\Sigma + \sigma_\varepsilon^2} + \sigma_u^2 \\
\Sigma^2 + \Sigma\sigma_\varepsilon^2 &= \rho^2\Sigma\sigma_\varepsilon^2 + \sigma_u^2\Sigma + \sigma_\varepsilon^2\sigma_u^2 \\
\Sigma^2 + \Sigma\sigma_\varepsilon^2 &= \Sigma(\rho^2\sigma_\varepsilon^2 + \sigma_u^2) + \sigma_\varepsilon^2\sigma_u^2 \\
\Sigma^2 + \Sigma(\sigma_\varepsilon^2 - \rho^2\sigma_\varepsilon^2 - \sigma_u^2) - \sigma_\varepsilon^2\sigma_u^2 &= 0 \\
\Sigma &= -\frac{\sigma_\varepsilon^2 - \rho^2\sigma_\varepsilon^2 - \sigma_u^2}{2} \pm \sqrt{\frac{(\sigma_\varepsilon^2 - \rho^2\sigma_\varepsilon^2 - \sigma_u^2)^2}{4} - \sigma_\varepsilon^2\sigma_u^2} \\
\Sigma &= -\frac{\sigma_\varepsilon^2 - \rho^2\sigma_\varepsilon^2 - \sigma_u^2 \pm \sqrt{\sigma_\varepsilon^4 - 2\rho^2\sigma_\varepsilon^4 + 2\sigma_\varepsilon^2\sigma_u^2 + \rho^4\sigma_\varepsilon^4 + 2\rho^2\sigma_\varepsilon^2\sigma_u^2 + \sigma_u^4}}{2} \\
\Sigma &= \frac{1}{2}(\sigma_u^2 - (1-\rho^2)\sigma_\varepsilon^2 + \sqrt{\left[(1-\rho^2)\sigma_\varepsilon^2 - \sigma_u^2\right] + 4\sigma_\varepsilon^2\sigma_u^2}.
\end{aligned}
$$

Finally, this yields the following distribution

$$
\begin{aligned}
&f(y_t \mid s_{it}) \sim \mathcal{N}\left(y_{it|t}, \frac{\Sigma\sigma_\varepsilon^2}{\Sigma + \sigma_\varepsilon^2}\right) \\
&\qquad y_{it|t} = y_{it|t-1} + K(s_{it} - y_{it|t-1}), \qquad K = \frac{\Sigma}{\Sigma + \sigma_\varepsilon^2}, \\
&\qquad \Sigma = \frac{1}{2}(\sigma_u^2 - (1-\rho^2)\sigma_\varepsilon^2 + \sqrt{\left[(1-\rho^2)\sigma_\varepsilon^2 - \sigma_u^2\right] + 4\sigma_\varepsilon^2\sigma_u^2}.
\end{aligned}
$$

In a next step, we use the distorted probability distribution to derive diagnostic expectations. The forecaster i then overweighs representative states by using the distorted posterior

$$f^{\theta}(y_t \mid s_{it}) = f(y_t \mid s_{it}) \left(\frac{f(y_t \mid s_{it})}{f(y_t \mid s_{it-1} \cup \{y_{it|t-1}\})} \right)^{\theta} \frac{1}{Z_t},$$

where Z_t is a normalization factor ensuring that $f^{\theta}(y_t \mid s_{it})$ integrates to one. For $s_{it} = y_{it|t-1}$ we have that $x_{it|t} = x_{it|t-1} = \rho x_{it-1|t-1}$, so the involved distributions look as follows

$$f(y_t \mid s_{it}) = (2\pi)^{-1/2} \left(\frac{\Sigma\sigma_\varepsilon^2}{\Sigma + \sigma_\varepsilon^2} \right)^{-1/2} \exp\left\{ -\frac{1}{2\frac{\Sigma\sigma_\varepsilon^2}{\Sigma+\sigma_\varepsilon^2}} (y_t - y_{it|t})^2 \right\},$$

$$f(y_t \mid s_{it} \cup \{y_{it|t-1}\}) = (2\pi)^{-1/2} \left(\frac{\Sigma\sigma_\varepsilon^2}{\sigma_u^2 + \sigma_\varepsilon^2} \right)^{-1/2} \exp\left\{ -\frac{1}{2\frac{\Sigma\sigma_\varepsilon^2}{\Sigma+\sigma_\varepsilon^2}} (y_t - y_{it|t-1})^2 \right\}.$$

Hence,

$$f^{\theta}(y_t \mid s_{it}) = (2\pi)^{-1/2} \left(\frac{\Sigma\sigma_\varepsilon^2}{\Sigma + \sigma_\varepsilon^2} \right)^{-1/2} \exp\left\{ -\frac{1}{2\frac{\Sigma\sigma_\varepsilon^2}{\Sigma+\sigma_\varepsilon^2}} Q^{\theta} \right\} \frac{1}{Z_t},$$

where

$$\begin{aligned}
Q^{\theta} &= (y_t - y_{it|t})^2 + \theta\left[(y_t - y_{it|t})^2 - (y_t - y_{it|t-1})^2\right] \\
&= y_t^2 - 2y_t y_{it|t} + y_{it|t}^2 + \theta y_t^2 - 2\theta y_t y_{it|t} + \theta y_{it|t}^2 - \theta y_t^2 + 2\theta y_t y_{it|t-1} - \theta y_{it|t-1}^2 \\
&= y_t^2 - 2y_t(y_{it|t} + \theta(y_{it|t} - y_{it|t-1})) + (1+\theta)y_{it|t}^2 - \theta y_{it|t-1}^2 \\
&= y_t^2 - 2y_t(y_{it|t} + \theta(y_{it|t} - y_{it|t-1})) + c(y_{it|t}, y_{it|t-1}).
\end{aligned}$$

From this, we can directly infer the solution by noting that $c(y_{it|t}, y_{it|t-1})$ is a constant that does not depend on y_t. By taking the normalization $\int f(y_t \mid s_{it})dy = 1$ into account, we find that this yields the distribution $f^{\theta}(y_t \mid s_{it}) \sim \mathcal{N}\left(y_{it|t}^{\theta}, \frac{\Sigma\sigma_\varepsilon^2}{\Sigma+\sigma_\varepsilon^2}\right)$ with

$$y_{it|t}^{\theta} = y_{it|t} + \theta(y_{it|t} - y_{it|t-1}).$$

Taking the Kalman filter $y_{it|t}$, I can re-write

$$\mathbb{F}_{it}^{\theta} y_t = y_{it|t}^{\theta} = y_{it|t} + \theta K(s_{it} - y_{it|t-1}), \qquad K = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_\varepsilon^2},$$

$$\Sigma = \frac{1}{2}(\sigma_u^2 - (1-\rho^2)\sigma_\varepsilon^2 + \sqrt{\left[(1-\rho^2)\sigma_\varepsilon^2 - \sigma_u^2\right] + 4\sigma_\varepsilon^2\sigma_u^2}. \quad \blacksquare$$

Proposition 2: *For $\rho > 0$, under the steady state diagnostic Kalman filter, the estimated coefficients of regression Eq.* ((3.1)) *and Eq.* ((3.2)) *at the consensus and individual level, β_1^c and β_1^p, are given by*

$$\beta_1^c = \frac{cov\left(y_{t+h} - y_{t+h|t}^{\theta}, y_{t+h|t}^{\theta} - y_{t+h|t-1}^{\theta}\right)}{var\left(y_{t+h|t}^{\theta} - y_{t+h|t-1}^{\theta}\right)} = \left(\sigma_\varepsilon^2 - \theta\Sigma\right) g(\sigma_\varepsilon^2, \Sigma, \rho, \theta), \tag{B.3}$$

$$\beta_1^p = \frac{cov\left(y_{it+h} - y_{it+h|t}^{\theta}, y_{it+h|t}^{\theta} - y_{it+h|t-1}^{\theta}\right)}{var\left(y_{it+h|t}^{\theta} - y_{it+h|t-1}^{\theta}\right)} = -\frac{\theta(1+\theta)}{(1+\theta)^2 + \theta^2\rho^2}, \tag{B.4}$$

where $g(\sigma_\varepsilon^2, \Sigma, \rho, \theta) > 0$ is a function of parameters. Thus, for $\theta \in (0, \sigma_\varepsilon^2/\Sigma)$ the diagnostic Kalman filter entails a positive consensus coefficient $\beta_1^c > 0$ and a negative individual coefficient $\beta_1^p < 0$.

Proof of Proposition 2. The rational consensus estimate for the current state is equal to $\int y_{it|t} = y_{t|t} = y_{t|t-1} + K(y_t - y_{t|t-1})$. Similarly, the diagnostic filter of the consensus estimate is equal to $\int y_{it|t}^\theta = y_{t|t}^\theta = y_{t|t} + \theta(y_{t|t} - y_{t|t-1})$. Note that $y_t = y_{t|t-1} + \frac{1}{K}(y_{t|t} - y_{t|t-1})$. The consensus forecast error under rationality is defined as follows:

$$\begin{aligned}
y_t - y_{t|t} &= y_t - y_{t|t-1} - K(y_t - y_{t|t-1}) \\
&= y_{t|t-1} + \frac{1}{K}(y_{t|t} - y_{t|t-1}) - y_{t|t-1} - K\left(y_{t|t-1} + \frac{1}{K}(y_{t|t} - y_{t|t-1}) - y_{t|t-1}\right) \\
&= \frac{1}{K}(y_{t|t} - y_{t|t-1}) - \frac{K}{K}(y_{t|t} - y_{t|t-1}) \\
&= \frac{1-K}{K}(y_{t|t} - y_{t|t-1})
\end{aligned}$$

Hence, the forecast error under diagnosticity is equal to

$$\begin{aligned}
y_t - y_{t|t}^\theta &= y_t - y_{t|t} + \theta(y_t - y_{t|t-1}) \\
&= y_t - y_{t|t-1} + K(y_t - y_{t|t-1}) + \theta\left(y_{t|t-1} + K(y_t - y_{t|t-1}) - y_{t|t-1}\right) \\
&= y_t - y_{t|t-1} - K(y_t - y_{t|t-1}) - \theta K(y_t - y_{t|t-1}) \\
&= y_{t|t-1} + \frac{1}{K}(y_{t|t} - y_{t|t-1}) - y_{t|t-1} - K\left(y_{t|t-1} + \frac{1}{K}(y_{t|t} - y_{t|t-1}) - y_{t|t-1}\right) \\
&- \theta K\left(y_{t|t-1} + \frac{1}{K}(y_{t|t} - y_{t|t-1}) - y_{t|t-1}\right) \\
&= \frac{1}{K}(y_{t|t} - y_{t|t-1}) + \frac{K}{K}(y_{t|t} - y_{t|t-1}) - \theta(y_{t|t} - y_{t|t-1}) \\
&= \left(\frac{1-K}{K} - \theta\right)(y_{t|t} - y_{t|t-1})
\end{aligned}$$

As a next step, we derive the diagnostic consensus forecast revision. Note that $y_{t|t-1} = \rho y_{t-1|t-1}$ and that $y_{t-1|t-2} = \rho y_{t|t-2}$. Hence, the forecast revision is equal to

$$\begin{aligned}
y_{t|t}^\theta - y_{t|t-1}^\theta &= y_{t|t} + \theta(y_{t|t} - y_{t|t-1}) - \rho y_{t-1|t-1}^\theta \\
&= y_{t|t} + \theta(y_{t|t} - y_{t|t-1}) - \rho y_{t-1|t-1} + \rho\theta(y_{t-1|t-1} - y_{t-1|t-2}) \\
&= y_{t|t} + \theta(y_{t|t} - y_{t|t-1}) - y_{t|t-1} + \rho\theta(y_{t-1|t-1} - y_{t-1|t-2}) \\
&= (1+\theta)(y_{t|t} - y_{t|t-1}) + \rho\theta(y_{t-1|t-1} - y_{t-1|t-2})
\end{aligned}$$

Therefore, the consensus CG coefficient is given by

$$\begin{aligned}
\beta_1^c &= \frac{\mathrm{cov}\left[y_t - y_{t|t}^\theta, y_{t|t}^\theta - y_{t|t-1}^\theta\right]}{var\left[y_{t|t}^\theta - y_{t|t}^\theta\right]} \\
&= \left(\frac{1-K}{K} - \theta\right)\frac{cov\left[y_{t|t} - y_{t|t-1}, (1+\theta)(y_{t|t} - y_{t|t-1}) + \rho\theta(y_{t-1|t-1} - y_{t-1|t-2})\right]}{var\left[(1+\theta)(y_{t|t} - y_{t|t-1}) + \rho\theta(y_{t-1|t-1} - y_{t-1|t-2})\right]}.
\end{aligned}$$

We have then in the nominator

$$\begin{aligned}
&\mathrm{cov}\left[y_{t|t} - y_{t|t-1}, (1+\theta)(y_{t|t} - y_{t|t-1}) + \rho\theta(y_{t-1|t-1} - y_{t-1|t-2})\right] \\
&= (1+\theta)\mathrm{var}[y_{t|t} - y_{t|t-1}] + \rho\theta\mathrm{cov}[y_{t|t} - y_{t|t-1}, y_{t-1|t-1} - y_{t-1|t-2}],
\end{aligned}$$

and in the denominator (note that $\text{var}(y_{t|t} - y_{t|t-1}) = \text{var}(y_{t-1|t-1} - y_{t-1|t-2})$)

$$\begin{aligned}
&\text{var}[(1+\theta)(y_{t|t} - y_{t|t-1}) + \rho\theta(y_{t-1|t-1} - y_{t-1|t-2})] \\
&= [(1+\theta)^2 + \rho^2\theta^2]\text{var}[y_{t|t} - y_{t|t-1}] - 2(1+\theta)\rho\theta\text{cov}[y_{t|t} - y_{t|t-1}, y_{t-1|t-1} - y_{t-1|t-2}].
\end{aligned}$$

To compute the covariance between adjacent rational revisions, note that $y_{t|t} = y_{t|t-1} + K(y_t - y_{t|t-1})$ and $y_{t|t-1} = \rho y_{t-1|t-1}$. Thus,

$$\begin{aligned}
y_{t|t-1} &= \rho y_{t-1|t-1} = \rho[y_{t-1|t-2} + K(y_{t-1} - y_{t-1|t-2})] \\
&= y_{t|t-2} + K(\rho y_{t-1} - y_{t|t-2}).
\end{aligned}$$

This implies that

$$\begin{aligned}
y_{t|t} - y_{t|t-1} &= y_{t|t-1} + K(y_t - y_{t|t-1}) - y_{t|t-2} - K(\rho y_{t-1} - y_{t|t-2}) \\
&= y_{t|t-1} - Ky_{t|t-1} - y_{t|t-2} + Ky_{t|t-2} + K(y_t - \rho y_{t-1}) \\
&= (1-K)(y_{t|t-1} - y_{t|t-2}) + Ku_t \\
&= (1-K)\rho(y_{t-1|t-1} - y_{t-1|t-2}) + Ku_t
\end{aligned}$$

As a result,

$$cov(y_{t|t} - y_{t|t-1}, y_{t-1|t-1} - y_{t-1|t-2}) = (1-K)\rho\text{var}[y_{t|t} - y_{t|t-1}].$$

Coming back to the formula of the beta coefficient, we have

$$\begin{aligned}
\beta_1^c &= \left(\frac{1-K}{K} - \theta\right)\frac{\text{cov}\left[y_{t|t} - y_{t|t-1}, (1+\theta)(y_{t|t} - y_{t|t-1}) + \rho\theta(y_{t-1|t-1} - y_{t-1|t-2})\right]}{\text{var}\left[(1+\theta)(y_{t|t} - y_{t|t-1}) + \rho\theta(y_{t-1|t-1} - y_{t-1|t-2})\right]} \\
&= \left(\frac{1-K}{K} - \theta\right)\frac{(1+\theta)\text{var}\left[y_{t|t} - y_{t|t-1}\right] + \rho\theta(1-K)\rho\text{var}\left[y_{t|t} - y_{t|t-1}\right]}{[(1+\theta)^2 + \rho^2\theta^2]\text{var}\left[y_{t|t} - y_{t|t-1}\right] - 2(1+\theta)\rho\theta(1-K)\rho\text{var}\left[y_{t|t} - y_{t|t-1}\right]} \\
&= \left(\frac{1-K}{K} - \theta\right)\frac{(1+\theta) - \rho^2\theta(1-K)}{[(1+\theta)^2 + \rho^2\theta^2] - 2(1+\theta)\rho^2\theta(1-K)},
\end{aligned}$$

which is positive if and only if $1 - K > \theta K$. This can be seen through two observations. The sign of this expression hinges on the first term

$$\begin{aligned}
\frac{1-K}{K} - \theta &> 0 \\
\frac{1-K-\theta K}{K} &> 0
\end{aligned}$$

This only holds if $1 - K > \theta K$ and if the second term is always positive. For that to hold, I show that the nominator and denominator are positive. For the nominator, we directly see that

$$(1+\theta) > \rho^2\theta(1-K),$$

because $0 < \rho^2 < 1$ and $0 < K < 1$. For the denominator, we have

$$\begin{aligned}
[(1+\theta)^2 + \rho^2\theta^2] - 2(1+\theta)\rho^2\theta(1-K) &> 0 \\
1 + 2\theta + \theta^2 + \rho^2\theta^2 &> 2(1+\theta)\rho^2\theta(1-K) \\
\frac{1 + 2\theta + \theta^2 + \rho^2\theta^2}{2(1+\theta)\rho^2} &> 1 - K
\end{aligned}$$

We know that $0 < 1-K < 1$ and thus, the nominator has to be larger than the denominator that this expression holds. This is easy to verify since $\rho > 0$ and $\theta > 0$, and by choosing an extremely small number $\epsilon > 0$ we have

$$\frac{1 + small}{small} > 1 - K,$$

and hence, this expression holds. It is then easy to show that $\beta_1^c > 0$ if and only if $\theta < \sigma_\varepsilon^2/\Sigma$.

Next, we consider individual level forecasts. The coefficient at the individual level of regression forecast errors on forecast revisions is equal to

$$\beta_1^p = \frac{\text{cov}\left[y_t - y_{it|t}^\theta, y_{it|t}^\theta - y_{it|t-1}^\theta\right]}{\text{var}\left[y_{it|t}^\theta - y_{it|t-1}^\theta\right]}.$$

Hence, we have that

$$y_t - y_{it|t}^\theta = y_t - y_{it|t} - \theta(y_{it|t} - y_{it|t-1}),$$

and note that $y_{it|t-1} = \rho y_{it-1|t-1}$

$$\begin{aligned}
y_{it|t}^\theta - y_{it|t-1} &= y_{it|t} + \theta(y_{it|t} - y_{it|t-1}) - \rho\left[y_{it-1|t-1} + \theta(y_{it-1|t-1} - y_{it-1|t-2}\right] \\
&= y_{it|t} + \theta(y_{it|t} - y_{it|t-1}) - y_{it|t-1} - \rho\theta(y_{it-1|t-1} - y_{it-1|t-2} \\
&= (1+\theta)(y_{it|t} - y_{it|t-1}) - \rho\theta(y_{it-1|t-1} - y_{it-1|t-2}).
\end{aligned}$$

Hence, we have

$$\begin{aligned}
\beta_1^p &= \frac{\text{cov}\left[y_t - y_{it|t}^\theta, y_{it|t}^\theta - y_{it|t-1}^\theta\right]}{\text{var}\left[y_{it|t}^\theta - y_{it|t-1}^\theta\right]} \\
&= \frac{\text{cov}\left[(y_t - y_{it|t}) - \theta(y_{it|t} - y_{it|t-1}), (1+\theta)(y_{it|t} - y_{it|t-1}) - \rho\theta(y_{it-1|t-1} - y_{it-1|t-2}\right]}{\text{var}\left[(1+\theta)(y_{it|t} - y_{it|t-1}) - \rho\theta(y_{it-1|t-1} - y_{it-1|t-2}\right]} \\
&= \frac{\text{cov}\left[y_t - y_{it|t}, (1+\theta)(y_{it|t} - y_{it|t-1}) - \rho\theta(y_{it-1|t-1} - y_{it-1|t-2})\right] - \theta(1+\theta)\text{var}\left[y_{it|t} - y_{it|t-1}\right]}{(1+\theta)^2\text{var}\left[y_{it|t} - y_{it|t-1}\right] + \rho^2\theta^2\text{var}\left[y_{it|t} - y_{it|t-1}\right]} \\
&\quad \frac{+\theta^2\rho\text{cov}\left[y_{it|t} - y_{it|t-1}, y_{it-1|t-1} - y_{it-1|t-2}\right]}{-2(1+\theta)\rho\theta\text{cov}\left[y_{it|t} - y_{it|t-1}, y_{it-1|t-1} - y_{it-1,t-2}\right]} \\
&= \frac{-\theta(1+\theta)\text{var}\left[y_{it|t} - y_{it|t-1}\right]}{(1+\theta)^2\text{var}\left[y_{it|t} - y_{it|t-1}\right] + \rho^2\theta^2\text{var}\left[y_{it|t} - y_{it|t-1}\right]} \\
&= \frac{-\theta(1+\theta)}{(1+\theta)^2 + \rho^2\theta^2}
\end{aligned}$$

Overreaction is larger ($\beta_1^p < 0$) for series with lower persistence. Intuitively, when persistence is low, rational beliefs respond less to news and there is more scope for overreaction.

C. Estimation of the Machine Efficient Benchmark

The machine efficient benchmark, or ML expectations, specified in Equation ((4.2)) is to be estimated

$$y_{t+h} = \alpha_h + \sum_{j=0}^{p} \left(\phi_{jh} y_{t-j} + \boldsymbol{\delta}_{jh} \boldsymbol{G}_{t-j}\right) + u_{t+h}. \tag{C.1}$$

Here, y_{t+h} refers to the Aaa or Baa credit spread at horizon $t+h$ $(h > 0)$. α_h refers to the respective constant, ϕ_{jh} denotes the autoregressive coefficients for lag j, and $\boldsymbol{\delta}_{jh}$ summarizes the coefficients with respect to the additional information summarized in the r_G-dimensional vector $\boldsymbol{G}_t$. This additional information are latent common factors from two data datasets: a real-time macroeconomic dataset and a monthly financial dataset. The construction of the dynamic factors are described in C1 and I discuss the exact model specification further below. Lastly, u_{t+h} denotes the respective error term.

Alternatively, I also explore a specification in which I control for subjective expectations. This specification directly controls for the forecaster's information set at time t and this has not to be proxied with publicly available data. Nevertheless, this restricts both samples drastically for the forecasting exercise and is thus only an alternative specification. It reads as follows

$$y_{t+h} = \alpha_h^{(k)} + \beta_{jh}^{(k)} \mathbb{F}_t^{(k)}\left[y_{t+h}\right] + \sum_{j=0}^{p} \left(\phi_{jh}^{(k)} y_{t-j} + \boldsymbol{\delta}_{jh}^{(k)} \boldsymbol{G}_{t-j} + \boldsymbol{\gamma}_{jh}^{(k)} \boldsymbol{W}_{t-j}\right) + u_{t+h}^{(k)}, \tag{C.2}$$

where the kth percentile of the survey forecast distribution is added to the specification. Hence, the superscript (k) denotes the coefficients corresponding to adding the k-th percentile of the survey forecast distribution to the specification. Furthermore, an r_W-dimensional vector $\boldsymbol{W}_t$ of additional information is added to the specification, specified below.

Both model specifications can be conviently summarized as follows

$$y_{t+h} = \mathcal{Z}_t' \boldsymbol{B}_h + u_{t+h}, \tag{C.3}$$

where we want to find an estimator $\hat{\boldsymbol{B}}_h$ to construct a true out-of-sample forecast. Further below, I inspect different estimators for estimating $\hat{\boldsymbol{B}}_h$: Bayesian shrinkage estimator in C2, Elastic Net (EN) estimator in C3, and Bayesian Additive Regression Trees (BART) estimator in C4.

Model specification. The exact model specification is the same for both credit spreads. Hence, for y_{t+h} equal to the respective credit spread the forecasting model considers the following variables in $\mathcal{Z}_t = \left(1, y_t, \ldots, y_{t-p}, \boldsymbol{G}_t, \ldots, \boldsymbol{G}_{t-p}, \boldsymbol{W}_t, \ldots, \boldsymbol{W}_{t-p}\right)$, which is of dimension $r = 1 + (1 + r_G + r_W)p$.

(i) Intercept.

(ii) Lags of the dependent variable $(y_t, \ldots, y_{t-p})$.

(iii) Factors in $\boldsymbol{G}_t = \left(\boldsymbol{f}_{t,final}^{M}, \ldots, \boldsymbol{f}_{t-j,final}^{M}, \boldsymbol{f}_{t,final}^{F}, \ldots, \boldsymbol{f}_{t-j,final}^{F}\right)$ from two large datasets separately:

- $\boldsymbol{f}_{t-j,final}^{M}$, for $j = 0, \ldots, p$ are factors formed from a real-time macroeconomic indicators dataset with 76 real-time macroeconomic series; includes both monthly and quarterly series, with a monthly series converted to quarterly according to the method described in the data appendix A3,
- $\boldsymbol{f}_{t-j,final}^{F}$, for $j = 0, \ldots, p$ are factors formed from a financial dataset with 147 financial series described in the data appendix A4.

(iv) Additional information in $\boldsymbol{W}_t$:

- $\mathbb{F}^{(\mu)}\left[y_{t-1}\right]$, lagged value of the mean of the survey forecast distribution,

- $\mathbb{F}^{(25)}\left[y_{t-1}\right]$, lagged value of the 25th percentile of the survey forecast distribution,
- $\mathbb{F}^{(50)}\left[y_{t-1}\right]$, lagged value of the 25th percentile of the survey forecast distribution,
- $\mathbb{F}^{(75)}\left[y_{t-1}\right]$, lagged value of the 25th percentile of the survey forecast distribution,
- $var\left(\mathbb{F}\left[y_t\right]\right)$, lagged cross-sectional variance of the survey forecast distribution,
- $skew\left(\mathbb{F}\left[y_t\right]\right)$, lagged cross-sectional skeweness of the survey forecast distribution,
- $(VXO_t, VXO_t^2, VXO_t^3)$, defined as the CBOE S&P 100 volatility index included in levels, squared, and cubic terms.

Hence, for a usual estimation setup the dimensionality of the baseline model is of medium size. The preferred baseline specification uses $g^C = 10$ factors from the respective data category, thus $r_G = 2 \times 10$, and features two lags, $p = 2$ and includes no additional information, $r_W = 0$. Hence, $r = 1 + (1 + 20 + 0) \times 2 = 43$. In the alternative specification, $r_W = 9$ regressors are added to the specification.

C1. Dynamic Factor Estimation

Assume that $\mathcal{X}_t^C = (x_{1t}^C, \ldots, x_{N^C t}^C)'$ generically denote a dataset of economic information in some category C that is available in real-time analysis. Here, category C refers to real-time macroeonomic data or monthly financial data which are treated separately. It is assumed that each x_{jt}^C $(\forall j)$ has been suitably transformed (such as by taking logs and differencing) so as to render the series statitonary. I assume that $\mathcal{X}_t^C$ has an approximate factor structure taking the form

$$\mathcal{X}_t^C = \mathbf{\Lambda}^C \boldsymbol{f}_t^C + \boldsymbol{\nu}_t^C, \tag{C.4}$$

where $\boldsymbol{f}_t^C$ is an $q^C \times 1$ vector of latent common factors, $\mathbf{\Lambda}^C$ is a corresponding $N^C \times q^C$ factor loadings matrix, and $\boldsymbol{\nu}_t^C$ is a $N^C \times 1$ vector of idiosyncratic errors. In an approximate dynamic factor structure setting, the idiosyncratic errors $\boldsymbol{\nu}_t^C$ are permitted to have a limited amount of cross-sectional correlation. The number of factors $q^C << N^C$ is usually significant smaller than the total number of series, N^C, which facilitates the use of high-dimensional time series datasets. In the estimation procedure, I re-estimate the factor at each time point in the sample recursively over time using the entire history of data available in real-time prior to each out-of-sample forecast. The following steps are performed in forming the macro and financial factors:

(i) Transform each series to stationarity (by taking logs and differencing) according to the transformations in Table A2 and A3.

(ii) As a next step, I scale the factors along the procedure proposed in Huang, Jiang and Tong (2018). They proposed to scale each variable in the dataset as follows. Run the following regression for variable x_{jt}^C

$$y_{t+h} = \omega_{hj0}^C + \omega_{hjx}^C x_{jt}^C + \nu_{jt+h}^C, \quad \nu_{jt+h}^C \sim \mathcal{N}\left(0, (\sigma_{jt+h}^C)^2\right). \tag{C.5}$$

Then we scale each variable with its forecasting power (i.e., predictive regression slope). Hence, for estimating the factors we use $\tilde{x}_{jt}^C = \omega_{hjx}^C x_{jt}^C$.

(iii) Throughout, the factors are estimated over $\tilde{x}_{jt}^C$ by the method of static principal components (PCA). The factor structure follows Equation ((C.4)) but replaces $\mathcal{X}_t^C$ with $\tilde{\mathcal{X}}_t^C = (\tilde{x}_{1t}^C, \ldots, \tilde{x}_{Nt}^C)'$. The aim is to find the vector of latent common factors $\boldsymbol{f}_t^C$ and the corresponding matrix of factor loadings $\mathbf{\Lambda}^C$. Specifically, the $T \times g^C$ matrix $\hat{\boldsymbol{f}}_t^C$ is $\sqrt{T}$ times the g^C eigenvectors corresponding to the g^C largest eigenvalues of the $T \times T$ matrix $\tilde{\mathcal{X}}\tilde{\mathcal{X}}'/(TN)$ in decreasing order. In large samples (when $\sqrt{TN} \to \infty$), Bai and Ng (2006) show that the estimates of $\hat{\boldsymbol{f}}_t$ can be treated as though they were observed in the subsequent forecasting regressions.

(iv) Afterwards, I collect the common latent factors into the matrix $T \times g^C$ $\boldsymbol{f}^C_{t,raw}$, where each principal component is a column.

(v) Square the raw variables and repeat steps (ii)-(iv). Collect the common factors from the squared data matrix into a matrix $\boldsymbol{f}^C_{t,sqr}$, where each principal component is a column.

(vi) Then, I square the first factor in $\boldsymbol{f}^C_{raw}$ and call this $\boldsymbol{f}^{2,C}_{t,raw1}$.

(vii) Finally, I construct the matrix of factors as $\boldsymbol{f}^C_{t,final} = (\boldsymbol{f}^C_{t,raw}, \boldsymbol{f}^C_{t,sqr,1}, \boldsymbol{f}^{2,C}_{raw1})$, where $\boldsymbol{f}^C_{t,sqr1}$ denotes the first column of $\boldsymbol{f}^C_{t,sqr}$.

C2. Bayesian Shrinkage Estimator

As a first estimator, I rely on a linear estimator with a Gaussian error term. Estimation is performed in a Bayesian setting and extended with a shrinkage prior and a possible stochastic volatility specification. For the shrinkage prior, I rely on the horseshoe (HS) prior as proposed in Carvalho, Polson and Scott (2010), which offers the advantage of being free of user-chosen hyperparameters but is still highly adaptive and robust to a variety of situations. Furthermore, it is straightforward to add stochastic volatility to the model specification (Kim, Shephard and Chib, 1998; Jacquier, Polson and Rossi, 2002; Kastner and Frühwirth-Schnatter, 2014). Then, the model reads as follows

$$y_{t+h} = \mathcal{Z}_t' \boldsymbol{B}_h + \varepsilon_{t+h}, \quad \varepsilon_{t+h} \sim \mathcal{N}(0, \sigma^2_{t+h}), \tag{C.6}$$

where in the case of stochastic volatility, the law of motion for $\ln \sigma^2_{t+h}$ is a centered autoregressive process

$$\ln \sigma^2_{t+h} = \mu^\sigma_{t+h} + \varphi^\sigma_{t+h}(\ln \sigma^2_{t+h-1} - \mu^\sigma_{t+h}) + \xi^\sigma_{t+h}, \quad \xi^\sigma_{t+h} \sim \mathcal{N}\left(0, (\sigma^\sigma_{t+h})^2\right). \tag{C.7}$$

The horseshoe prior is given by for element j element of the vector $\boldsymbol{B}_h$ in the next equation. For the sake of brevity, I suppress the subscripts h referring to the hth forecasting horizon. Then, B_j denotes the jth element and the prior is given by

$$B_j \mid \lambda_j \sim \mathcal{N}\left(\underline{\mathrm{B}}_j, \lambda_j^2 \tau^2\right), \quad \lambda_j \sim C^+(0,1), \quad \tau \sim C^+(0,1), \tag{C.8}$$

where $C^+(0, a)$ denotes the half-Cauchy distribution on the positive reals with scale parameter a. λ_j denotes the *local* shrinkage paramter that is coefficient specific and τ is a *global* shrinkage term that pulls all elements in $\boldsymbol{B}_{ih}$ towards zero. Makalic and Schmidt (2015) provide a simple and efficient sampling scheme based on auxiliary variables that lead to conjugate conditional posterior distributions of all parameters. Regarding the priors of the stochastic volatility specification, I closely follow Kastner and Frühwirth-Schnatter (2014) and its implementation in Kastner (2016) and define $\mu^\sigma_{i,t+h} \sim \mathcal{N}(b_\mu, B_\mu)$, where $b_\mu = 0$ and $B_\mu = 100^2$ to be rather uninformative. Furthermore, the persistence paramter $\varphi^\sigma_{i,t+h} \in (-1,1)$, thus $(\varphi^\sigma_{i,t+h} + 1)/2 \sim \mathcal{B}(a_0, b_0)$, where $a_0 = 25$ and $b_0 = 1.5$ are positive hyperparameters and $\mathcal{B}(a,b)$ denotes the Beta-distribution with shape paremter a and b. The prior specification implies a prior mean of 0.94 and a prior standard deviation 0.04, implying a rather persistent volatility process. For the volatility of the log-variance, Kastner (2016) useses a Gamma distributed prior, i.e., $(\sigma^\sigma_{i,t+h})^2 \sim \mathcal{G}(1/2, 1/2B_\sigma)$, where the hyperparameter B_σ is not very influential and set to unity.

C3. Elastic Net Estimator

As a second estimator, I rely on the use the Elastic Net (EN) estimator, which combines Least Absolute Shrinkage and Selection Operator (LASSO) and ridge type penalties (Zou and Hastie, 2005) implemented by Kuhn (2022). Suppose the goal is to estimate the coefficients of the following linear model

$$y_{t+h} = \mathcal{Z}_t' \boldsymbol{B}_h + \varepsilon_{t+h}, \tag{C.9}$$

where all the independent variables are collected in the matrix $\mathcal{Z}_t' = (1, z_{1t}, \ldots, z_{Kt})'$ and all coefficients are summarized in the vector $\boldsymbol{B}_h = (b_0, b_1, \ldots, b_K)'$. Before the estimation all elements in $\boldsymbol{X}_t$ are standardized such that sample means are zero and sample standard deviation equals unity. The coefficient estimates are then put back in their original scale by multiplying the slope coefficients by their respective standard deviation, and adding back the mean (scaled by slope coefficient over standard deviation).

The naïve EN estimator $\hat{\boldsymbol{B}}_h$ incorporates both an L_1 and L_2 penalty and minimizes the equation

$$\begin{aligned}\hat{\boldsymbol{B}}_h &= \arg\min L\left(\lambda_{1h}, \lambda_{2h}, \boldsymbol{B}_h\right) \\ &= \arg\min_{b_0, b_1, \ldots, b_K} \left\{ \sum_{\tau=1}^{T} \left(y_{\tau+h} - \mathcal{Z}_\tau' \boldsymbol{B}_h\right)^2 + \underbrace{\lambda_{1h} \sum_{j=1}^{K} |b_{jh}|}_{LASSO} + \underbrace{\lambda_{2h} \sum_{j=1}^{K} b_{jh}^2}_{RIDGE} \right\}.\end{aligned} \tag{C.10}$$

By minimizing the RMSE over the training samples, I choose the optimal λ_{1h} and λ_{2h} values simultaneously.

C4. Bayesian Additive Regression Trees

A a third estimator, I use Bayesian Additive Regression Trees (BART, Chipman, George and McCulloch, 2010), which is a Bayesian approach to nonparametric function estimation using regression trees. Tree-based regression models have an ability to flexibly fit interactions and nonlinearities. Models composed of sums of regression trees have an even greater flexibility than single trees to capture interactions and nonlinearities as well as additive effects. The BART model is implemented by Kapelner and Bleich (2016). Hence, the BART model can be expressed as follows:

$$y_{t+h} = f(\mathcal{Z}_t) + \varepsilon_{t+h}, \quad \varepsilon_{t+h} \sim \mathcal{N}(0, \sigma_{t+h}^2) \tag{C.11}$$

where f is a potentially nonlinear function, which is approximated with a tree-structure

$$f(\mathcal{Z}_t) \approx \sum_{s=1}^{S} g_s\left(\mathcal{Z}_t \mid \mathcal{T}_s, \boldsymbol{\mu}_s\right). \tag{C.12}$$

There are S distinct regression trees, each composed of a single regression tree function g_s and the corresponding tree structure $\mathcal{T}_s$, and the parameters at the terminal nodes $\boldsymbol{\mu}_s$. The dimension of $\boldsymbol{\mu}_s$ is denoted by b_s and describes the complexity of the tree. The structure of a tree $\mathcal{T}_s$ includes information on how any observation recurses down the tree. For each nonterminal node of the tree, there is a splitting rule according to which the space of explanatory variables are selected into various disjoint regions using a sequence of binary rules. These take the form of $\{\mathcal{Z} \in \mathcal{A}_s\}$ or $\{\mathcal{Z} \notin \mathcal{A}_r\}$ with $\mathcal{A}_r$ being a partition set for $r = 1, \ldots, b$ and $\mathcal{Z} = (\mathcal{Z}_1, \ldots, \mathcal{Z}_T)$ a full-data matrix of dimension $T \times K$. Then the splitting rule is defined through a splitting value c, creating partition sets of the form $\{\mathcal{Z}_{1:T,i} \leq c\}$ or $\{\mathcal{Z}_{1:T,i} > c\}$. This process continues until a terminal node is reached. Then, the observation receives the leaf value of the terminal mode. The sum of the S leaf values becomes its predicted value. The tree's leaf value is given by

$$g(\mathcal{Z} \mid \mathcal{T}_s, \boldsymbol{\mu}_s) = \sum_{r=1}^{b} \mu_{s,r} \mathbb{I}\left[\mathcal{Z} \in \mathcal{A}_r\right]. \tag{C.13}$$

BART can be distinguished from other ensemble tree-of-trees models due to its underlying probability model. As a Bayesian model, BART consists of a set of priors for the structure and the leaf parameters and a likelihood for data in the terminal nodes. In the practical application, I perform cross-validation over the hyperparameters. In particular, the number of regression trees $S = \{50, 200\}$, the hyperparameters regulating the variability of the leaf parameter, and the error variance.

D. Estimation of the Macroeconomic Model

In this section, I briefly describe the estimation strategy of the macroeconomic model. The estimation of the VAR is based on a Bayesian framework with the Minnesota prior (Doan, Litterman and Sims, 1984; Litterman, 1986) and features stochastic volatility (Kim, Shephard and Chib, 1998; Jacquier, Polson and Rossi, 2002; Kastner and Frühwirth-Schnatter, 2014). Hence, following Equation ((5.1)), the reduced-form VAR(p) model reads

$$\boldsymbol{y}_t = \boldsymbol{c} + \boldsymbol{A}_1 \boldsymbol{y}_{t-1} + \ldots + \boldsymbol{A}_p \boldsymbol{y}_{t-p} + \boldsymbol{u}_t, \quad \boldsymbol{u}_t \sim \mathcal{N}_n(\boldsymbol{0}, \boldsymbol{\Sigma}_t), \tag{D.1}$$

where p is the lag order, $\boldsymbol{c}$ is an $n \times 1$ vector of constants, $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_p$ are $n \times n$ coefficient matrices, and $\boldsymbol{u}_t$ denotes an $n \times 1$ vector of reduced-form Gaussian distributed innovations with time-varying covariance matrix $\boldsymbol{\Sigma}_t$, factorized as follows $\boldsymbol{\Sigma}_t = \boldsymbol{H}^{-1}\boldsymbol{\Lambda}_t \boldsymbol{H}^{-1'}$. Collect all VAR coefficients in $\boldsymbol{\alpha} = (\boldsymbol{c}', \boldsymbol{A}_1', \ldots, \boldsymbol{A}_p')'$. $\boldsymbol{\Lambda}_t$ is a diagonal matrix with generic jth element λ_{jt} and $\boldsymbol{H}^{-1}$ is a lower-triangular matrix with ones on its main diagonal. By taking the logarithm of the elements on the main diagonal of the matrix $\boldsymbol{\Lambda}_t$, these elements follow a centered autoregressive process

$$\ln \lambda_{i,t} = \mu_i + \varphi_j(\ln \lambda_{i,t-1} - \mu_j) + \xi_{i,t}, \quad \xi_{i,t} \sim \mathcal{N}(0, \sigma^2_{i,\xi}), \quad i = 1, \ldots, n. \tag{D.2}$$

Gather the coefficients of the stochastic volatility specification in the following vectors: $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)$, $\boldsymbol{\varphi} = (\varphi_1, \ldots, \varphi_n)$, and $\boldsymbol{\sigma^2_\xi} = (\sigma^2_{1,\xi}, \ldots, \sigma^2_{n,\xi})$.

Estimation. For the estimation, we pursue the approach by Chan and Eisenstat (2018) and Chan (2022). For that, we re-write the system in its structural form:

$$\boldsymbol{H}\boldsymbol{y}_t = \tilde{\boldsymbol{x}}_t \tilde{\boldsymbol{\alpha}} + \boldsymbol{\varepsilon}_t, \qquad \boldsymbol{\varepsilon}_t \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Lambda}_t), \tag{D.3}$$

where $\tilde{\boldsymbol{x}}_t = (1, \boldsymbol{y}_{t-1}', \ldots, \boldsymbol{y}_{t-p}')$. We can easily recover the reduced-form parameters by $\boldsymbol{\alpha} = H^{-1}\tilde{\boldsymbol{\alpha}}$, the reduced-form covariance matrix $\boldsymbol{\Sigma}_t = \boldsymbol{H}^{-1}\boldsymbol{\Lambda}_t \boldsymbol{H}^{-1'}$, and reduced-form shocks by $\boldsymbol{u}_t = H^{-1}\boldsymbol{\varepsilon}_t$. Consequently, re-write the ith equation of the system as

$$y_{i,t} = \tilde{\boldsymbol{w}}_{i,t}\boldsymbol{h}_i + \tilde{\boldsymbol{x}}_t \tilde{\boldsymbol{\alpha}}_i + \varepsilon_{i,t}, \qquad \varepsilon_{i,t} \sim \mathcal{N}(0, \lambda^2_{i,t}), \tag{D.4}$$

where $\tilde{\boldsymbol{w}}_{i,t} = (-y_{1,t}, \ldots, -y_{i-1,t})$ and $\boldsymbol{h}_i$ are the elements first $i-1$ elements in the ith row of $\boldsymbol{H}$. Note that $y_{i,t}$ depends on the contemporaneous variables $y_{1,t}, \ldots, y_{i-1,t}$. I estimate the system in its triangular form and if we let $\boldsymbol{x}_{i,t} = (\tilde{\boldsymbol{w}}_{i,t}, \tilde{\boldsymbol{x}}_t)$, I can simplify to

$$y_{i,t} = \boldsymbol{x}_{i,t}\boldsymbol{\theta}_i + \varepsilon_{i,t}, \qquad \varepsilon_{i,t} \sim \mathcal{N}(0, \lambda^2_{i,t}), \tag{D.5}$$

where $\boldsymbol{\theta}_i = (\boldsymbol{h}_i', \tilde{\boldsymbol{\alpha}}_i')$ is of dimension $k_i = np + i$. This allows to estimate the VAR equation-by-equation and afterwards reduced-form coefficients can be backed out. Important to note here is that we specify priors directly on the *structural* coefficients and not the *reduced-form* coefficients. This variant of VAR estimation has no order invariance issues as in Carriero, Clark and Marcellino (2019) and Carriero et al. (2022).

Prior Specification. I have to elicit prior distribution on $(\boldsymbol{\theta}, \boldsymbol{\mu}, \boldsymbol{\varphi}, \boldsymbol{\sigma}^2_\xi)$. I assume that the parameters are a priori independent across equations, such that $p(\boldsymbol{\theta}, \boldsymbol{\mu}, \boldsymbol{\varphi}, \boldsymbol{\sigma}^2_\xi) = \prod_{i=1}^n p((\boldsymbol{\theta}_i, \mu_i, \varphi_i, \sigma^2_{i,\xi})$.

I assume that for $i = 1, \ldots, n$:

$$\boldsymbol{\theta}_i \sim \mathcal{N}(\boldsymbol{m}_i, \boldsymbol{V}_i). \tag{D.6}$$

Following Litterman (1986) and Sims and Zha (1998), we consider a Minneosta-type prior shrinkage prior setup for the VAR coefficients. First, we partition $\boldsymbol{m}_i = (\boldsymbol{m}_{i,\boldsymbol{h}}', \boldsymbol{m}_{i,\tilde{\alpha}}')$ and $\boldsymbol{V}_i = \operatorname{diag}\left(\boldsymbol{V}_{i,\boldsymbol{h}}, \boldsymbol{V}_{i,\tilde{\alpha}}\right)$, where $\boldsymbol{m}_{i,\boldsymbol{h}}$ and $\boldsymbol{V}_{i,\boldsymbol{h}}$ are the hyperparameters corresponding to $\boldsymbol{h}_i$, whereas $\boldsymbol{m}_{i,\tilde{\alpha}}$ and $\boldsymbol{V}_{i,\tilde{\alpha}}$ are those related to $\tilde{\boldsymbol{\alpha}}_i$. Hence, we set all elements in $\boldsymbol{m}_{i,\boldsymbol{h}} = \boldsymbol{0}$ and $\boldsymbol{V}_{i,\boldsymbol{h}} = 10$ which is relatively uninformative. For the hyperparameters related to $\tilde{\boldsymbol{\alpha}}_i$, we set $\boldsymbol{m}_{i,\tilde{\alpha}} = \boldsymbol{0}$ to shrink VAR coefficients towards zero. The coefficient associated with the first own lag is set to one for log-level data. Regarding the prior covariance matrix, we assume that $\boldsymbol{V}_{i,\tilde{\alpha}}$ is diagonal with the kth element $(\boldsymbol{V}_{i,\tilde{\alpha}})_k$ set to be

$$(\boldsymbol{V}_{i,\tilde{\alpha}})_k = \begin{cases} \left(\frac{\kappa_1}{l_3^{\kappa}}\right)^2, & \text{for the coefficient on the } l\text{th lag of variable } i, \\ \frac{s_i^2}{s_j^2}\left(\frac{\kappa_1\kappa_2}{l_3^{\kappa}}\right)^2, & \text{for the coefficient on the } l\text{th lag of variable } j,\ j \neq i, \\ s_i^2(\kappa_1\kappa_4)^2, & \text{for the deterministic terms,} \end{cases} \tag{D.7}$$

We use the following hyperparameter values: $\kappa_1 = 0.5$, $\kappa_2 = \kappa_3 = 1$, and $\kappa_4 = 1000^2$.

Regarding the priors of the stochastic volatility specification, I closely follow Kastner and Frühwirth-Schnatter (2014) and its implementation in Kastner (2016) and define $\mu_i \sim \mathcal{N}(b_\mu, B_\mu)$, where $b_\mu = 0$ and $B_\mu = 100^2$ to be rather uninformative. Furthermore, the persistence paramter $\varphi_i \in (-1, 1)$, thus $(\varphi_i + 1)/2 \sim \mathcal{B}(a_0, b_0)$, where $a_0 = 25$ and $b_0 = 1.5$ are positive hyperparameters and $\mathcal{B}(a, b)$ denotes the Beta-distribution with shape paremter a and b. The prior specification implies a prior mean of 0.94 and a prior standard deviation 0.04, implying a rather persistent volatility process. For the volatility of the log-variance, Kastner (2016) useses a Gamma distributed prior, i.e., $\sigma^2_{i,\xi} \sim \mathcal{G}(1/2, 1/2B_\sigma)$, where the hyperparameter B_σ is not very influential and set to unity.

E. Details on Structural Scenario Analysis Counterfactuals

Building on the work of Waggoner and Zha (1999), the structural scenario analysis framework of Antolin-Diaz, Petrella and Rubio-Ramirez (2021) provides a general framework on how to impose specific paths on observed variables in a VAR model as conditional forecasts with and without constraints on the set of offsetting – or *driving* – shocks. This has been adapted to the case of impulse response analysis with structural scenario analysis (SSA) (Breitenlechner, Georgiadis and Schumann, 2022; Boeck and Zörner, 2023). Again, iterate the VAR model in Equation Equation ((5.1)) forward and re-write it as

$$\boldsymbol{y}_{T+1,T+h} = \boldsymbol{b}_{T+1,T+h} + \boldsymbol{M}'\boldsymbol{\varepsilon}_{T+1,T+h}, \tag{E.1}$$

where the $nh\times 1$ vector $\boldsymbol{y}_{T+1,T+h} = (\boldsymbol{y}'_{T+1}, \boldsymbol{y}'_{T+2}, \ldots, \boldsymbol{y}'_{T+h})'$ denotes future values of the endogenous variables, $\boldsymbol{b}_{T+1,T+h}$ an autoregressive component that is due to initial conditions as of period T, and the $nh \times 1$ vector $\boldsymbol{\varepsilon}_{T+1,T+h} = (\boldsymbol{\varepsilon}'_{T+1}, \boldsymbol{\varepsilon}'_{T+2}, \ldots, \boldsymbol{\varepsilon}'_{T+h})'$ future values of the structural shocks. The $nh \times nh$ matrix $\boldsymbol{M}$ reflects the impulse responses and is a function of the structural VAR parameters. The definition of $\boldsymbol{M}$ is as follows

$$\boldsymbol{M} = \begin{bmatrix} \boldsymbol{M}_0 & \boldsymbol{M}_1 & \ldots & \boldsymbol{M}_{h-1} \\ \boldsymbol{0} & \boldsymbol{M}_0 & \ldots & \boldsymbol{M}_{h-2} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \ldots & \boldsymbol{M}_0 \end{bmatrix}, \tag{E.2}$$

where $\boldsymbol{M}_0 = \boldsymbol{S}$ and $\boldsymbol{M}_i = \sum_{j=1}^{i} \boldsymbol{M}_{i-j}\boldsymbol{B}_j$ with $\boldsymbol{B}_j = \boldsymbol{0}$ if $j > p$. From this representation it is clear that the matrix $\boldsymbol{M}$ only depends on the structural parameters. Furthermore, note that $\boldsymbol{M}'\boldsymbol{M}$ only depends on the reduced-form parameters. Thus, one only needs the history of observables and the reduced-form parameters to characterize the distribution of the unconditional forecast.

Then, the unconditional forecast is distributed

$$\boldsymbol{y}_{T+1,T+h} = \mathcal{N}\left(\boldsymbol{b}_{T+1,T+h}, \boldsymbol{M}'\boldsymbol{M}\right). \tag{E.3}$$

In the framework of Antolin-Diaz, Petrella and Rubio-Ramirez (2021), structural scenarios involve

i) *Conditional-on-observables* forecasting, i.e., specifying paths for a subset of observables in $\boldsymbol{y}_{T+1,T+h}$ that depart from their unconditional forecast, and/or

ii) *Conditional-on-shocks* forecasting, i.e., specifying the subset of structural shocks $\boldsymbol{\varepsilon}_{T+1,T+h}$ that are allowed to deviate from their unconditional distribution to produce the specified path of the observables in (i).

In the following, we will discuss how to implement both options. Therefore, one should note that

$$\tilde{\boldsymbol{y}}_{T+1,T+h} \sim \mathcal{N}\left(\boldsymbol{\mu}_y, \boldsymbol{\Sigma}_y\right), \tag{E.4}$$

denotes the distribution of the future values of the *constrained* observables. The goal is to determine $\boldsymbol{\mu}_y$ and $\boldsymbol{\Sigma}_y$ such that the constraints in (i) and (ii) are satisfied simultaneously.

Under (i), *conditional-on-observables* forecasting can be implemented as follows. Let $\overline{\boldsymbol{C}}$ be a $k_o \times nh$ selection matrix, with k_o denoting the number of restrictions. Then, *conditional-on-observables* restrictions can be written as

$$\overline{\boldsymbol{C}}\tilde{\boldsymbol{y}}_{T+1,T+h} \sim \mathcal{N}\left(\overline{\boldsymbol{f}}_{T+1,T+h}, \overline{\boldsymbol{\Omega}}_f\right), \tag{E.5}$$

where the $k_o \times 1$ vector $\overline{\boldsymbol{f}}_{T+1,T+h}$ is the mean of the distribution of the observables constrained under the conditional forecast, and the $k_o \times k_o$ matrix $\overline{\boldsymbol{\Omega}}_f$ is the associated variance-covariance matrix.

Under (ii), *conditional-on-shocks* forecasting can be implemented as follows. Let $\boldsymbol{\Xi}$ be a $k_s \times nh$ selection matrix, with k_s denoting the number of restrictions. Then, *conditional-on-shocks* restrictions can be written as

$$\Xi\tilde{\boldsymbol{\varepsilon}}_{T+1,T+h} \sim \mathcal{N}\left(\boldsymbol{g}_{T+1,T+h}, \boldsymbol{\Omega}_g\right), \tag{E.6}$$

where the $k_s \times 1$ vector $\boldsymbol{g}_{T+1,T+h}$ is the mean of the distribution of the shocks constrained under the conditional forecast and the $k_s \times k_s$ matrix $\boldsymbol{\Omega}_g$ is the associated variance-covariance matrix. Under invertability, the shocks can always be expressed as a function of observed variables and allows us to re-write the restrictions:

$$\begin{aligned} \Xi\boldsymbol{M}'^{-1}\tilde{\boldsymbol{y}}_{T+1,T+h} &= \Xi\boldsymbol{M}'^{-1}\boldsymbol{b}_{T+1,T+h} + \Xi\tilde{\boldsymbol{\varepsilon}}_{T+1,T+h} \\ \underline{\boldsymbol{C}}\tilde{\boldsymbol{y}}_{T+1,T+h} &= \underline{\boldsymbol{C}}\boldsymbol{b}_{T+1,T+h} + \Xi\tilde{\boldsymbol{\varepsilon}}_{T+1,T+h}, \end{aligned} \tag{E.7}$$

and thus

$$\underline{\boldsymbol{C}}\tilde{\boldsymbol{y}}_{T+1,T+h} = \underline{\boldsymbol{C}}\boldsymbol{b}_{T+1,T+h} + \Xi\tilde{\boldsymbol{\varepsilon}}_{T+1,T+h} \sim \mathcal{N}\left(\underline{\boldsymbol{f}}_{T+1,T+h}, \underline{\boldsymbol{\Omega}}_f\right), \tag{E.8}$$

where $\underline{\boldsymbol{\Omega}}_f = \boldsymbol{\Omega}_g$.

Now we can combine the k_o restrictions on the observables under *conditional-on-observables* forecasting and the k_s restrictions on the structural shocks under *conditional-on-shocks* forecasting. This amounts to $k = k_o + k_s$ total restrictions. We define the $k \times nh$ matrices $\boldsymbol{C} = [\overline{\boldsymbol{C}}', \underline{\boldsymbol{C}}']'$ and $\boldsymbol{D} = [\boldsymbol{M}\overline{\boldsymbol{C}}', \Xi']'$, which allows us to write

$$\boldsymbol{C}\tilde{\boldsymbol{y}}_{T+1,T+h} = \boldsymbol{C}\boldsymbol{b}_{T+1,T+h} + \boldsymbol{D}\tilde{\boldsymbol{\varepsilon}}_{T+1,T+h} \sim \mathcal{N}\left(\boldsymbol{f}_{T+1,T+h}, \boldsymbol{\Omega}_f\right), \tag{E.9}$$

where the $k \times 1$ vector $\boldsymbol{f}_{T+1,T+h} = [\overline{\boldsymbol{f}}'_{T+1,T+h}, \underline{\boldsymbol{f}}'_{T+1,T+h}]'$ stacks the means of the distribution and the $k \times k$ matrix $\boldsymbol{\Omega}_f = \mathtt{diag}(\overline{\boldsymbol{\Omega}}_f, \underline{\boldsymbol{\Omega}}_f)$ denotes the associated variance-covariance matrix.

Following the framework in Antolin-Diaz, Petrella and Rubio-Ramirez (2021) and given the restrictions specified above, we can derive solutions for $\boldsymbol{\mu}_y$ and $\boldsymbol{\Sigma}_y$. Define the restricted future shocks

$$\tilde{\boldsymbol{\varepsilon}}_{T+1,T+h} \sim \mathcal{N}(\boldsymbol{\mu}_\varepsilon, \boldsymbol{\Sigma}_\varepsilon), \tag{E.10}$$

where $\boldsymbol{\Sigma}_\varepsilon = \boldsymbol{I}_n h + \boldsymbol{\Psi}_\varepsilon$, such that $\boldsymbol{\mu}_\varepsilon$ and $\boldsymbol{\Psi}_\varepsilon$ denote the deviation of the mean and covariance matrix from their unconditional counterparts. Using Equation (E.9), we match the first and second moment to get

$$\boldsymbol{f}_{T+1,T+h} = \boldsymbol{C}\boldsymbol{b}_{T+1,T+h} + \boldsymbol{D}\boldsymbol{\mu}_\varepsilon, \tag{E.11}$$

$$\boldsymbol{\Omega}_f = \boldsymbol{D}(\boldsymbol{I}_n h + \boldsymbol{\Psi}_\varepsilon)\boldsymbol{D}'. \tag{E.12}$$

Depending on k, the number of restrictions, and nh, the length of $\tilde{\boldsymbol{y}}_{T+1,T+h}$, the systems of Equation (E.11) and Equation (E.12) may have multiple solutions ($k < nh$), one solution ($k = nh$), or no solution ($k > nh$). Since $k < nh$ is the most interesting case, the solution are given by

$$\boldsymbol{\mu}_\varepsilon = \boldsymbol{D}^*\left(\boldsymbol{f}_{T+1,T+h} - \boldsymbol{C}\boldsymbol{b}_{T+1,T+h}\right), \tag{E.13}$$

$$\boldsymbol{\Psi}_\varepsilon = \boldsymbol{D}^*\boldsymbol{\Omega}_f\boldsymbol{D}^{*'} - \boldsymbol{D}^*\boldsymbol{D}\boldsymbol{D}'\boldsymbol{D}^{*'}, \tag{E.14}$$

where $\boldsymbol{D}^*$ is the Moore-Penrose inverse of $\boldsymbol{D}$. Equation (E.13) shows that the path of the implied structural shocks under the conditional forecast depend on its deviation from the unconditional forecast. Furthermore, Equation (E.14) shows that the variance of the implied future structural shocks depends on the uncertainty the researcher attaches to the conditional forecast. If the uncertainty is zero ($\boldsymbol{\Omega}_f = \boldsymbol{0}$), then $\boldsymbol{\Sigma}_\varepsilon = \boldsymbol{0}$. This means that a unique path for $\boldsymbol{\mu}_\varepsilon$ can be found.

Combining Equation (E.3), Equation (E.13), and Equation (E.14), we get

$$\boldsymbol{\mu}_y = \boldsymbol{b}_{T+1,T+h} + \boldsymbol{M}'\boldsymbol{D}^*\left(\boldsymbol{f}_{T+1,T+h} - \boldsymbol{C}\boldsymbol{b}_{T+1,T+h}\right), \tag{E.15}$$

$$\boldsymbol{\Sigma}_y = \boldsymbol{M}'\boldsymbol{M} - \boldsymbol{M}'\boldsymbol{D}^*\left(\boldsymbol{\Omega}_f - \boldsymbol{D}\boldsymbol{D}'\right)\boldsymbol{D}^{*'}\boldsymbol{M}. \tag{E.16}$$

As before, if $\boldsymbol{\Omega}_f = \boldsymbol{0}$, then $\boldsymbol{\Sigma}_y = \boldsymbol{0}$ and thus there is no uncertainty about the path of the observables under the imposed restrictions.

E1. Restrictions in the VAR

The VAR in section 5.3 features $n = 9$ variables with belief distortions ordered last and the EBP ordered after real GDP, real consumption, real investment, and prices. The remaining variables in the VAR after the EBP are the respective credit spread, the S&P 500, the federal funds rate, and the belief distortion series. I constrain the effect of the financial shock on belief distortions to be zero. Denote with $\boldsymbol{e}_i$ a $n \times 1$ vector of zeros with unity at the i-th position.

Under (i), *conditional-on-observable* forecasting, we impose

$$\overline{\boldsymbol{C}} = \boldsymbol{I}_h \otimes \boldsymbol{e}_9', \tag{E.17}$$

$$\overline{\boldsymbol{f}}_{T+1,T+h} = \boldsymbol{0}_{h\times 1}, \tag{E.18}$$

$$\overline{\boldsymbol{\Omega}}_f = \boldsymbol{0}_{h\times h}. \tag{E.19}$$

These equations impose that the conditional forecast that underlies the impulse response of belief distortions (which is ordered last in the VAR) is constrained to be zero over all horizons $T+1, \ldots, T+h$. Furthermore, we do not allow for any uncertainty.

Under (ii), *conditional-on-shocks* forecasting, we impose

$$\boldsymbol{\Xi}_{h\times nh} = \boldsymbol{I}_h \otimes \boldsymbol{e}_5' \tag{E.20}$$

$$\underline{\boldsymbol{f}}_{T+1,T+h} = \boldsymbol{g}_{T+1,T+h} = [1, \boldsymbol{0}_{1\times h-1}]' \tag{E.21}$$

$$\underline{\boldsymbol{\Omega}}_f = \boldsymbol{\Omega}_g = \boldsymbol{0}_{h\times h} \tag{E.22}$$

Equation (E.20) selects the financial shock via the EBP ordered fifth in $\boldsymbol{\varepsilon}_t$ for the entire impulse response horizon. We constrain that the financial shock is equal to unity in the first period and zero afterwards, as indicated by Equation (E.21). The first element constrains it to be unity on impact in period $T+1$ and zero for $T+2, T+3, \ldots, T+h$. Hence, in $T+1, T+3, \ldots, T+h$ all structural shocks except the financial shock are allowed to vary. Lastly, Equation (E.22) specifies that we allow for no uncertainty. It is also interesting to consider the stacked matrices $\boldsymbol{C}$ and $\boldsymbol{D}$ which look as follows

$$\boldsymbol{C} = \begin{pmatrix} \overline{\boldsymbol{C}}_{h\times nh} \\ \underline{\boldsymbol{C}}_{h(n-1)\times nh} \end{pmatrix}_{hn\times nh}, \qquad \boldsymbol{D} = \begin{pmatrix} \overline{\boldsymbol{C}}_{h\times nh}\boldsymbol{M}'_{nh\times nh} \\ \boldsymbol{\Xi}_{h(n-1)\times nh} \end{pmatrix}_{hn\times nh}, \tag{E.23}$$

where $\underline{\boldsymbol{C}} = \boldsymbol{\Xi}\boldsymbol{M}'^{-1}$.

E2. How Plausible is the Counterfactual?

Generally, structural scenario analysis counterfactuals based on SVARs are not prone to the Lucas critique (Lucas, 1976). However, if the implied shocks are so *unusual* the analysis might becomes subject to the Lucas critique anyway. Hence, measures of plausibility of the created counterfactual scenario are a remedy. We use two measures: the q-divergence proposed in Antolin-Diaz, Petrella and Rubio-Ramirez (2021) and adapted to the case of impulse response functions by Breitenlechner, Georgiadis and Schumann (2022) and the modesty statistic proposed by Leeper and Zha (2003). These measures intend to measure by how much the structural scenario deviates from its unconditional counterpart. When this deviation becomes too large, the scenario might be implausible.

Antolin-Diaz, Petrella and Rubio-Ramirez (2021) propose to use the Kullback-Leibler (KL) divergence as a measure how plausible a scenario is. Denote with $\mathcal{D}(\mathcal{N}_{SS}||\mathcal{N}_{UF})$ the KL divergence between the distributions of the structural scenario analysis $\mathcal{N}_{SS}$ and the unconditional distribution $\mathcal{N}_{UF}$. While it is straightforward to compute $\mathcal{D}(\mathcal{N}_{SS}||\mathcal{N}_{UF})$, it is difficult to grasp whether any value for the KL divergence is large or small. In other words, the KL divergence can be easily used to rank scenarios, but it is hard to understand how far away they are from the unconditional forecast. Therefore, Antolin-Diaz, Petrella and Rubio-Ramirez (2021) propose to compare the KL divergence with the divergence between two binomial distributions, one with probability q and the other with probability $p = 0.5$. The idea is to compare the implied counterfactual distribution with their unconditional distribution, which translates into a comparison of the binomial distributions of a fair and a biased coin. If the probability q is near to to p, then this suggests that the distribution of the offsetting shocks is not at all far from the unconditional distribution. Antolin-Diaz, Petrella and Rubio-Ramirez (2021) suggest calibrating the KL divergence from $\mathcal{N}_{UF}$ to $\mathcal{N}_{SS}$ to a parameter q that would solve the following equation $\mathcal{D}\left(\mathcal{B}(nh, 0.5)||\mathcal{B}(nh, q)\right) = \mathcal{D}\left(\mathcal{N}_{SS}||\mathcal{N}_{UF}\right)$. The solution to the equation is

$$q = 0.5 * \left(1 + \sqrt{1 - \exp\left(-\frac{2z}{nh}\right)}\right) \quad \text{with} \quad z = \mathcal{D}\left(\mathcal{N}_{SS}||\mathcal{N}_{UF}\right). \tag{E.24}$$

As Breitenlechner, Georgiadis and Schumann (2022) point out, in the context of impulse responses the KL divergence has to be slightly adjusted, because Antolin-Diaz, Petrella and Rubio-Ramirez (2021) propose their measure in the context of conditional forecasts relative to an unconditional forecast. As before, the unconditional scenario is the case with only a single shock of unity size, which occurs in $T + 1$ with certainty. More formally, $\boldsymbol{\varepsilon}_{T+1,T+h} = (\boldsymbol{e}_5', \boldsymbol{0}_{n(h-1)\times 1})'$ denotes the *unconditional* impulse response of a financial shock. $\boldsymbol{e}_i$ denotes the unit vector with unity on the i-th position. For the structural scenario analysis counterfactual, we impose the restrictions specified above (i.e., belief distortions do not react to a financial shock). Hence, we set

$$\texttt{UF:} \qquad \boldsymbol{\mu}_{UF} = \boldsymbol{M}'(\boldsymbol{e}_5', \boldsymbol{0}_{n(h-1)\times 1})' \tag{E.25}$$

$$\texttt{SS:} \qquad \boldsymbol{\mu}_{SS} = \mu_y, \tag{E.26}$$

where $\boldsymbol{\mu}_y$ is given by Equation ((E.15)). Since we impose this with certainty, $\boldsymbol{\Psi} = \boldsymbol{0}$ such that the shocks have their unconditional variance. Hence, $\boldsymbol{\Sigma}_{UF} = \boldsymbol{\Sigma}_{SS} = \boldsymbol{\Sigma}_{\varepsilon} = \boldsymbol{I}$. The KL divergence between the distribution of the shocks under the unconditional and conditional scenario is then given by

$$\mathcal{D}(\mathcal{N}_{SS}||\mathcal{N}_{UF}) = \frac{1}{2}\left(\operatorname{tr}\left(\boldsymbol{\Sigma}_{SS}^{-1}\boldsymbol{\Sigma}_{UF}\right) + (\boldsymbol{\mu}_{SS} - \boldsymbol{\mu}_{UF})'\boldsymbol{\Sigma}_{SS}^{-1}(\boldsymbol{\mu}_{SS} - \boldsymbol{\mu}_{UF}) - nh + \ln\left(\frac{\det \boldsymbol{\Sigma}_{SS}}{\det \boldsymbol{\Sigma}_{UF}}\right)\right), \tag{E.27}$$

where $\boldsymbol{\mu}_{\varepsilon}$ and $\boldsymbol{\Sigma}_{\varepsilon}$ are given by Equation ((E.13)) and Equation ((E.14)). Furthermore, we discard any SSA counterfactuals when the offsetting shocks are particularly unlikely. We set this to be above $q > 0.9$.

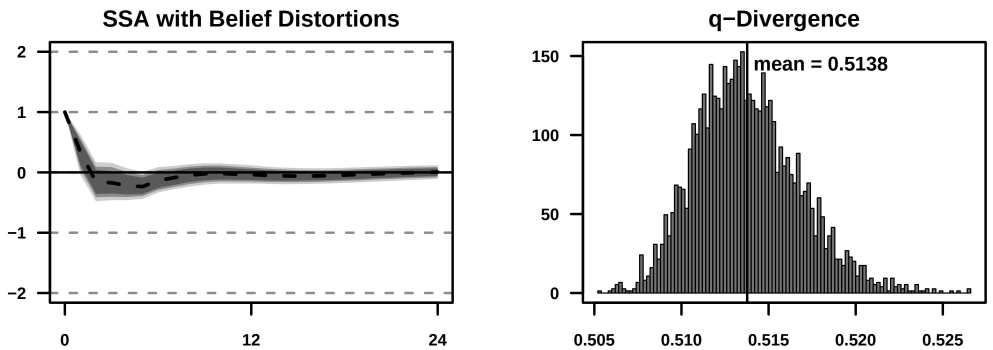
The second plausibility measure is the one of *modest intervention* or *modesty statistic* used in Leeper and Zha (2003). The measure reports how unusual the path for policy shocks is relative to the typical size of these shocks, which are needed to impose the counterfactual restriction. For instance, if the counterfactual implies a sequence of shocks close to their unconditional mean, the policy intervention is considered *modest*, in the sense that the shocks are unlikely to induce agents to revise their beliefs about policy rules and the structure of the economy. Instead, if the counterfactual involves an unlikely sequence of shocks the analysis is likely to be prone to the critique by Lucas (1976). The offsetting shocks are considered to be modest if the statistic is smaller than two in absolute value.

E3. Additional Results

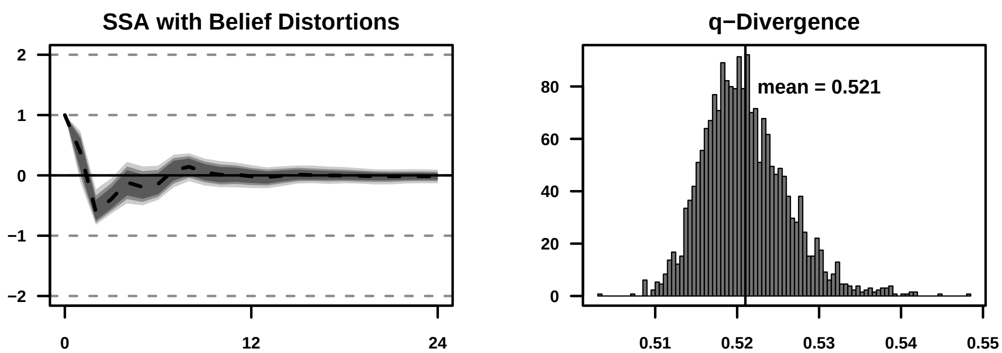
The plausibility of the counterfactuals obtained by the structural scenario analysis depends on the offsetting structural shocks. I show here the modesty statistic of Leeper and Zha (2003) and the q-divergence proposed by Antolin-Diaz, Petrella and Rubio-Ramirez (2021). Both are presented in Figure E1. The top panel shows the modesty statistic, which are the implied offsetting shocks that impose the counterfactual constraint for belief distortions. The offsetting shocks are *modest* if the statistic is smaller than two in absolute values. This is confirmed and thus the materialisation is unlikely to induce agents to adjust their expectation formation and beliefs about the structure of the economy showing no sign for the Lucas critique. In the lower panel, the q-divergence indicates how strongly the distribution of offsetting shocks in the counterfactual deviate

Figure E1: Plausbility Statistics of Counterfactuals.

(**a**) Model with Aaa Belief Distortion.



(**b**) Model with Baa Belief Distortion.



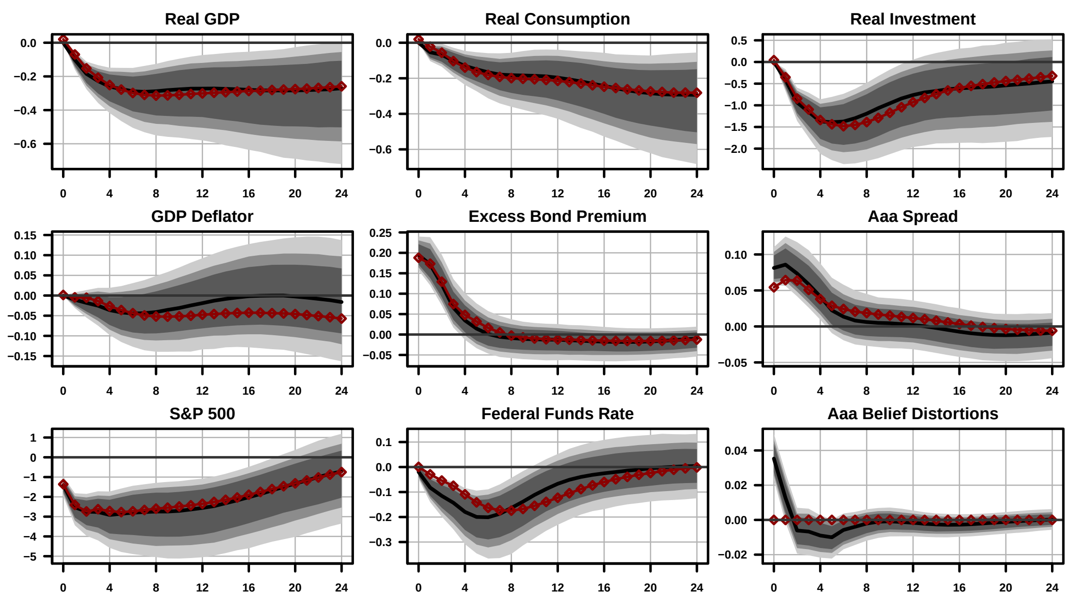
Notes: The upper panel shows the modesty statistic of Leeper and Zha (2003) and the lower panel shows the distribution of the q-divergence proposed by Antolin-Diaz, Petrella and Rubio-Ramirez (2021). The modesty statistic reports the implied shocks that impose the counterfactual constraint for belief distortions. The black dashed line denotes the posterior median responses while gray shaded areas depict the 68/80/90 percent confidence intervals.

from their unconditional distribution translated into a comparison of the binomial distribution of a fair and a biased coin. Again, the test does not indicate that the distribution of offsetting shocks in the counterfactual is notably different from the unconditional distribution.

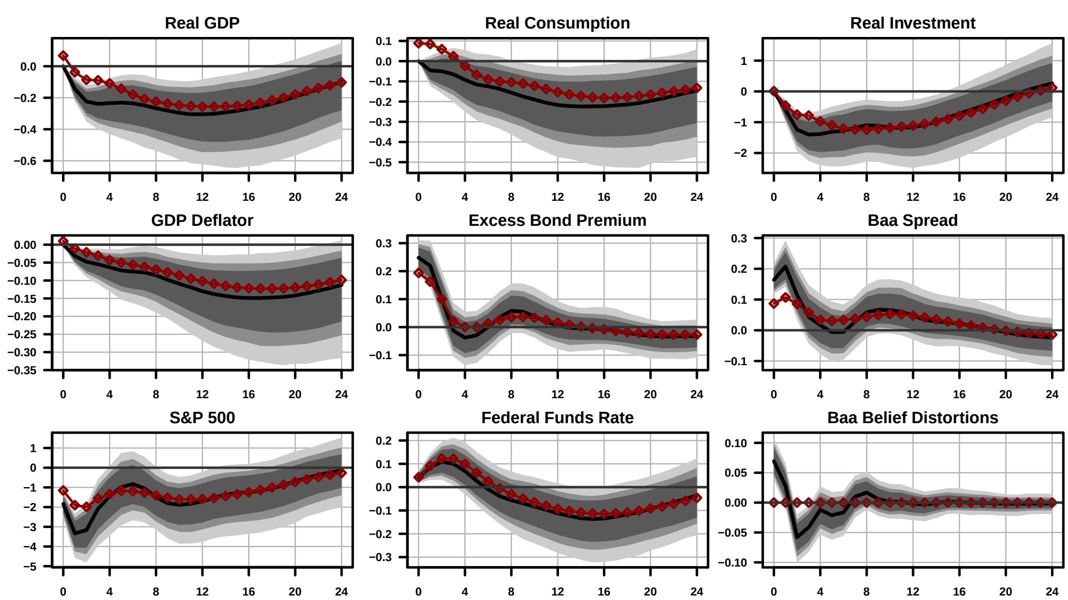
Lastly, in Figure E2 I report the impulse responses of all variables in the model including its counterfactual response when shutting down the transmission channel of the financial shock via belief distortions. This complements Figure 5, in which I only show a subset of the responses for brevity.

Figure E2: Counterfactual Impulse Response Functions to a Financial Shock (Extended Baseline).

(a) Model with Aaa Belief Distortion.



(b) Model with Baa Belief Distortion.



Notes: Impulse response functions of the extended baseline VAR. Black line denotes median response while gray shaded areas denote the 68/80/90 percent confidence intervals. Dark red lines indicate counterfactual impulse response function. The responses of real GDP, real consumption, real investment, GDP deflator, and the stock market index are scaled in percent, while the excess bond premium, the credit spread, the federal funds rate, and belief distortions are scaled in percentage points.

F. Diagnostics of the Belief Distortion Series

As discussed in the paper, I perform several validity checks on the belief distortion series. First, I investigate whether autocorrelation is present in the distortion series. Figure F1 depicts the autocorrelation function. I also perform Granger causality tests, depicted in Table F1. Last, I compute correlations to other structural shocks from the literature, which I show in Figure F2 and Table F2. In particular, I compare the belief distortions in the Aaa and Baa credit spread to high-frequency monetary policy shocks in Jarociński and Karadi (2020) (labelled *HFI Monetary Policy*), the narrative fiscal policy shocks by Romer and Romer (2010) (labelled *RR Fiscal Policy*), the uncertainty indicators based on Jurado, Ludvigson and Ng (2015) (labelled *Financial Uncertainty*, *Macro Uncertainty* and *Real Uncertainty* as well as first differences thereof, indicated with *(diff)*), the economic policy uncertainty indicator by Baker, Bloom and Davis (2016) (labelled *Economic Policy Uncertainty* as well as first differences thereof, indicated with *(diff)*), the extended high-frequency monetary policy instrument by Miranda-Agrippino and Ricco (2021) (labelled *HFI Monetary Policy Ext1* and *HFI Monetary Policy Ext2*), the extended monetary policy measured constructed by Romer and Romer (2004) and extended by Breitenlechner (2018) (labelled *RR Monetary Policy 1* and *RR Monetary Policy 2*), high-frequency oil supply and supply news shocks by Känzig (2021) (labelled *HFI Oil Supply* and *HFI Oil News*), and the structural oil supply and demand as well as the aggregate demand shock by Kilian (2009) (labelled *Oil Supply*, *Aggregate Demand (Oil)*, and *Oil Demand*). Additionally, I also compare the belief distortion series to the excess bond premium (Gilchrist and Zakrajšek, 2012), its first difference, and the residuals of the VAR specified to construct financial shocks in Gilchrist and Zakrajšek (2012). Lastly, I compare the measure of belief distortions to the financial shocks constructed in Caldara et al. (2016), indicated by `CFAGZ1` (σ-EBP identification) and `CFAGZ2` (EBP-σ identification). Correlations are depicted in Figure F2 and in Table F2.

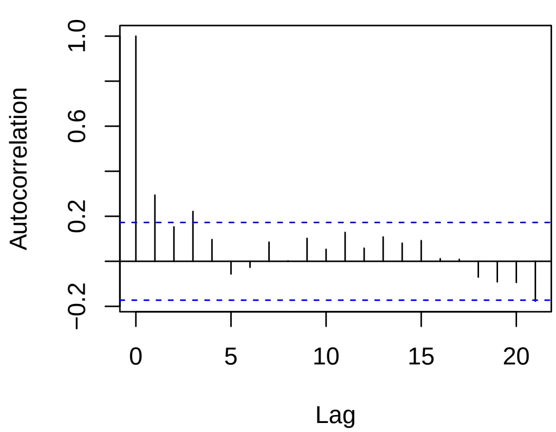
The diagnostics of the belief distortion series reveal the following. For both series, there is no evidence that the series is serially autocorrelated. Granger causality tests for different lag lengths shows that there is no predictive causality running from any of those variables to the constructed series of belief distortions. Regarding the correlation structure to other indicators from the literature, an interesting picture emerges. When compared to other identified, plausible exogenous macroeconomic shocks (monetary, fiscal, and oil), correlations are low ($\rho < 0.20$) and statistically insignificant. However, turning to the financial shocks correlations are higher and mostly statistically significant. The same holds true when inspecting correlations to uncertainty indicators. All these correlations are negative. For the Aaa belief distortion series correlations are safely below $\rho < 0.40$ and for the Baa belief distortion series go up to around $\rho \approx 0.60$. Furthermore, those correlations show statistical significance. In particular, statistical significance is high for the uncertainty indicators in first differences since they are in their original form quite persistent time series. Furthermore, it is interesting that only the financial shock from Caldara et al. (2016) with the σ-EBP identification is statistically significant while the other is not.

Table F1: Granger Causality Tests.

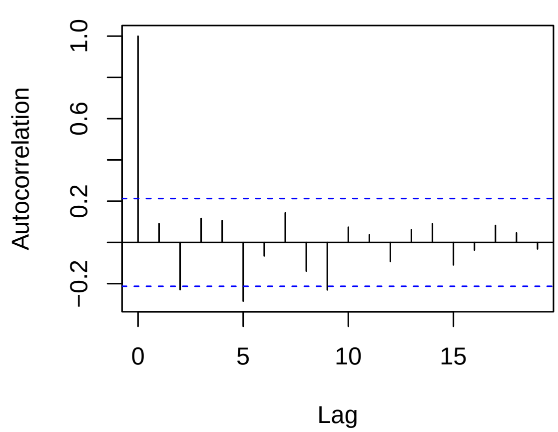
Variable	Aaa	Baa	Aaa	Baa	Aaa	Baa	Aaa	Baa
	$p = 1$		$p = 2$		$p = 3$		$p = 4$	
FEDFUNDS	0.56	0.89	0.75	0.66	0.88	0.75	0.55	0.78
GS1	0.51	0.88	0.70	0.96	0.80	0.59	0.81	0.27
GS5	0.44	0.55	0.46	0.52	0.61	0.69	0.62	0.25
GS10	0.52	0.53	0.37	0.40	0.54	0.57	0.56	0.39
GDPC1	0.15	0.09	0.23	0.18	0.36	0.34	0.40	0.38
UNRATE	0.51	0.28	0.50	0.47	0.71	0.71	0.58	0.44
BUSLOANS	0.67	0.98	0.54	0.26	0.55	0.39	0.65	0.48
SP500	0.21	0.77	0.04	0.13	0.03	0.36	0.05	0.25
NASDAQCOM	0.37	0.96	0.11	0.58	0.12	0.91	0.10	0.35
GDPDEF	0.43	0.52	0.55	0.72	0.66	0.87	0.83	0.86
CPIAUCSL	0.27	0.63	0.25	0.55	0.36	0.60	0.51	0.27

Notes: Table shows p-values of a series of Granger causality tests of the respective belief distortion series using a selection of macroeconomic and financial variables. Series are transformed to stationarity according to transformations provided in Table A1. The lag order is given above and in terms of deterministics, only a constant term is included.

Figure F1: Autocorrelation Function of Belief Distortion Series.



(a) Aaa Belief Distortions.



(b) Baa Belief Distortions.

Figure F2: Correalation to Other Structural Shocks.

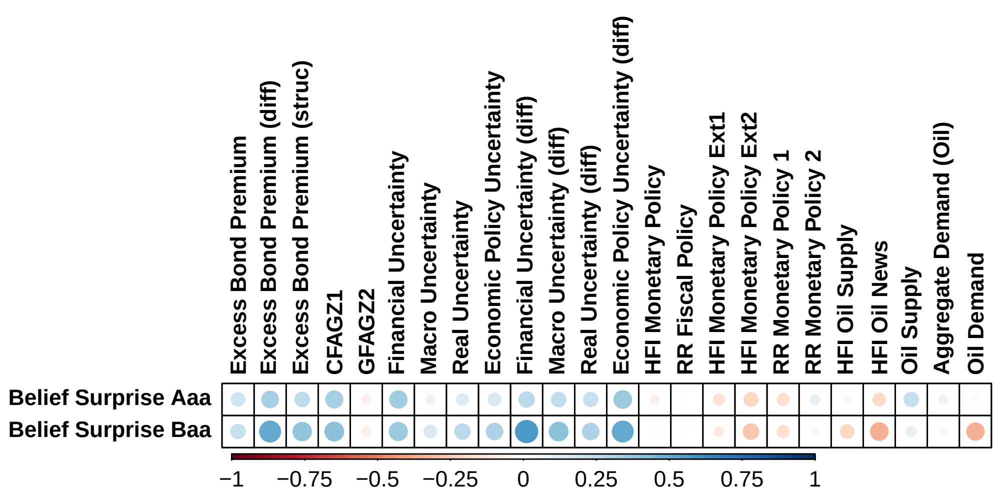


Table F2: Correlations to Other Structural Shocks.

Variable	Aaa Belief Distortion		Baa Belief Distortion	
	Est.	p-val	Est.	p-val
Financial Shocks				
`Excess Bond Premium`	0.12	0.20	0.24	0.05
`Excess Bond Premium (diff)`	0.33	0.00	0.49	0.00
`Excess Bond Premium (struc)`	0.22	0.02	0.34	0.00
`GFAGZ-1`	0.38	0.00	0.41	0.00
`GFAGZ-2`	−0.07	0.44	−0.03	0.82
Uncertainty Indicators				
`Financial Uncertainty`	0.25	0.00	0.37	0.00
`Macro Uncertainty`	0.06	0.46	0.18	0.09
`Real Uncertainty`	0.14	0.11	0.29	0.01
`Economic Policy Uncertainty`	0.15	0.10	0.28	0.01
`Financial Uncertainty (diff)`	0.32	0.00	0.57	0.00
`Macro Uncertainty (diff)`	0.24	0.01	0.43	0.00
`Real Uncertainty (diff)`	0.25	0.00	0.37	0.00
`Economic Policy Uncertainty (diff)`	0.37	0.00	0.43	0.00
Macroeconomic Shocks				
`HFI Monetary Policy`	0.03	0.79	0.08	0.59
`RR Fiscal Policy`	0.07	0.55	0.07	0.64
`HFI Monetary Policy Ext1`	−0.09	0.41	−0.12	0.48
`HFI Monetary Policy Ext2`	−0.18	0.07	−0.21	0.13
`RR Monetary Policy 1`	−0.15	0.13	−0.14	0.25
`RR Monetary Policy 2`	0.10	0.36	0.12	0.48
`HFI Oil Supply`	−0.01	0.94	−0.16	0.16
`HFI Oil News`	−0.19	0.03	−0.28	0.01
`Oil Supply`	0.18	0.14	0.09	0.64
`Aggregate Demand (Oil)`	0.12	0.33	−0.12	0.56
`Oil Demand`	−0.01	0.92	−0.30	0.13

Notes: Table shows estimated Pearson's moment correlations coefficients (Est.) and according p-values (p-val) of a test against zero correlation.

G. Convergence Diagnostics

In this section, I evaluate convergence of the model presented in section 4. To proceed, I look at three different convergence diagnostics. In an ideal setting, the sampler returns independent draws. The stronger the autocorrelation in the sampler, the more draws are needed. To evaluate the extent of autocorrelation in the MCMC chain, I use three different statistics. First, I compute inefficiency factors indicating how many draws are needed for drawing one identically and independently distributed draw. Second, I have a look at the Raftery and Lewis's diagnostic statistic (Raftery and Lewis, 1992). It is also a measure of autocorrelation and returns a dependence factor which should not exceed 5 in the ideal setting. Third, I examine Geweke's convergence diagnostic (Geweke et al., 1991). This is a test of equality of the means of the first 10% and last 50% of the MCMC chain. Here, I report the share of Z-scores exceeding the critical value of 1.96.

For all models, convergence is safely achieved. While inefficiency factors are around 2-3, the dependence factors are even lower and do not exceed 2. Furthermore, when looking at the share of Z-scores exceeding the critical value of 1.96, it does not seem to be an issue. In the last column of Table G1, I report the percentage of retained draws of stationary draws. This percentage share fluctuates more, but the sampler always retain at least 20% of all draws for posterior analysis.

Table G1: Convergence Statistics.

Model	Inefficiency Factor	Dependence factor	Geweke's Z-scores	% draws retained
Baseline Model with Belief Distortions				
– Aaa	2.73	1.53	0.05	78.50
– Baa	2.93	1.67	0.03	59.50
Extended Baseline Model with Belief Distortions				
– Aaa	2.79	1.55	0.06	74.66
– Baa	3.00	1.60	0.06	52.60
Model with Survey Forecast Errors				
– Aaa	2.76	1.57	0.08	87.16
– Baa	3.11	1.75	0.04	62.12
Model with Machine Learning Forecast Errors				
– Aaa	2.57	1.48	0.07	80.06
– Baa	3.52	1.77	0.04	59.62

H. Additional Results: Forecasting

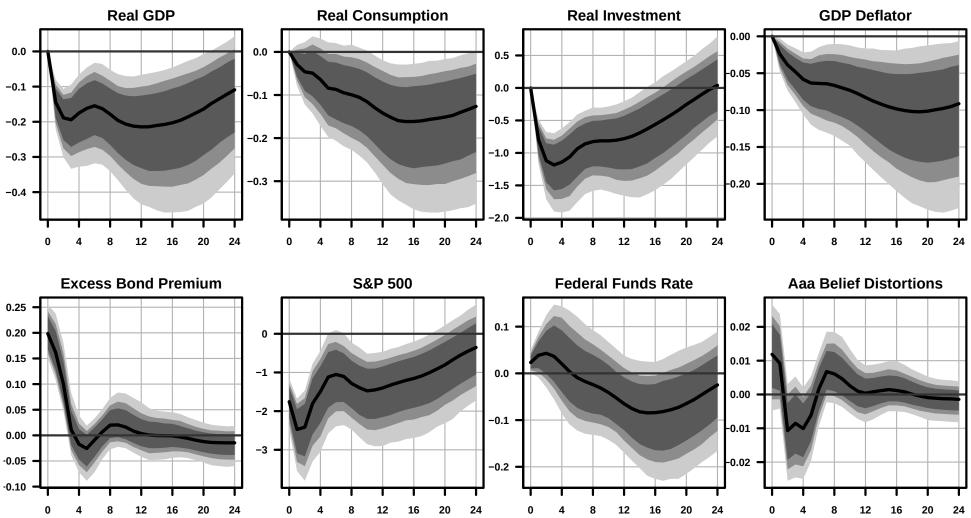
Table H1: Forecasting Evaluation.

	h=1		**h=2**		**h=3**		**h=4**	
	Aaa spread	Baa spread	Aaa spread	Baa spread	Aaa spread	Baa spread	Aaa spread	Baa spread
`RW`	**0.192**	0.391	0.288	0.617	0.342	0.764	0.391	0.866
		Autoregressive Models						
`AR(1)`	0.193	0.397	**0.287**	0.610	0.339	0.728	0.377	0.783
`AR(2)`	0.193	0.389	0.292	0.616	0.343	0.800	0.380	0.843
`AR(1)-SV`	0.195	0.391	0.298	0.615	0.361	0.755	0.417	0.851
`AR(2)-SV`	0.194	**0.370**	0.300	0.610	0.359	0.748	0.417	0.858
		Autoregressive Distributed Lag Models						
`ADL(1)`	0.196	0.551	0.308	0.728	0.379	0.885	0.421	0.863
`ADL(2)`	0.195	0.527	0.303	0.682	0.361	0.835	0.398	0.785
`ADL(1)-SV`	0.199	0.374	0.318	0.620	0.351	0.803	0.461	0.866
`ADL(2)-SV`	0.201	0.372	0.301	0.612	0.396	0.796	0.473	0.866
		Machine Learning Approaches						
`EN(1)`	0.211	0.607	0.315	0.810	0.424	0.976	0.497	0.872
`EN(2)`	0.219	0.535	0.321	0.683	0.353	0.859	0.397	0.792
`BART(1)`	0.218	0.433	0.297	**0.605**	**0.322**	0.645	0.344	**0.696**
`BART(2)`	0.219	0.453	0.293	0.606	0.323	**0.632**	**0.343**	0.709
		Subjective and Objective Expectations						
$\text{RMSE}_{\mathbb{F}}$	0.230	0.449	0.296	0.615	0.343	0.720	0.386	0.788
$\text{RMSE}_{\mathbb{E}}$	0.193	0.364	0.278	0.535	0.309	0.772	0.369	0.852
min[RMSE]	0.192	0.370	0.287	0.605	0.322	0.632	0.343	0.696
$\text{RMSE}_{\mathbb{E}}/\text{RMSE}_{\mathbb{F}}$	0.837	0.810	0.941	0.870	0.901	1.073	0.957	1.081
$\text{RMSE}_{\mathbb{E}}/\min\,[\text{RMSE}]$	1.005	0.984	0.968	0.885	0.959	1.222	1.076	1.225

Notes: True out-of-sample performance in terms of RMSEs. The bold figures indicate the best performing model for a given variable and time horizon. The following models nested in Equation ((4.2)) are considered: `RW` - random walk, `AR` - autoregressive model, `ADL` - autoregressive distributed lag model, `EN` - Elastic Net, `BART` - Bayesian Additive Regression Trees. The (first) number in the parentheses indicates the lag length. `SV` refers to stochastic volatility.

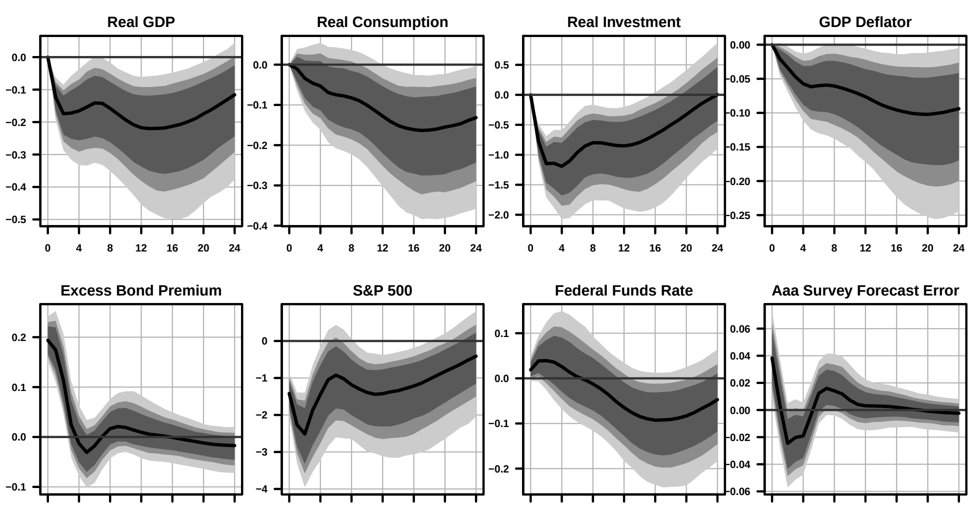
I. Additional Results: Macroeconomic Model

Figure I1: Impulse Response Functions to a Financial Shock (Belief Distortions, Short Sample).



Notes: Impulse response functions of the baseline VAR with the short sample starting in 1999Q1. Black line denotes median response while gray shaded areas denote the 68/80/90 percent confidence intervals. The responses of real GDP, real consumption, real investment, GDP deflator, and the stock market index are scaled in percent, while the excess bond premium, the federal funds rate, and belief distortions are scaled in percentage points.

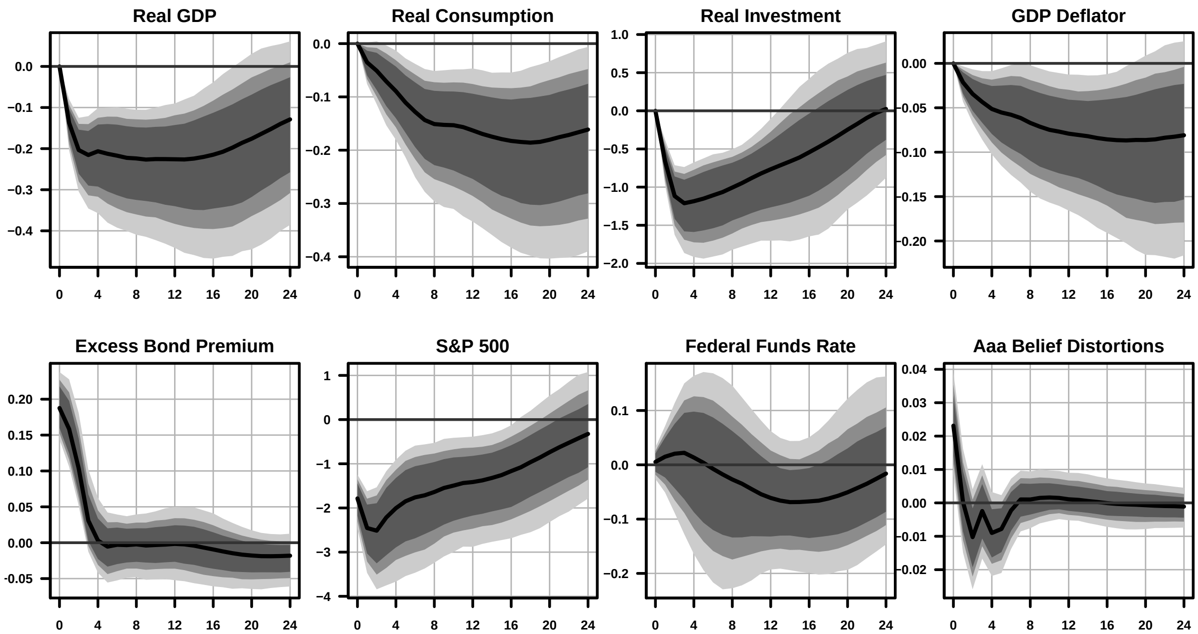
Figure I2: Impulse Response Functions to a Financial Shock (Survey Forecast Errors, Short Sample).



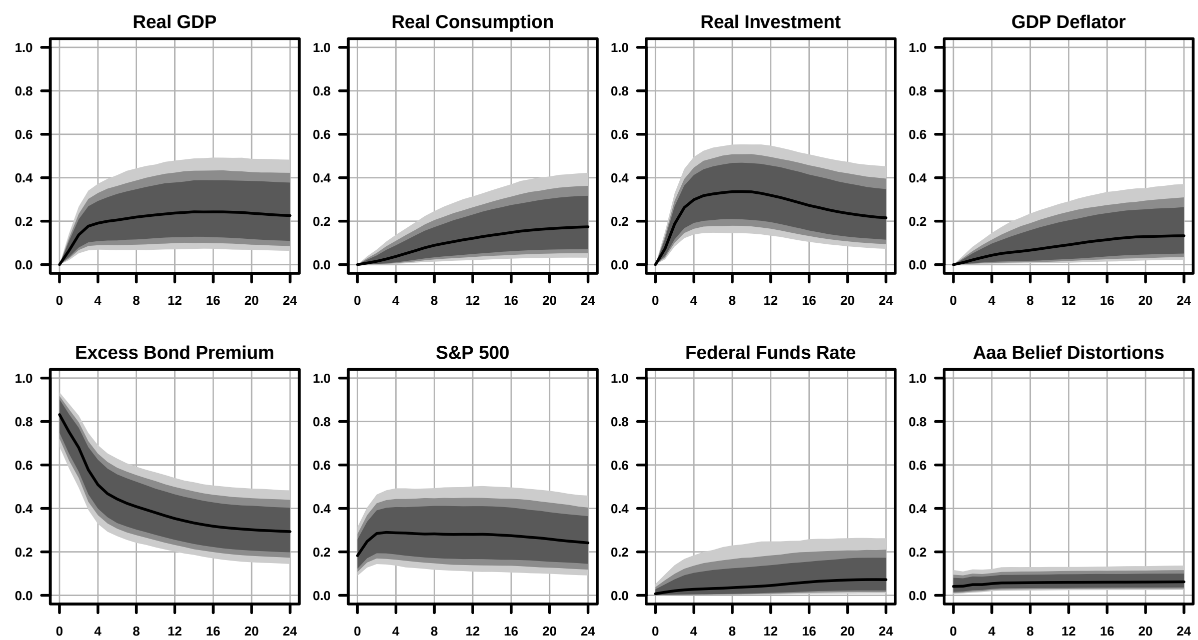
Notes: Impulse response functions of the baseline VAR with the short sample starting in 1999Q1. Black line denotes median response while gray shaded areas denote the 68/80/90 percent confidence intervals. The responses of real GDP, real consumption, real investment, GDP deflator, and the stock market index are scaled in percent, while the excess bond premium, the federal funds rate, and belief distortions are scaled in percentage points.

Figure I3: Alternative Specification with Aaa Belief Distortions.

(a) Impulse Response Functions.



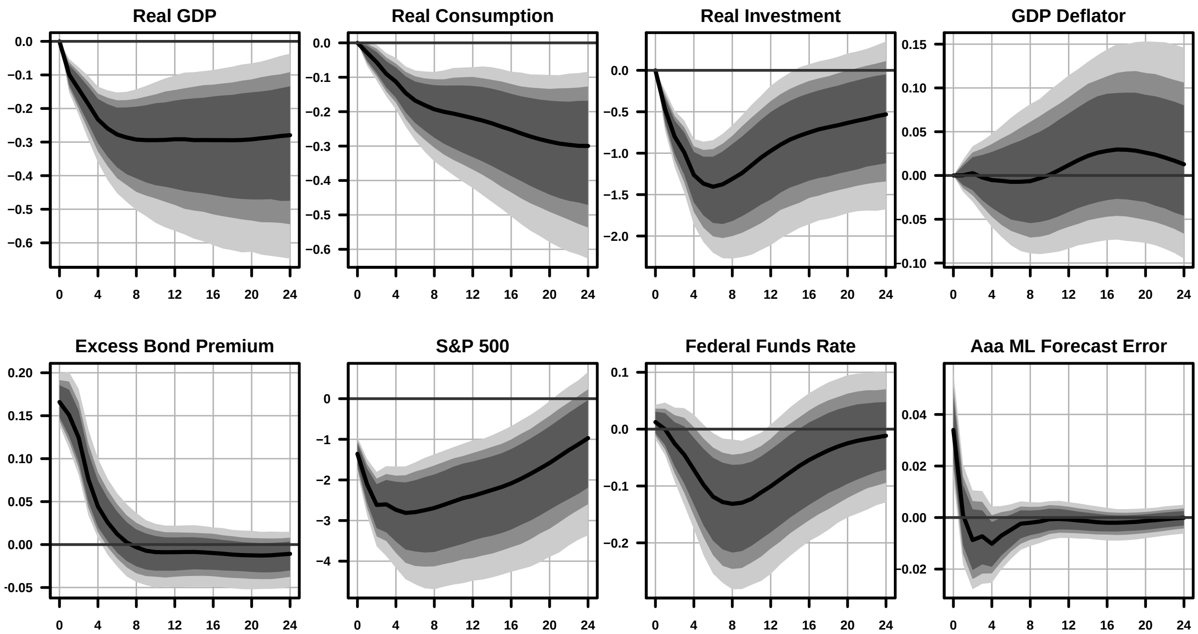
(b) Forecast Error Variance Decompositions.



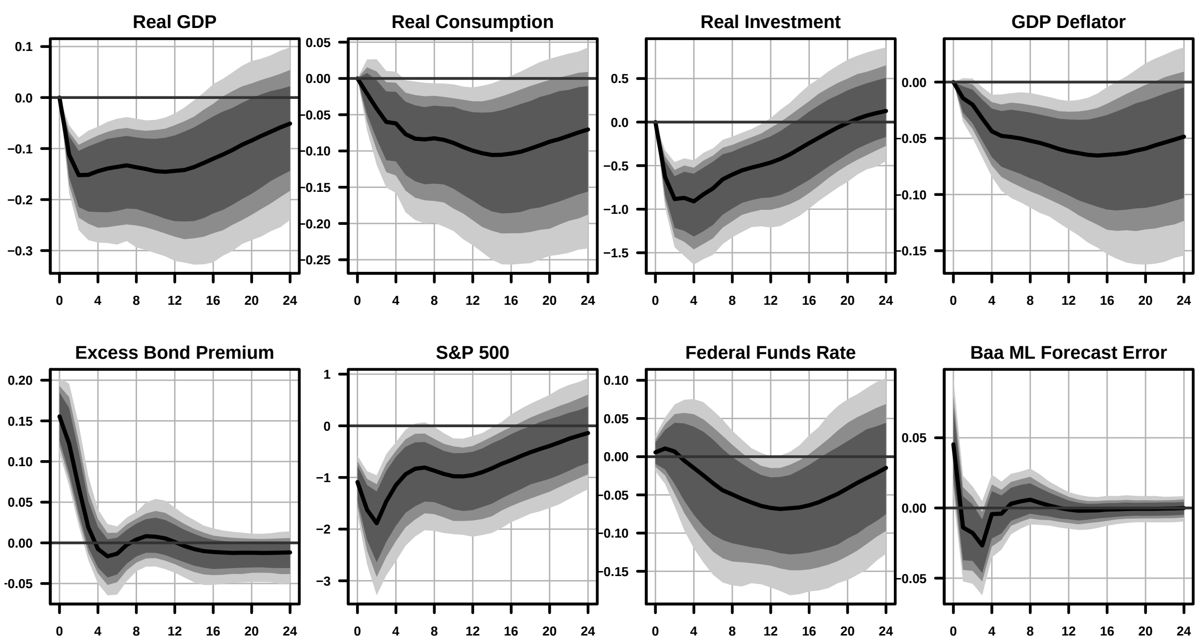
Notes: Baseline VAR with belief distortions of the alternative specification. Sample ranges from 1995Q2 to 2020Q1. Black line denotes median response while gray shaded areas denote the 68/80/90 percent confidence intervals. The impulse responses of real GDP, real consumption, real investment, GDP deflator, and the stock market index are scaled in percent, while the excess bond premium, the federal funds rate, and belief distortions are scaled in percentage points.

Figure I4: Impulse Response Functions to a Financial Shock (ML Forecast Errors).

(a) Model with Aaa Belief Distortion.



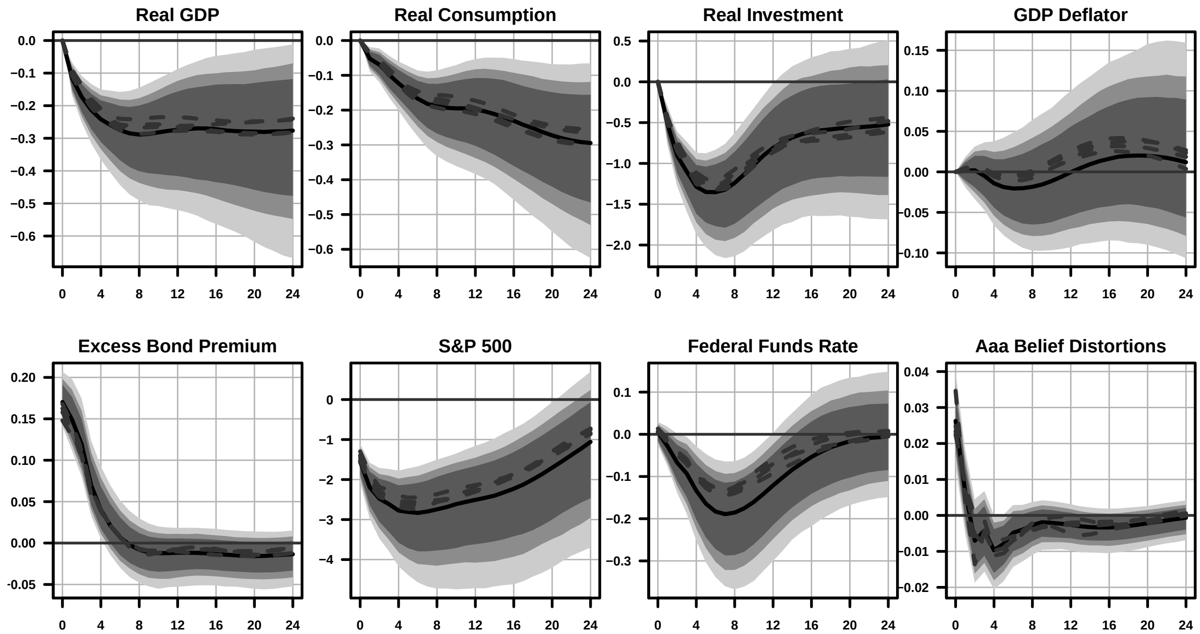
(b) Model with Baa Belief Distortion.



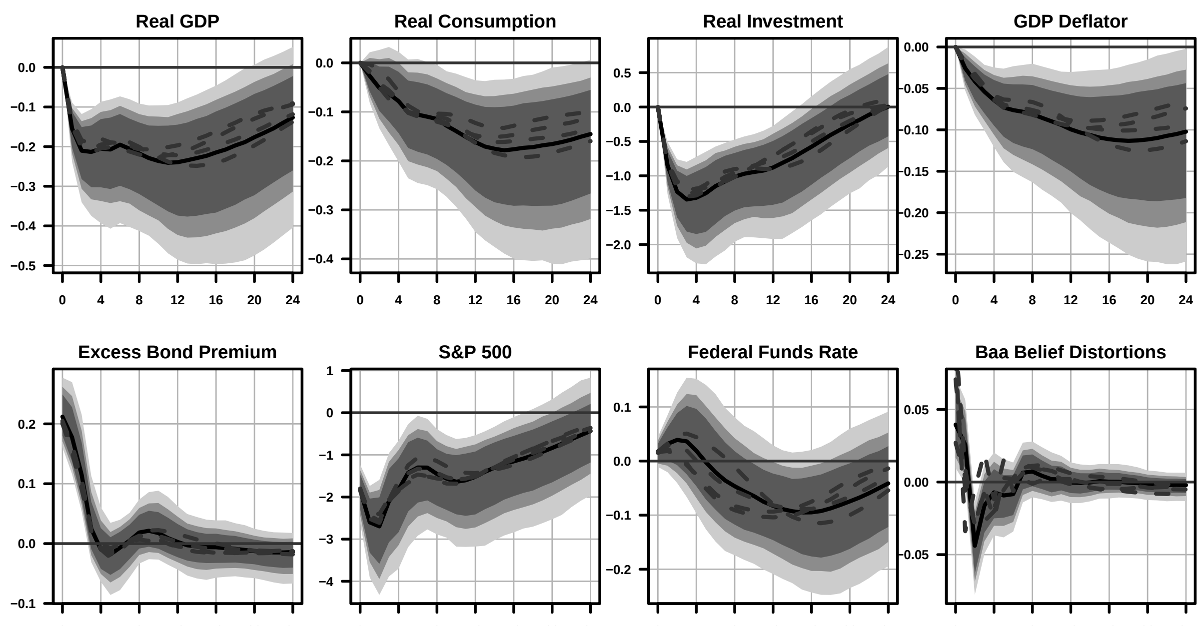
Notes: Impulse response functions of the VAR with ML forecast errors. Black line denotes median response while gray shaded areas denote the 68/80/90 percent confidence intervals. The responses of real GDP, real consumption, real investment, GDP deflator, and the stock market index are scaled in percent, while the excess bond premium, the federal funds rate, and belief distortions are scaled in percentage points.

Figure I5: Robustness: Impulse Response Functions to a Financial Shock (Baseline, Higher-Order Distortions).

(a) Model with Aaa Belief Distortion.



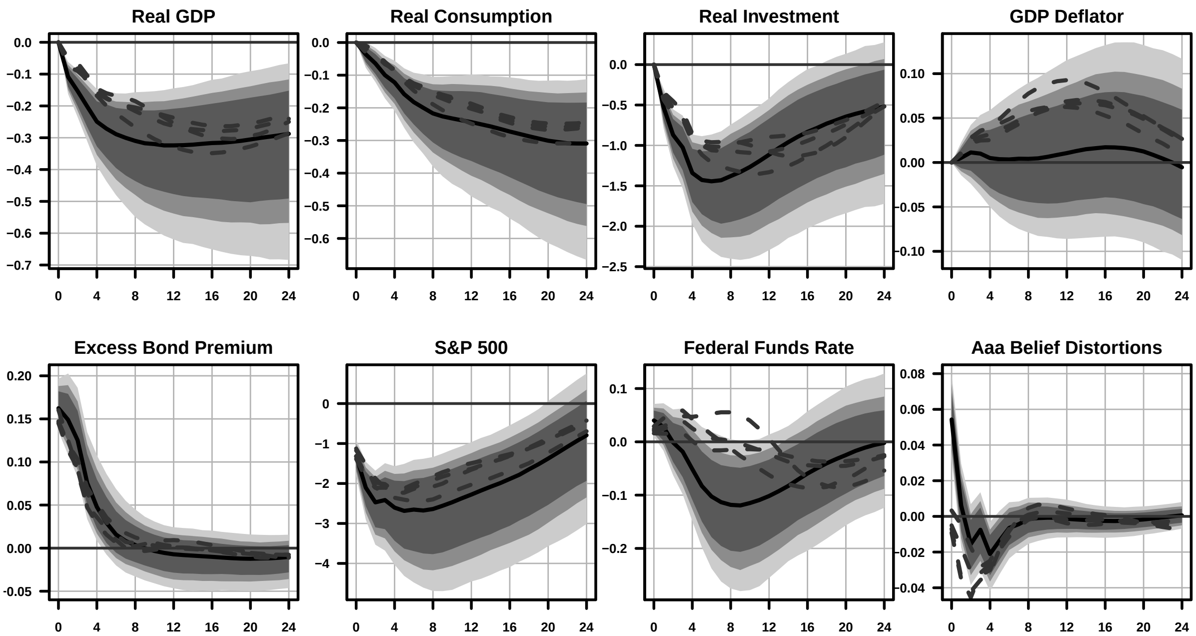
(b) Model with Baa Belief Distortion.



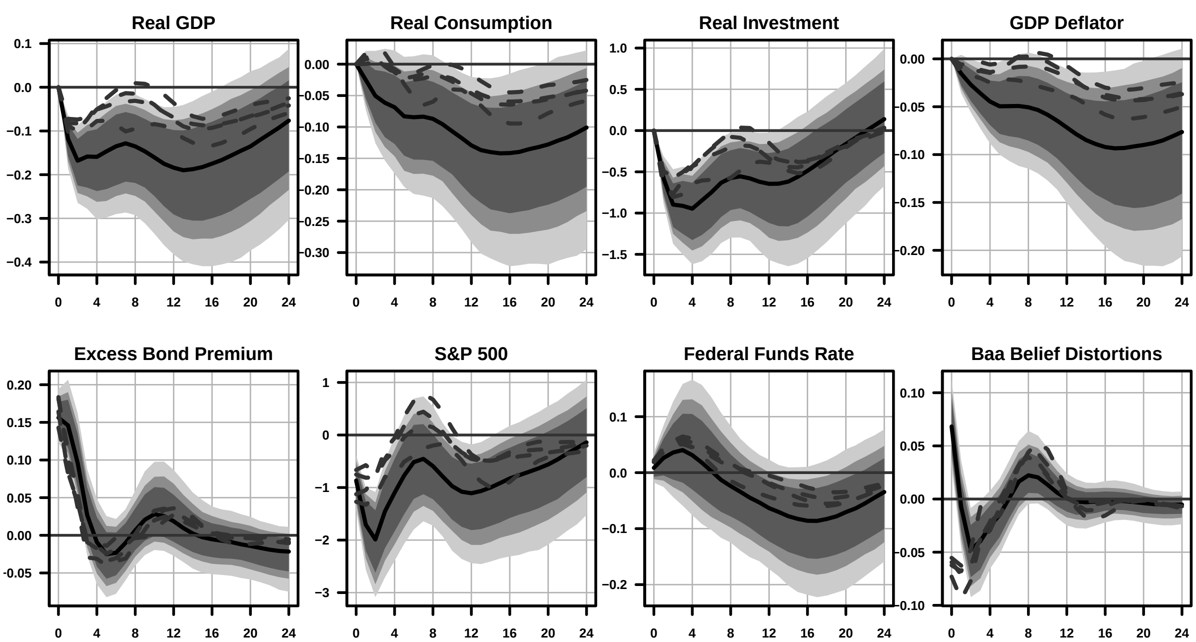
Notes: Impulse response functions of the baseline VAR. Black line denotes median response while gray shaded areas denote the 68/80/90 percent confidence intervals. Dashed, grey lines denote the median responses of the VAR with higher-order belief distortions (two-, three-, and four-step ahead) and mean belief distortions over all horizons (mean over one-, two-, three-, and four-step ahead). The responses of real GDP, real consumption, real investment, GDP deflator, and the stock market index are scaled in percent, while the excess bond premium, the federal funds rate, and belief distortions are scaled in percentage points.

Figure I6: Robustness: Impulse Response Functions to a Financial Shock (Survey Forecast Errors, Higher-Order Distortions).

(a) Model with Aaa Belief Distortion.



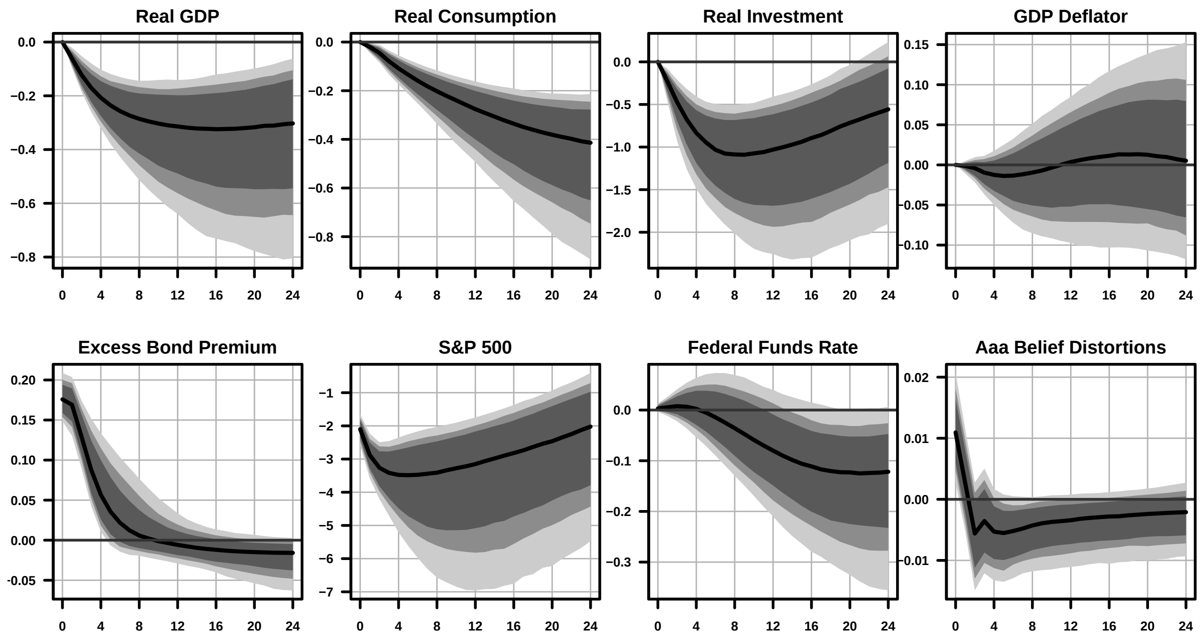
(b) Model with Baa Belief Distortion.



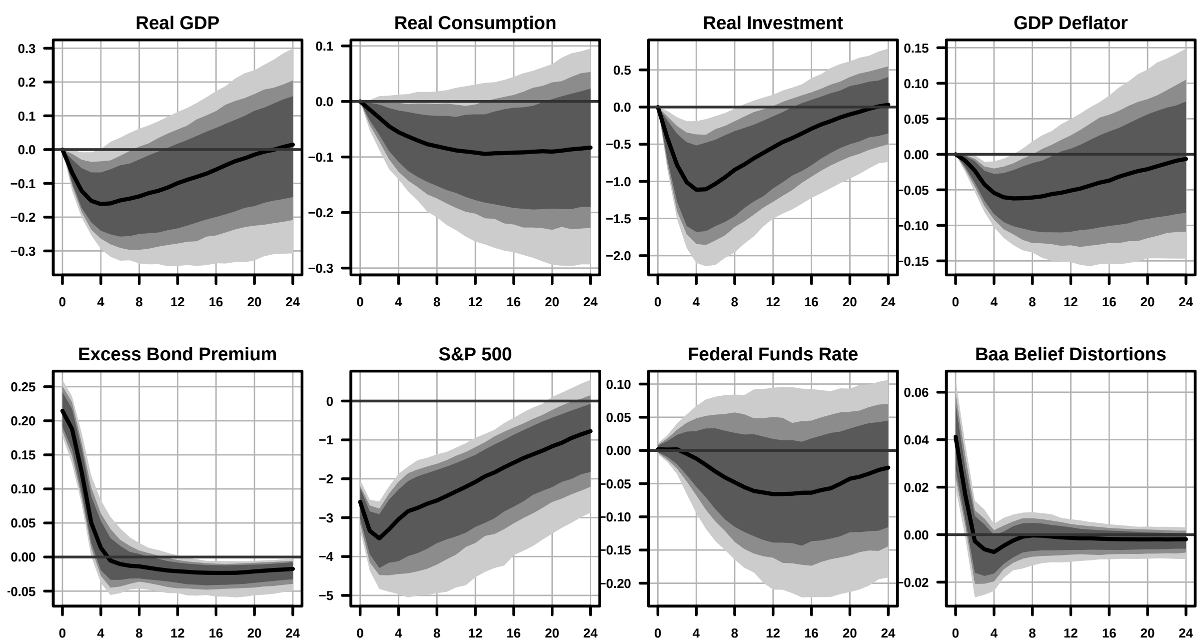
Notes: Impulse response functions of the VAR featuring survey forecast errors. Black line denotes median response while gray shaded areas denote the 68/80/90 percent confidence intervals. Dashed, grey lines denote the median responses of the VAR with higher-order belief distortions (two-, three-, and four-step ahead) and mean belief distortions over all horizons (mean over one-, two-, three-, and four-step ahead). The responses of real GDP, real consumption, real investment, GDP deflator, and the stock market index are scaled in percent, while the excess bond premium, the federal funds rate, and belief distortions are scaled in percentage points.

Figure I7: Robustness: Impulse Response Functions to a Financial Shock (Baseline, Horseshoe Shrinkage Prior).

(a) Model with Aaa Belief Distortion.



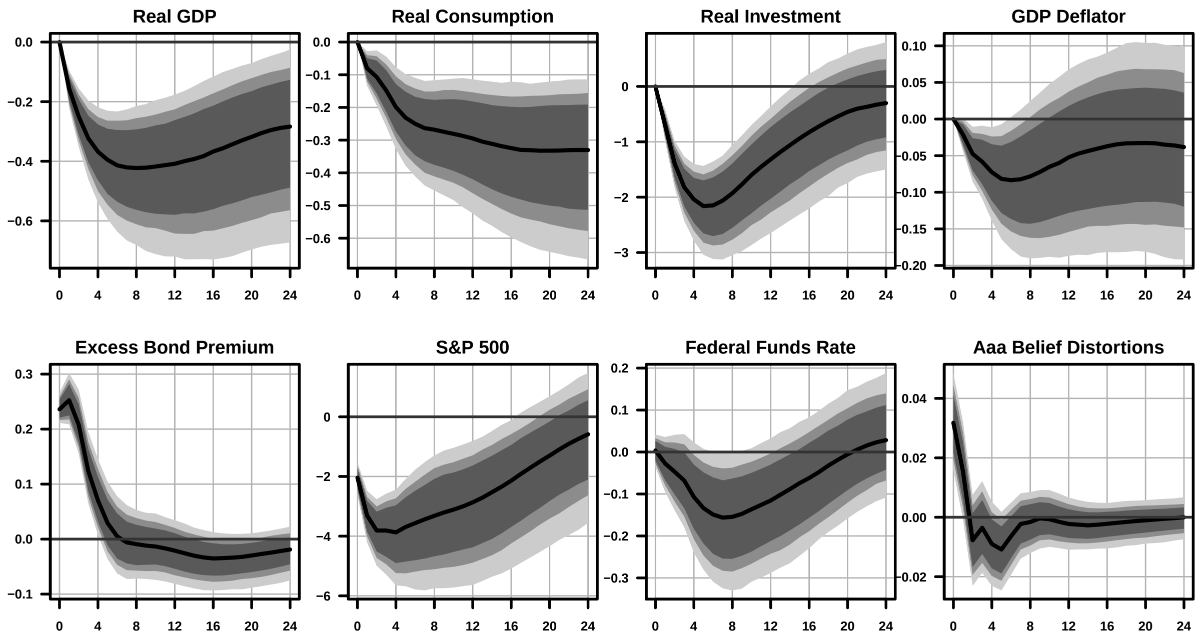
(b) Model with Baa Belief Distortion.



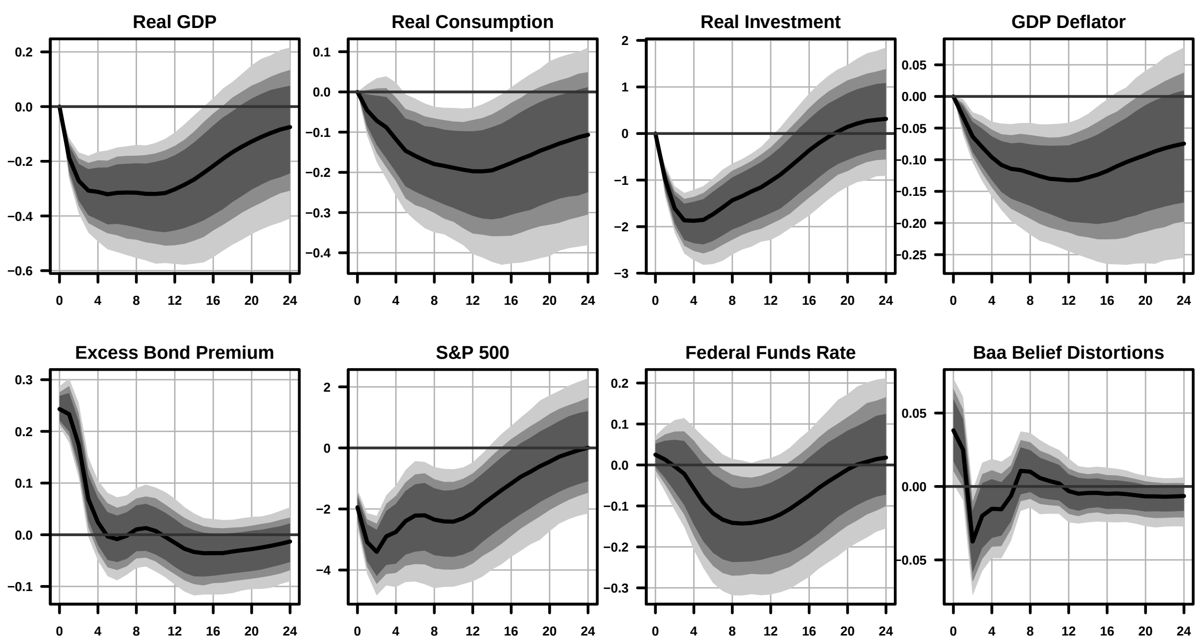
Notes: Impulse response functions of the baseline VAR estimated with the horseshoe shrinkage prior. Black line denotes median response while gray shaded areas denote the 68/80/90 percent confidence intervals. The responses of real GDP, real consumption, real investment, GDP deflator, and the stock market index are scaled in percent, while the excess bond premium, the federal funds rate, and belief distortions are scaled in percentage points.

Figure I8: Robustness: Impulse Response Functions to a Financial Shock (Baseline, No Stochastic Volatility).

(a) Model with Aaa Belief Distortion.



(b) Model with Baa Belief Distortion.



Notes: Impulse response functions of the baseline VAR estimated without stochastic volatility. Black line denotes median response while gray shaded areas denote the 68/80/90 percent confidence intervals. The responses of real GDP, real consumption, real investment, GDP deflator, and the stock market index are scaled in percent, while the excess bond premium, the federal funds rate, and belief distortions are scaled in percentage points.

Appendix References

Antolin-Diaz J., Petrella I. and Rubio-Ramirez J F. (2021) "Structural Scenario Analysis with SVARs". *Journal of Monetary Economics* Vol. 117, pp. 798–815.

Bai J. and Ng S. (2006) "Confidence Intervals for Diffusion Index Forecasts and Inference for Factor-Augmented Regressions". *Econometrica* Vol. 74(4), pp. 1133–1150.

Baker S R., Bloom N. and Davis S J. (2016) "Measuring Economic Policy Uncertainty". *Quarterly Journal of Economics* Vol. 131(4), pp. 1593–1636.

Bianchi F., Ludvigson S C. and Ma S. (2022) "Belief Distortions and Macroeconomic Fluctuations". *American Economic Review* Vol. 112(7), pp. 2269–2315.

Boeck M. and Zörner T O. (2023) *Natural Gas Prices and Unnatural Propagation Effects: The Role of Inflation Expectations in the Euro Area*. Tech. rep. Available at SSRN: https://ssrn.com/abstract=4376796.

Bordalo P., Gennaioli N., Ma Y. and Shleifer A. (2020) "Overreaction in Macroeconomic Expectations". *American Economic Review* Vol. 110(9), pp. 2748–82.

Boysel S. and Vaughan D. (2019) *fredr: An R Client for the 'FRED' API*. R package version 1.0.0.9000.

Breitenlechner M. (2018) *An Update of Romer and Romer (2004) Narrative U.S. Monetary Policy Shocks up to 2012Q4*. Online Notes (available here https://eeecon.uibk.ac.at/ breitenlechner/data/UpdateRR04.pdf.

Breitenlechner M., Georgiadis G. and Schumann B. (2022) "What Goes Around Comes Around: How Large Are Spillbacks from US Monetary Policy?" *Journal of Monetary Economics* Vol. 131, pp. 45–60.

Caldara D., Fuentes-Albero C., Gilchrist S. and Zakrajšek E. (2016) "The Macroeconomic Impact of Financial and Uncertainty Shocks". *European Economic Review* Vol. 88, pp. 185–207.

Carriero A., Chan J., Clark T E. and Marcellino M. (2022) "Corrigendum to "Large Bayesian Vector Autoregressions with Stochastic Volatility and Non-Conjugate Priors"[J. Econometrics 212 (1)(2019) 137–154]". *Journal of Econometrics* Vol. 227(2), pp. 506–512.

Carriero A., Clark T E. and Marcellino M. (2019) "Large Bayesian vector autoregressions with stochastic volatility and non-conjugate priors". *Journal of Econometrics* Vol. 212(1), pp. 137–154.

Carvalho C M., Polson N G. and Scott J G. (2010) "The Horseshoe Estimator for Sparse Signals". *Biometrika* Vol. 97(2), pp. 465–480.

Center for Research in Security Prices . (2022) *Annual Update 1926-2022*. https://wrds-www.wharton.upenn.edu/pages/get-data/center-research-security-prices-crsp/. Accessed from Wharton Research Data Services: 2022-05-22.

Chan J C. (2022) "Asymmetric Conjugate Priors for Large Bayesian VARs". *Quantitative Economics* Vol. 13(3), pp. 1145–1169.

Chan J C. and Eisenstat E. (2018) "Bayesian Model Comparison for Time-Varying Parameter VARs with Stochastic Volatility". *Journal of Applied Econometrics* Vol. 33(4), pp. 509–532.

Chipman H A., George E I. and McCulloch R E. (2010) "BART: Bayesian Additive Regression Trees". *Annals of Applied Statistics* Vol. 4(1), pp. 266 –298.

Data F R E. (2022) *Federal Reserve Economic Data Base*. https://fred.stlouisfed.org. Accessed: 2022-05-01.

Doan T., Litterman R. and Sims C. (1984) “Forecasting and Conditional Projection using Realistic Prior Distributions”. *Econometric Reviews* Vol. 3(1), pp. 1–100.

Fama E F. and French K R. (1992) “The Cross-Section of Expected Stock Returns”. *Journal of Finance* Vol. 47(2), pp. 427–465.

Fama E F. and French K R. (1993) “Common risk factors in the returns on stocks and bonds”. *Journal of Financial Economics* Vol. 33(1), pp. 3–56.

Federal Reserve Philadelphia . (2022) *Real-Time Data Set for Macroeconomicsts*. https://www.philadelphiafed.org/surveys-and-data/real-time-data-research/real-time-data-set-full-time-series-history. Accessed: 2022-05-10.

French K. (2022) *Research Returns Data (1926-2022)*. https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html. Accessed: 2022-05-22.

Geweke J. et al. (1991) *Evaluating the Accuracy of Sampling-Based Approaches to the Calculation of Posterior Moments*. Vol. 196. Federal Reserve Bank of Minneapolis, Research Department Minneapolis, MN.

Gilchrist S. and Zakrajšek E. (2012) “Credit Spreads and Business Cycle Fluctuations”. *American Economic Review* Vol. 102(4), pp. 1692–1720.

Huang D., Jiang F. and Tong G. (2018) *Real Time Macro Factors in Bond Risk Premium*. Tech. rep. Lee Kong Chian School of Business.

Jacquier E., Polson N G. and Rossi P E. (2002) “Bayesian Analysis of Stochastic Volatility Models”. *Journal of Business & Economic Statistics* Vol. 20(1), pp. 69–87.

Jarociński M. and Karadi P. (2020) “Deconstructing Monetary Policy Surprises – The Role of Information Shocks”. *American Economic Journal: Macroeconomics* Vol. 12(2), pp. 1–43.

Jurado K., Ludvigson S C. and Ng S. (2015) “Measuring Uncertainty”. *American Economic Review* Vol. 105(3), pp. 1177–1216.

Känzig D R. (2021) “The Macroeconomic Effects of Oil Supply News: Evidence from OPEC Announcements”. *American Economic Review* Vol. 111(4), pp. 1092–1125.

Kapelner A. and Bleich J. (2016) “bartMachine: Machine Learning with Bayesian Additive Regression Trees”. *Journal of Statistical Software* Vol. 70(4), pp. 1–40.

Kastner G. (2016) “Dealing with Stochastic Volatility in Time Series Using the R Package stochvol”. *Journal of Statistical Software* Vol. 69(5), pp. 1–30.

Kastner G. and Frühwirth-Schnatter S. (2014) “Ancillarity-Sufficiency Interweaving Strategy (ASIS) for Boosting MCMC Estimation of Stochastic Volatility Models”. *Computational Statistics & Data Analysis* Vol. 76, pp. 408–423.

Kilian L. (2009) “Not All Oil Price Shocks are Alike: Disentangling Demand and Supply Shocks in the Crude Oil Market”. *American Economic Review* Vol. 99(3), pp. 1053–69.

Kim S., Shephard N. and Chib S. (1998) “Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models”. *Review of Economic Studies* Vol. 65(3), pp. 361–393.

Kuhn M. (2022) *caret: Classification and Regression Training*. R package version 6.0-92.

Leeper E M. and Zha T. (2003) “Modest Policy Interventions”. *Journal of Monetary Economics* Vol. 50(8), pp. 1673–1700.

Litterman R B. (1986) “Forecasting with Bayesian vector autoregressions—five years of experience”. *Journal of Business & Economic Statistics* Vol. 4(1), pp. 25–38.

Lucas R E. (1976) “Econometric policy evaluation: A critique”. *Carnegie-Rochester Conference Series on Public Policy* Vol. 1, pp. 19–46.

Ludvigson S C., Ma S. and Ng S. (2021) “Uncertainty and business cycles: exogenous impulse or endogenous response?” *American Economic Journal: Macroeconomics* Vol. 13(4), pp. 369–410.

Makalic E. and Schmidt D F. (2015) “A Simple Sampler for the Horseshoe Estimator”. *IEEE Signal Processing Letters* Vol. 23(1), pp. 179–182.

Miranda-Agrippino S. and Ricco G. (2021) “The Transmission of Monetary Policy Shocks”. *American Economic Journal: Macroeconomics* Vol. 13(3), pp. 74–107.

Raftery A E. and Lewis S M. (1992) “[Practical Markov Chain Monte Carlo]: Comment: One Long Run with Diagnostics: Implementation Strategies for Markov Chain Monte Carlo”. *Statistical Science* Vol. 7(4), pp. 493–497.

Romer C D. and Romer D H. (2004) “A New Measure of Monetary Shocks: Derivation and Implications”. *American Economic Review* Vol. 94(4), pp. 1055–1084.

Romer C D. and Romer D H. (2010) “The Macroeconomic Effects of Tax Changes: Estimates Based on a New Measure of Fiscal Shocks”. *American Economic Review* Vol. 100(3), pp. 763–801.

Shiller R J. (2022) *Online Data Robert Shiller (1890-2022)*. `http://www.econ.yale.edu/~shiller/data.htm`. Accessed: 2022-05-23.

Sims C A. and Zha T. (1998) “Bayesian Methods for Dynamic Multivariate Models”. *International Economic Review*, pp. 949–968.

Waggoner D F. and Zha T. (1999) “Conditional Forecasts in Dynamic Multivariate Models”. *Review of Economics and Statistics* Vol. 81(4), pp. 639–651.

Zou H. and Hastie T. (2005) “Regularization and Variable Selection via the Elastic Net”. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* Vol. 67(2), pp. 301–320.