Essays on Information in Macroeconomics and Finance:

Author: Ethan Struby

Boston College Electronic Thesis or Dissertation, 2017

Boston College

Morrissey College of Arts and Sciences

Graduate School

Department of Economics

ESSAYS ON INFORMATION IN MACROECONOMICS AND FINANCE

a dissertation

by

ETHAN STRUBY

submitted in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

May 2017

Essays on Information in Macroeconomics and Finance

by

ETHAN STRUBY

Dissertation Committee:
RYAN CHAHROUR (Chair)
PETER IRELAND
SUSANTO BASU

Abstract

Expectations formation is central to macroeconomics. Households, firms, and policymakers must form expectations not only about fundamentals, but about what other agents' beliefs are, because others' beliefs will determine their actions. The three essays in this dissertation examine empirically and theoretically how agents use both public and private information to form expectations. The first two essays combine a models of optimizing behavior and forecasting with data on the macroeconomy, financial prices, and macroeconomic forecasts to examine the extent to which economic agents learn about the macroeconomy from financial prices and monetary policy actions. The third essay examines theoretically how members of a committee use public and private information to form beliefs when they care both about having accurate forecasts and coordinating actions with others. All three essays emphasize that frictions in expectations formation are a salient feature of the world, and understanding the extent and importance of those frictions is important for both positive and normative questions in macroeconomics and finance.

Beliefs about the future determine the willingness of financial market participants to save and invest, and theory suggests they should value more highly assets which are expected to pay higher returns during recessionary periods when consumption is otherwise low. Hence, financial prices

reflect macroeconomic expectations. In the first essay, titled **Macroeconomic Disagreement in Treasury Yields**, I explore how agents with idiosyncratic, private information form beliefs about both the macroeconomy and the beliefs of other agents. Using data on United States Treasury debt, the macroeconomy, and individual inflation forecasts, I estimate the precision of bond traders' information about the macroeconomy and how much they disagree with each other. I allow for traders to learn both from private signals and from asset prices, which aggregate the beliefs of all the traders in the market. I find that bond prices are moderately informative about macroeconomic variables, but are the source of most of the information traders have about monetary policy and the beliefs of others. In contrast to studies which assume full information, risk premia are much less important than slow-adjusting interest rate expectations for explaining the behavior of long-run yields. The most important signal for bond traders appears to be the Federal Reserve's short-run rate, which encodes information about the macroeconomy and the central bank's intended future policy. Nevertheless, the fact that traders held disparate beliefs about the macroeconomy, and especially about the long-run inflation target of the Federal Reserve, elevated long-term yields on average.

The first essay demonstrates empirically that financial market participants learn about the macroeconomy from monetary policy actions. However, it is silent on how monetary policymakers form beliefs about the macroeconomy, or how the information in monetary policy rates endogenously affects macroeconomic outcomes. In the second essay **Your Guess is as Good as Mine: Central Bank Information and Monetary Policy**, I use data on private sector forecasts and forecasts from the Federal Reserve Board staff to examine the typical assumption of common information between firms and monetary policymakers. Using forecasts from a survey of professional forecasters and from the Federal Reserve Board staff, I show evidence against the typical assumption of common information between monetary policymakers and the private sector, and also that policymakers are, at best, only weakly better at forecasting than private forecasters. Based on this evidence, I augment an otherwise standard monetary policy model by relaxing the common information assumption. Instead, I assume there is idiosyncratic, private information among price-setting firms, and between firms and the central banker. Firms combine private information about aggregate conditions with the observed monetary policy rate to form expectations about fundamentals and the beliefs of rival firms. The central banker must form expectations about firms' beliefs because those beliefs will determine inflation and overall economic activity. But as a result of their differences in information sets, firms must form expectations about other firms' expectations, and what the central banks' expectations of their expectations are. I examine the ability of this model to fit the data and find that the model can capture features of both firm and central bank inflation expectations, but in the absence of imperfect information among households, it is difficult to simultaneously match the forecast data and data on real activity. This result points to the sensitivity of models with dispersed information to the underlying assumptions about how central bankers will respond to exogenous shocks.

The second chapter emphasized how the assumptions economists make regarding monetary policymakers' information is critical for understanding their actions. Motivated by this example, my third chapter **Information Investment in a Coordination Game** explores theoretically how members of a committee who are uncertain about others' beliefs decide on a binary action, and how their decision to pay close attention to public or private signals is related to their desire to accurately forecast versus coordinating their behavior with others. I show that when it is assumed that information decisions among committee members are symmetric - everyone pays the same amount of attention to the same things - there is a unique outcome of the coordination game. However, I further show that it is difficult to guarantee that committee members will all choose a symmetric allocation of information. Aside from the direct cost of acquiring better information, allocating attention to more accurate signals can harm welfare when coordination motives are dominant. In a set of numerical exercises, however, I show that it is possible for a unique equilibrium to exist, and that actions that do not have a large impact on the payoffs of committee members (such as changing the size of the committee) may nevertheless have large impacts on the accuracy of the committee's forecasts. This suggests a possible tension between the welfare of the committee, which benefits from consensus, and the welfare of those affected by the committee's actions, which likely depends on whether the committee takes the objectively correct action.

My dissertation has important implications for both academic economists and policymakers. Understanding the sources of business cycle fluctuations and the determinants of asset prices requires grappling with the fact that people have differences in beliefs. Empirical evidence suggests that agents' beliefs are shaped by both idiosyncratic forces and by public announcements and policy decisions, and economists' models need to reflect these features of the world. Policy, too, is affected by the information available to policymakers, and to understand how policymakers have acted in the past and should act in the future, it is necessary to take seriously the ways their belief formation deviates from the full information rational expectations benchmark.

To Joules

and in memory of my mother, Marsha

Acknowledgments

This dissertation would not have been possible without the advice, help, and support of many mentors, colleagues, and friends.

I would like to thank in particular my committee members, Ryan Chahrour, Peter Ireland, and Susanto Basu, for their guidance. Ryan, thanks for the many patient hours spent helping me figure out numerous modeling, coding, and data issues; thanks also for putting up with my occasional outbreaks of impostor syndrome during your office hours. I have always appreciated your honesty and encouragement. Peter, thank you for your clear advice, your faith in my abilities, and the occasional music discussion. Susanto, thank you for lending me your seemingly endless breadth of economics knowledge and intuition, and for reminding me to read the classics.

I thank the Department of Economics and the Graduate School of the Morrissey College of Arts and Sciences for their support during my time at Boston College. My professors were a source of wonderful instruction, comments, and advice. I would especially like to thank Fabio Schiantarelli for advice on numerous topics, both academic and not, and especially for his support on the job market. I would also like to apologize to him for my terrible handwriting. I am grateful for the help of the Economics Department staff, especially Gail Sullivan for her patience and skill in helping negotiate bureaucracy within and without BC, and Rick Sanger for friendly conversation. I would also like to thank my fellow graduate students for their friendship and comments on research. Special thanks to John Lindner for being a great office mate, sounding board, midafternoon-snack-enabler, and friend. Thanks to John's wife Anna Marie for her friendship and relentless positivity. Thanks also to Michael Connolly and Dominique Brabant for their friendship and thoughtful comments on many presentations and drafts. Thank you especially to Ana Lariau, Francesca Toscano, Igor Karagodsky, and Alex Eiermann for job market commiseration, and to Jake Penglase, Solvejg Wewel, Giacomo Candian, Pierre De Leo, and my many other friends in the department.

I am grateful for many friends outside of the Economics Department who supported me. Particularly, I would like to thank Tom Hogan for nearly fifteen years of friendship and reassurance that things were going to work out for the best - I promise I'm going to relax, eventually. Thanks also to Tom's family for putting up with me for so long. Sorry about eating all the peanuts. Thanks to Laura Hogan for helping Tom survive his PhD and for her friendship since high school. Thanks to Kate Koegel for being an excellent debate partner and friend, and to our high school debate coach Melissa Reynolds for her skill and patience in teaching me (and Tom, and Kate, and many others) how to do research and construct arguments. I'm also grateful for my friends from Grinnell College for constantly encouraging me to be a more thoughtful and engaged citizen of the academic community and of the world. Special thanks to Harry Krejsa for constant helpful distractions and political discussions, and to Amy Henning and Meredith Nechitilo for years of friendship, advice, and reminders to not take myself too seriously.

I would never have been able to make it this far without the love and support of my family, particularly my siblings, their spouses, and my nieces and nephews. Thanks to my father, Carl Struby, for support and for the appropriate amount of worry about my math grades in high school. I would also like to acknowledge the memory of my mother, Marsha Babicki Struby, an Economics major at the University of Kansas. I still miss your advice and conversation every day.

Finally, despite her claims to the contrary, there is no way I would have been able to finish this dissertation without the love, support, and friendship of Julia Stewart. Julia has constantly encouraged and challenged me to be more thoughtful in all things, and was endlessly tolerant and supportive during many stretches of difficulty with research. I could not ask for a more patient or kind partner, and I am so grateful for your presence in my life. I love you.

Contents

List of Tables

List of Figures

Chapter 1

Macroeconomic Disagreement in Treasury Yields

Abstract

I estimate an affine term structure model of Treasury yields featuring idiosyncratic, private information about macroeconomic and policy conditions. The estimation uses data on U.S. long-term bond yields and inflation forecasts to identify the precision of bond traders' information and their degree of disagreement. The results imply: (1) Bond prices are moderately informative about the macroeconomy, but very informative about policy and others' beliefs (2) Bond traders' beliefs are still dispersed despite a large number of endogenous public signals (3) The short term interest rate is more informative than other interest rates (4) Accounting for agents' learning dynamics dramatically reduces the magnitude and volatility of risk premia relative to a full information benchmark. Overall, I find that the failure of common knowledge adds an average of 60 basis points to ten year yields, with most time variation in this wedge attributable to disagreement about the Federal Reserve's inflation target.

I gratefully acknowledge advice and suggestions from Susanto Basu, Ryan Chahrour, and Peter Ireland at every stage. Thanks also to Rosen Valchev, Dongho Song, Michael Connolly, and Giacomo Candian, seminar participants at the FDIC, the SEC, and Carleton College, and participants at the Boston College Macro Lunch and dissertation workshop.

1.1 Introduction

Professional forecasters generally disagree about the evolution of the macroeconomy. Survey evidence reveals that even during normal times, there is nontrivial disagreement about what inflation will be in the current quarter (figure 1.1). How does macroeconomic disagreement affect the price of long-term assets? And what can the dynamics of asset prices reveal about beliefs? Accounting for differences in beliefs may affect our assessment of the causes of bond price fluctuations; whether we care about bond prices for their own sake, to evaluate the effects of policy, or to inform the implications of equilibrium models, it is important to understand the process of belief formation that determines bond prices. Furthermore, understanding what asset prices tell us about the information financial market participants have can help us assess the empirical plausibility of more structural macroeconomic models with dispersed beliefs.

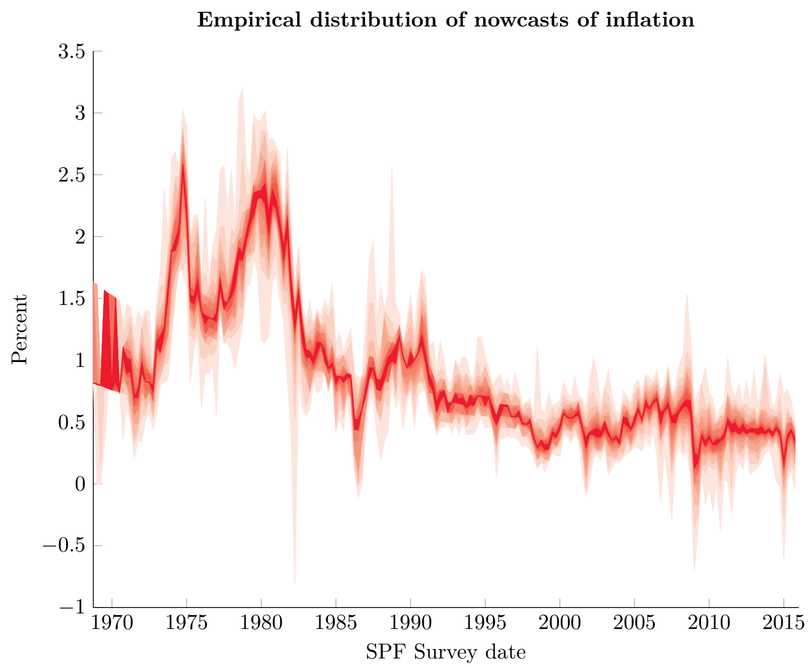


Figure 1.1: Distribution of forecasts from the SPF of current-quarter inflation as measured by change in the log GDP deflator.

To answer these questions, I estimate an affine term structure model (ATSM) of Treasury yields, where short term interest rates and macroeconomic variables are described by a structural vector autoregression. I relax the usual assumption in the asset pricing literature that agents share a common information set. Instead, I assume their information is *dispersed*: bond traders optimally combine noisy, idiosyncratic signals with the prices they observe to form expectations. Because they care about other traders' (unobserved) beliefs, they must form expectations of not only fundamentals, but also expectations about others' expectations, and about others' expectations of others' expectations, etc. To account for this "forecasting the forecasts of others" problem, the model solution is cast in terms of a fixed point problem between the evolution of agents' beliefs about the macroeconomy and the prices used to form those beliefs. I estimate the model on U.S. data from Q41971-Q42007, including data on individual forecasts from professional forecasters to discipline the belief of agents in the model.

The estimated model allows me to quantify the sources of agents' information. My estimates imply roughly half of what bond traders know about macroeconomic factors (deviations of inflation from the Federal Reserve's implicit target and the output gap) comes from observing asset prices, rather than their private information. Asset prices are more informative about policy risks - here, risks related to the inflation target - and are the source of nearly everything traders know about the beliefs of others. The price most informative for agents in that of one-period bonds, which is determined in the model by a Taylor-type rule. The short rate, combined with agents' idiosyncratic information contains nearly all the information agents have about fundamentals and others' beliefs, with longer yields adding only small fraction of additional information. The importance of the policy rate in expectations formation adds to growing evidence of a signaling channel of monetary policy (for example, Melosi (2017), Tang (2013)).

I use the estimated model to understand the determinants of bond prices. Because agents' expectations about others' expectations ("higher order" expectations) differ from average expectations about fundamentals, prices differ from those that would obtain if traders counterfactually held common beliefs. The difference is a wedge directly attributable to dispersed information. I find that this higher order wedge plays a direct role in yields; on average, it contributed 60 basis points to ten year yields over the sample period. In my setting, this wedge can be meaningfully decomposed into components driven by different macroeconomic variables. I find that the majority of time variation in the wedge for long-term debt is attributable to changes in higher-order beliefs about monetary policy, particularly, policymakers' long-run inflation target.

After estimating the model, I decompose yields into average expected short rates over the life of the bond (the "expectations hypothesis" component), the higher order wedge, and "classical" compensation for risk. The model attributes the vast majority of movement in long-term yields to rate expectations which adjust slowly relative to what a full information version of the model would suggest. This is a consequence of traders' inference problem; they attribute some changes in fundamentals to noise in their idiosyncratic signals, and they misattribute transitory shocks to extremely persistent changes in the Federal Reserve's inflation target. This suggest much of the "excess sensitivity" of long term yields to short-term macroeconomic news (noted by Gurkaynak et al. (2005)) is attributed by the model to violations of the auxiliary assumption of full information rational expectations. "Classical" risk premia are estimated to be quite small and nearly constant across all maturities, in sharp contrast to their full information counterparts.

The dispersed information affine term structure model is builds on the work of Barillas and Nimark (2015). They assume yields are driven by three latent factors in the yield curve and identify rate expectations using interest rate forecasts. They find a somewhat larger role for the higher order wedge in explaining yields. In contrast to their latent factor approach, I explicitly model the relationship between short rates, macroeconomic variables, financial risk, and monetary policy. I particularly generalize the structural VAR of Ireland (2015) to incorporate dispersed information. In the model, the central bank is assumed to set an inflation target, and to set short rates in response to deviations from that target, the output gap, and changes in risk premia. The dynamics of these fundamentals are identified using structural assumptions. Changes in market prices of risk (and thus risk premia) are governed by changes in a single variable, consistent with Cochrane and Piazzesi (2005) and Bauer (2016). Shocks to this variable are correlated with macroeconomic shocks, and the level of the variable that governs movement in risk premia is allowed to affect macroeconomic dynamics; hence, the model allows for more links between the macroeconomy and financial markets than the estimated Taylor rules of Ang, Dong and Piazzesi (2007).

Because the model explicitly accounts for links between monetary policy and the prices of bonds, the results shed light on the relationship between policy uncertainty and yields. This paper is particularly related to the branch of the macro-finance literature that relates long-maturity bond price movements to changes in the monetary policy framework. For example, Gurkaynak, Sack and Swanson (2005) suggest incorporating learning about a (possibly time varying) long-run inflation target can help macro-finance models explain the effect of transitory shocks on long-term bonds. This paper extends this idea to the entire term structure. Moreover, I allow for pervasive information frictions about macroeconomic variables, and the estimated results quantify how important learning dynamics are for fluctuations in the prices of bonds at different maturities. My results also complement those of Wright (2011). He estimates term premia across different countries using both a term structure model and forecast surveys, and links declines in measured term premia to falling inflation uncertainty due to changes in the conduct of monetary policy. In my paper, the relationship between monetary policy and uncertainty is self-contained; dispersion of beliefs about the level of the inflation target could be interpreted as disagreement about policymakers' tolerance for inflation in the long run.[1] My results imply the decline in long-run yields in the United States is explained by falling average rate expectations, driven by a decline in the Federal Reserve's inflation target. Moreover, higher-order beliefs about the inflation target became a smaller component of prices of long-run debt over the course of the Great Moderation, implying that disagreement about

[1]Doh (2012) estimates a model where agents have a noisy signal of trend inflation, which he interprets as the credibility of the inflation target.

the target contributed less to yields. To the extent the target could be interpreted as the level of inflation central bankers were perceived as being credibly able to achieve, the results of my paper are complementary to Wright's.

My finding of small risk premia and persistent short rate expectations, stands in contrast to the literature that assumes bond yields are determined by full information rational expectations. My results add to growing evidence that accounting for information frictions tends to make time varying risk premia less important for explaining yields. Critically, the "slow" adjustment of rate expectations holds even with optimal Bayesian learning where agents have full understanding of the model structure and access to a large number of signals. This stands in contrast to other papers (for example, Dewachter and Lyrio (2008)) who assume traders' forecasts are based on a model-inconsistent prior. Moreover, my structural results are consist with the more agnostic approach of Piazzesi, Salomao and Schneider (2013) who construct subjective beliefs without modeling inference. Unlike their paper, however, I am able to numerically characterize the information content of different signals.

The findings of this paper should also be of interest to researchers working with more detailed dynamic general equilibrium or financial models featuring information frictions. The term structure model makes relatively modest structural and functional form assumptions. Unlike many exogenous information models, I do not restrict agents from learning from prices they encounter (in line with the "market consistent information" assumption advocated by Graham and Wright (2010)). The empirical results allow me to generate estimates of the plausible degree of information dispersion about the macroeconomy consistent with asset price movements, and of how informative prices are for agents. Despite having less structure, the ATSM is consistent with the pricing implications of many general equilibrium models: Barillas and Nimark (2015) show the dispersed-information ATSM nests an equilibrium model with wealth-maximizing traders (as in Barillas and Nimark (2016)). A number of authors have also embedded ATSM in DSGE models with the term structure (for example

Jermann (1998), Wu (2001), Doh (2012)). A result of the estimation is that belief dispersion is sustained and important, despite agents' access to a large number of informative common signals. Furthermore, the estimated results point to prices as an important source of information for agents making investment decisions. While this feature of prices has a long intellectual history mentioned in the next section, it has not been explored as much in the recent literature on macroeconomic models with dispersed information. My results suggest structural macroeconomic models with dispersed information should not ignore the role of prices, especially asset prices, as a source of information about the macroeconomy and monetary policy.

In the next section, I discuss in more detail the relationship of this paper to the existing literature on asset pricing, especially asset pricing with non-full information rational expectations. Section 1.3 presents some reduced-form and graphical evidence of information frictions in financial forecasts. I outline the asset pricing side of the model, the macroeconomic VAR, and solution and estimation strategy in sections 1.4 and 1.5. I then discuss the parameter estimates, the information content of signals, and the model's interpretation of the sources of yield fluctuations before concluding.

1.2 Related literature

The model in this paper is an affine term structure model that assumes the absence of arbitrage, combined with a structural macroeconomic VAR. Although it is most closely related to the specific model of Ireland (2015), a number of authors have estimated models combining structural macroeconomic elements with a no-arbitrage finance model, assuming that agents have full information rational expectations. Ang, Dong and Piazzesi (2007) estimate Taylor rules in such a setting. Rudebusch and Wu (2008) link yields to a dynamic New Keynesian model. Unlike these papers, I focus on disagreement about the macroeconomy and do not assume agents have full information.

My emphasis on learning from prices means that this paper is closely related to the noisy rational expectations literature. An influential classic in this literature is Grossman and Stiglitz (1980). The basic model of Grossman and Stiglitz was been extended by Hellwig (1980) and Admati (1985). More recently, Mondria and Quintana-Domeque (2012) use rational inattention in a two-market setup similar to Admati (1985), but where prices are endogenously noisy. Hassan and Mertens (2014) embed the Hellwig noisy rational expectations model in a DSGE model to study the equity premium. My model is less structural than these models to facilitate estimation while retaining a complicated inference problem with many assets and fundamentals.

A second classic asset pricing literature, associated with Harrison and Kreps (1978), focuses on disagreement about fundamentals but abstracts from learning. Harrison and Kreps assume there are discrete agent types with different priors who "agree to disagree." Institutional or financial frictions prevent agents from making side bets which would make each type satisfied with their positions. Hence, agents may want to pay more for an asset today in order to resell it later. This form of heterogeneity has been used by Barsky (2009) to explain the Japanese asset bubble; Simsek (2013) generalizes the model and shows how it can rationalize financial innovations similar to those that emerged in the subprime mortgage market. Cao (2011) develops a dynamic version of the Harrison and Kreps model with collateral constraints. He shows incorrect beliefs can be sustained in equilibrium because agents "profit by speculating." Unlike the papers in this literature, I assume any differences in agents' beliefs are driven by differences in observed signals. In this way, the model in this paper is consistent with the "Harsanyi doctrine" (Harsanyi (1968)); agents will have full information about the structure of the model and its parameters, and form expectations optimally. Only differences in information gives rise to differences in belief.[2]

This paper falls primarily into the recent literature on deviations from full information rational expectations in asset pricing. Much of this literature retains the assumption of common information and thus ignores the "forecasting the forecasts of others" problem or assume agents are not Bayesian learners. Piazzesi, Salomao and Schneider (2013) use forecasts to construct subjective bond risk

[2]Aumann (1976) points out that two agents with common priors whose posteriors are common knowledge cannot "agree to disagree." Here, posterior beliefs of particular agents about the state will not be common knowledge, and thus need not be the same despite a common prior. Posterior beliefs about prices will, of course, be common knowledge because they are commonly observed.

premia, but they abstract from agents' inference procedure. Piazzesi and Schneider (2007) examine how different assumptions about information affect risk premia in a representative investor setting. Sinha (2016) shows how adaptive learning can account for perceived failures of the expectations hypothesis. Dewachter and Lyrio (2008) use the restrictions implied by a three-equation Dynamic New Keynesian model to govern the evolution of macroeconomic variables that are priced in an affine term structure model where agents have misspecified priors. Collin-Dufresne et al. (2016) examine how Bayesian learning about parameters related to long-run risks, rare disasters, and model uncertainty can help a general equilibrium model explain risk premia.

The asset pricing literature that allows for differences in belief tends to abstract from higher order beliefs, the macroeconomy, or both. Giacoletti et al. (2015) also develop an arbitrage-free term structure model with belief dispersion about the parameters that govern latent risk factors but explicitly ignore the "forecasting the forecasts of others" problem. Colacito et al. (2016) develop an equity pricing model that includes variance and skewness of professional forecasts, which they treat as exogenous processes. Makarov and Rytchkov (2012) show how the state space of a dynamic asset pricing model with dispersed information can be infinite-dimensional, and that information asymmetries affect the time-series properties of returns. Kasa et al. (2014) solve a particular asset pricing model with higher-order expectations in the frequency domain. Their focus is on how the model can generate failures of tests of present value models.

Finally, this paper is related to a literature that seeks to explain the beliefs implied by forecast surveys. Examples include Patton and Timmerman (2010), Andrade et al. (2014) and Crump et al. (2016). Like these papers, I use the cross-section of forecasts at different horizons to help identify agents' belief formation, in a setting where agents are assumed to understand that macroeconomic and financial variables are related to each other. However, I infer the beliefs of bond traders treating both forecasts and prices as endogenous.

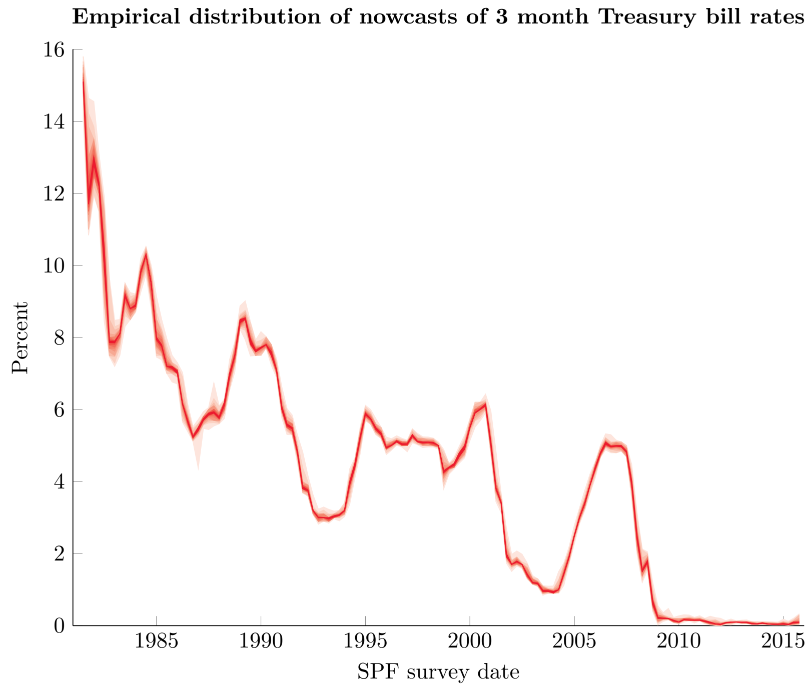


Figure 1.2: Distribution of SPF forecasts of *current*-quarter average rate on 3-month Treasury bill.

1.3 Dispersed information: evidence from forecasts

A number of authors (such as Mankiw et al. (2004) and Coibion and Gorodnichenko (2012a, 2015)) have documented evidence of information frictions using forecast data. With the exception of Coibion and Gorodnichenko (2015), most papers have focused on inflation expectations, generally finding average forecast errors are predictable and beliefs appear to adjust slowly to shocks. In this section, I briefly discuss some evidence for the presence of dispersed information about the evolution of Treasury bond prices in particular.

I take data on forecasts from the Survey of Professional Forecasters (SPF), a quarterly survey originally conducted by the American Statistical Association and the NBER before being taken over by the Federal Reserve Bank of Philadelphia in 1990. The survey is generally sent out after the first month of each quarter (after the initial release of the National Income and Product Accounts) to a panel of forecasters in the financial services industry, non-financial private sector, and academia.

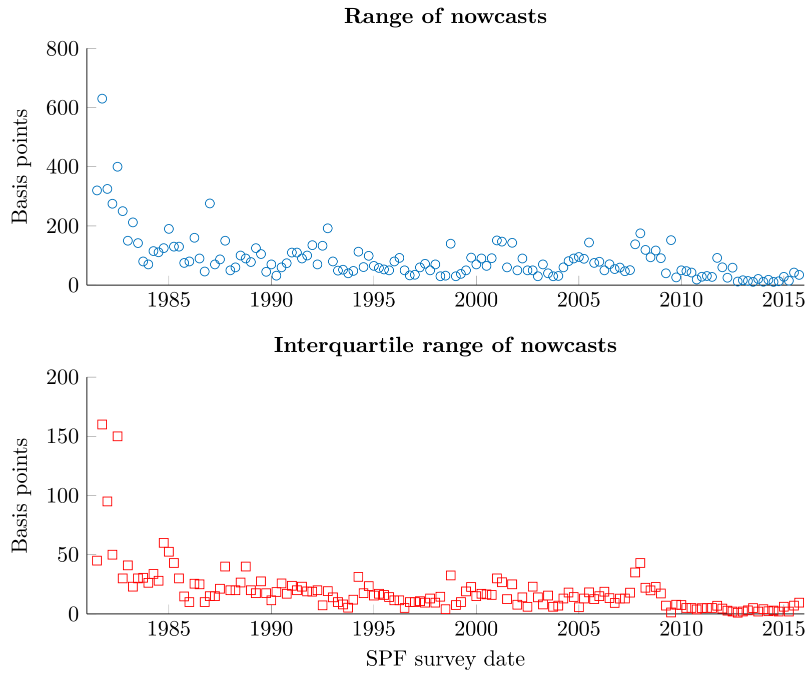


Figure 1.3: Full range (top) and interquartile range (bottom) of current-quarter average rate on 3-month Treasury bill.

The SPF began to survey its panel about 3-month Treasury bill rates in 1981. The 5th through 95th percentiles of the current-quarter forecasts are found in figure 1.2, and the range of forecasts in figure 1.3. Despite the fact that the rate for a Treasury bill in the secondary market is observable freely in real time to survey participants, there is still a fair amount of disagreement among forecasters within the current quarter - that is to say, the forecasters surveyed in the SPF disagree about what the average yield of Treasury bills will be over the course of the next two months. The interquartile range of forecasts, even including the zero lower bound period where Treasury bill rates were also effectively zero, is still nearly 20 basis points, with the overall range of forecasts often in the neighborhood of 100-200 basis points. To place these ranges in context, the yield on 3-month Treasuries was about 436 basis points between 1981-2015. From 2008-2015, the average was 24 basis points. The pictures for one-quarter ahead forecasts, shown in appendix 1.A, are similar. The striking fact that emerges is that the evolution of the price of an essentially risk-free asset is the subject of nontrivial disagreement among professional forecasters.[3]

To more formally test for information frictions, I adapt the empirical strategy of Coibion and Gorodnichenko (2015) to the forecasts of bond prices in the SPF.[4] For simplicity, assume Treasury bill rates follow an AR(1) process but agents observe idiosyncratic, noisy signals about the realization of that process.[5] Innovations and signal noise are assumed to be normally distributed and mean zero:

$$\begin{aligned} r_t = &\quad \rho r_{t-1} + \varepsilon_t \text{ with } \rho \in [0, 1) \\ r_{it} = &\qquad r_t + e_{it} \end{aligned}$$

Assuming agents are Bayesian learners, their conditional expectations can be written as:

$$\begin{aligned} E_t^i r_t = \kappa r_{it} + (1-\kappa) E_{t-1}^i r_t \\ E_t^i r_{t+h} = \rho^h E_t^i r_t \end{aligned}$$

Their expectation of the short rate is a weighted average of their current signal and their prior, where κ is the relative weight placed on the signal. Anticipating notation used later, I use $r_{t|t}^{(1)}$ to indicate the average expectation of r_t at time t.

Averaging across agents and rearranging gives the relationship between the forecast error for the *average* forecast and the revision of the average forecast at each horizon h:

$$r_{t+h} - r_{t+h|t}^{(1)} = \frac{1-\kappa}{\kappa} \left(r_{t+h|t}^{(1)} - r_{t+h|t-1}^{(1)} \right) + \sum_{j=1}^{h} \rho^{h-j} \varepsilon_{t+h} \tag{1.1}$$

where the error term is the sum of rational expectations errors. If signals were perfectly informative, $\kappa = 1$, and there would be no weight on forecast revisions in this regression. To the extent agents face information frictions, $\kappa < 1$. The simple reduced-form test of information frictions in financial forecasts amounts to projecting forecast revisions on forecast errors; the null hypothesis of full

[3]Recall that under rational expectations with common information, even if that information is imperfect, forecast distributions will be degenerate.

[4]Coibion and Gorodnichenko (2015) consider Treasury bill forecasts as part of their pooled regressions in a robustness test, but do not explicitly test for information frictions using financial asset forecasts alone.

[5]In the model developed in the next section, prices will be endogenous objects determined by fundamentals that agents have noisy signals about. A disconnect between the model and the results presented here is in the model that I will assume agents view the prices of current-quarter assets without error.

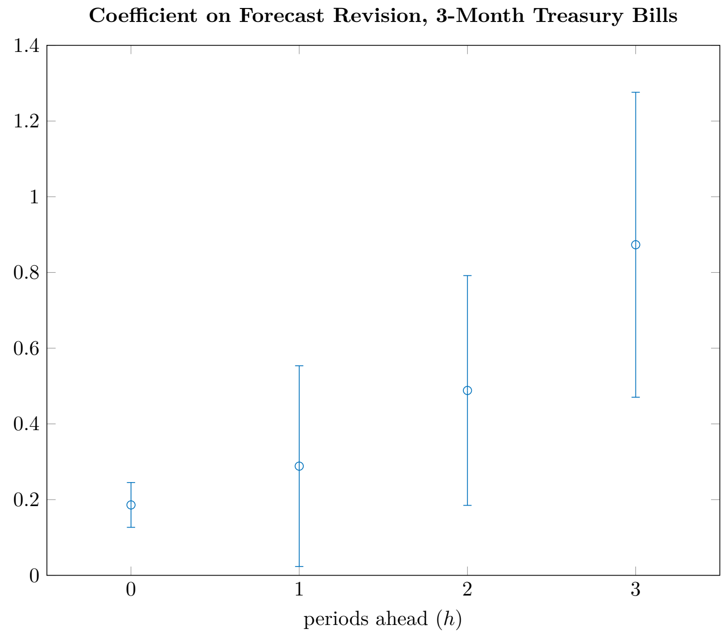


Figure 1.4: Coefficients from regression (1.2) for treasury bills. Bands represent 90 percent HAC confidence bands.

information rational expectations is equivalent to testing whether the regression coefficient is 0. Finding a significant positive coefficient, on the other hand, suggests information frictions. The regression takes the form

$$\text{Average Forecast Error}_{t,h} = \beta(\text{Average Forecast Revision}_{t,h}^{t,t-1}) + \bar{\varepsilon}_t \tag{1.2}$$

where $\bar{\varepsilon}_t$ is the sum of rational expectations errors as before.

The results of conducting this for different forecast horizons are shown in figure 1.4 for 3-month Treasury bills. The results are broadly consistent with Coibion and Gorodnichenko (2015)'s findings for inflation. The estimated coefficient is positive and at least marginally significant, suggesting average forecasts for financial variables reflect dispersed information among individuals. The response for 10 year bonds (not shown) are more mixed and have a high degree of uncertainty, probably reflecting the fact that the sample of available forecasts is much smaller. However, the point estimates are consistently positive and generally significant.

Graphical and reduced-form evidence suggests professional forecasts of macroeconomic and financial variables are inconsistent with those forecasters having full information rational expectations. Moreover, using the mean forecast for short-maturity debt, there is a positive, significant relationship between average forecast errors and average forecast revisions. This is inconsistent with full information rational expectations, but is consistent with a world where agents face information frictions.In the remainder of the paper, I explore the extent to which these facts are related and how much this disagreement affects prices.

1.4 The dispersed information model

In this section, I outline the model of asset prices and macroeconomic dynamics used to assess the effect of macroeconomic disagreement on prices. The asset pricing intuition and derivation in the next subsection closely follows that of Barillas and Nimark (2015); some additional details are found in appendix 1.B.

1.4.1 The term structure model with heterogeneous information

Intuition: the fundamental asset pricing relationship. Index bond traders by $j \in [0,1]$. Denote $E_t^j x_t = E[x_t|\Omega_t^j]$ as the expectation conditional on j's information set at time t (Ω_t^j). Call Ω_t the "full information" information set (i.e., the history of the realizations of all variables up to time t).

The basic asset pricing equation for a zero-coupon bond is

$$P_t^n = E_t[M_{t+1} P_{t+1}^{n-1}] \tag{1.3}$$

Standard results in asset pricing theory give that the nominal stochastic discount factor M_{t+1} exists and is positive if the law of one price holds and in the absence of arbitrage (Cochrane (2005)). If we relax the common information assumption, instead assuming there are a continuum of agents

$j \in (0, 1)$ with heterogeneous information sets, the pricing relationship for each agent j is:

$$P_t^n = E_t^j \left[M_{t+1}^j P_{t+1}^{n-1} \right] \tag{1.4}$$

Two things are specific to each agent j: Information sets (Ω_t^j) and stochastic discount factors (M_{t+1}^j). Centralized trading means that it is common knowledge that all agents face the same prices today and will face the same price tomorrow, while the fact that agents are atomistic implies they take prices as given. However, allowing information sets and forecasts of future prices to differ across agents, while assuming today's price is common knowledge, implies (1.4) can hold with equality only if stochastic discount factors also differ.

To decide their willingness to pay for a bond today, agents must form expectations of tomorrow's prices. Tomorrow's buyers face the same problem, so the decision to purchase a bond today depends on a conjecture about others' (future) beliefs - that is, they face the Townsend (1983) "forecasting the forecasts of others" problem. More specifically dispersion of information implies that asset prices depend on "higher order expectations" - expectations of expectations.[6] Assuming common knowledge of the pricing equation, joint lognormality of prices and stochastic discount factors, and deterministic conditional variances one can show (appendix 1.B) the log price of the bond takes the following form:

$$\begin{aligned} p_t^n = & \int E_t^j[m_{t+1}^j]dj \\ & + \int E_t^j \left[\left(\int E_{t+1}^k[m_{t+2}^k]dk + \int E_{t+1}^k[p_{t+2}^{n-2}]dk \right) \right] dj \\ & + \frac{1}{2}\mathrm{Var}(m_{t+1}^j + p_{t+1}^{n-1}) + \frac{1}{2}\mathrm{Var}(m_{t+2}^k + p_{t+2}^{n-2}) \end{aligned} \tag{1.5}$$

The price of a bond in period t is a function of the average expected stochastic discount factor in $t+1$ plus the average expectation of the average SDF and price at $t+2$, plus variances. Repeatedly recursively substituting allows us to write prices today as a function of average higher

[6] As alluded to in the literature review, the role of higher order beliefs in asset pricing is discussed by Allen et al. (2006), Bacchetta and Van Wincoop (2008), and Makarov and Rytchkov (2012). Barillas and Nimark (2015) consider the particular case of zero-coupon government debt.

order expectations about future SDFs and variance terms.[7] The model outlined below is consistent with the assumptions made here, but puts additional structure on the stochastic discount factor; doing so makes it easier to characterize how agents form higher order expectations and how those expectations affect bond prices.

Short rates and higher order expectations. Let x_t be a vector of exogenous factors and conjecture that the one-period risk free rate r_t is

$$r_t = \delta_0 + \delta_x' x_t \tag{1.6}$$

Assume there are d elements in x_t. I refer to x_t as "fundamentals" or "fundamental factors." Fundamentals follow a VAR(1):

$$x_{t+1} = \mu^P + F^P x_t + C\varepsilon_{t+1} \tag{1.7}$$

where $\varepsilon_{t+1} \sim N(0, I)$. These matrices are specified in detail in section 1.4.2.

Each period, agents observe private signals which are a linear combination of x_t and an idiosyncratic noise component:

$$x_t^j = Sx_t + Q\eta_t^j \tag{1.8}$$

where $\eta_t^j \sim N(0, I)$ is assumed to be independent across agents. For tractability, and in keeping with most of the dispersed information literature, I assume signal precision is the same across all agents, fixed at all times, and common knowledge.

By the pricing equation (1.4), bond prices will be related to stochastic discount factors, which themselves are a function of the fundamentals x_t. Future stochastic discount factors will be a function of (future) fundamentals. Combined with the fact that bond prices *today* are functions of higher order expectations about stochastic discount factors, the relevant state vector will be the

[7] Barillas and Nimark (2015) derive more implications of this result, such as the fact that the portion of individuals' expected excess returns due to differences in belief from the cross-sectional average must be orthogonal to public information.

hierarchy of average higher order expectations about fundamentals. pth order average expectations are defined recursively as

$$x_t^{(p)} \equiv \int E\left[x_t^{(p-1)}|\Omega_t^j\right] dj$$

and the hierarchy of average order expectations is collected in the vector X_t:

$$X_t \equiv \begin{bmatrix} x_t \\ x_t^1 \\ \vdots \\ x_t^{(p)} \\ \vdots \\ x_t^{(\bar{k})} \end{bmatrix}$$

where $\bar{k}$ is the maximum level of higher order expectation considered (Nimark (2007)).[8]

Conjecture and verify the price of a bond takes the form

$$p_t^n = A_n + B_n' X_t + \nu_t^n \tag{1.9}$$

where ν_t^n is a maturity-specific shock, which is *i.i.d.* across time and maturities. Further conjecture that X_t follows a VAR(1)

$$X_{t+1} = \mu_X + \mathcal{F} X_t + \mathcal{C} u_{t+1} \tag{1.10}$$

where u_{t+1} contains all aggregate shocks - the shocks to fundamentals ϵ_t and the vector of price shocks ν_t. X_t is a $(d \times (\bar{k}+1)) \times 1$ vector.

(Log) yields at time t of a zero coupon bond maturing in n periods are defined as $-\frac{1}{n}p_t^n$ where p_t^n is the log price of the bond. Assume bonds up to a finite maturity $\bar{n}$ are traded.[9] Collect yields in a vector y_t:

$$y_t \equiv \begin{bmatrix} -\frac{1}{2}p_t^2 & \cdots & -\frac{1}{n}p_t^n & \cdots & -\frac{1}{\bar{n}}p_t^{\bar{n}} \end{bmatrix}$$

[8] I set $\bar{k} = 15$. The majority of the weight in bond prices is on the first few orders of expectation; raising the order increases the computational burden substantially without improving the fit of the model.

[9] Hilscher et al. (2014) document that the vast majority of Treasury debt currently held by the public has maturity of less than ten years. I set $\bar{n} = 40$, i.e., 10 years is the maximum traded by agents or used to form forecasts.

Assume agents' information sets Ω_t^j include the history of their private signals x_t^j, the short rate r_t and a vector of bond yields out to maturity $\bar{n}$:

$$\Omega_t^j = \left\{x_t^j, r_t, y_t, \Omega_{t-1}^j\right\} \tag{1.11}$$

Having conjectured an affine form for bond prices and exogenous information, the signals that agents observe will be an affine function of the state. The filtering problem of an atomistic agent j has the following state-space representation:

$$\begin{aligned} X_{t+1} &= \quad \mu_X + \mathcal{F} X_t + \mathcal{C} u_{t+1} \\ \underbrace{\begin{bmatrix} x_t^j \\ r_t \\ y_t \end{bmatrix}}_{z_t^j} &= \quad \mu_z + D X_t + R \begin{bmatrix} u_t \\ \eta_t^j \end{bmatrix} \end{aligned} \tag{1.12}$$

I assume agents use the Kalman filter to form estimates of the state X_t. In a linear, Gaussian setting, the Kalman filter is equivalent to recursive Bayesian updating, so this amounts to assuming that agents use Bayes' rule to update their predictions (Harvey (1989)). I further assume agents have observed an infinitely long history of signals, so they use the steady state Kalman filter to make their predictions. This standard assumption avoids the need to keep track of individual signal histories. The matrices $\mathcal{F}, \mathcal{C}$ determine how higher order expectations evolve, which depends on the individual filtering problem of traders and the equilibrium expressions for prices. At the same time, prices depend on the evolution of (higher order) expectations, which depends on the state transition matrices $\mathcal{F}, \mathcal{C}$. Hence, we first take the bond price equations as given to derive the law of motion, and then show the law of motion is consistent with our conjecture for prices.

The details of the bond trader's Kalman filtering problem are in appendix 1.B.2. The filtering problem, aggregating across traders, implies a fixed point expression for $\mathcal{F}$ and $\mathcal{C}$ (appendix equation (1.42)).

SDFs and bond prices. To derive an expression for prices, I need to explicitly model the stochastic discount factor of bond traders. As is common in the Gaussian affine term structure literature, I assume stochastic discount factors are "essentially affine" as in Duffee (2002). The log stochastic discount factor is assumed to take the form:

$$m_{t+1}^{j} = -r_t - \frac{1}{2}\Lambda_t^{j\prime}\Sigma_a\Lambda_t^{j} - \Lambda_t^{j\prime}a_{t+1}^{j} \tag{1.13}$$

In the above expression, $\Lambda_t^{j\prime}$ are (time-varying) market prices of risks to holding bonds, and a_{t+1}^{j} is the vector of one-period-ahead bond price forecast errors, which have unconditional covariance matrix Σ_a.

$$a_{t+1}^{j} \equiv \begin{bmatrix} p_{t+1}^{1} - E_t^{j}[p_{t+1}^{1}] \\ \vdots \\ p_{t+1}^{\bar{n}-1} - E_t^{j}[p_{t+1}^{\bar{n}-1}] \end{bmatrix} \tag{1.14}$$

These errors occur because of shocks that were unanticipated by agents. Hence, the vector of forecast errors span the risks that agents must be compensated for.

Assume prices of risk Λ_t^{j} are an affine function of X_t^{j} and the vector of maturity shocks:

$$\Lambda_t^{j} = \Lambda_0 + \Lambda_x X_t^{j} + \Lambda_\nu E[\nu_t|\Omega_t^{j}] \tag{1.15}$$

where X_t^{j} is are trader j's expectations (from 0 to $\bar{k}$) of the latent factors

$$X_t^{j} \equiv \begin{bmatrix} x_t^{j} \\ E_t^{j}[x_t|\Omega_t^{j}] \\ \vdots \\ E_t^{j}[x_t^{(\bar{k})}|\Omega_t^{j}] \end{bmatrix} \tag{1.16}$$

As mentioned above, the prices of risk represent additional compensation required for traders to be willing to hold an additional unit of each type of risk. In the absence of Λ_x and Λ_v, that compensation would not vary over time and risk premia would be constant. If $\Lambda_t^{j} = \mathbf{0}$, agents would be risk-neutral.

Given the conjectured bond price equation (1.9):

$$p_t^{n} = A_n + B_n' X_t + \nu_t^{n}$$

Appendix 1.B shows how to arrive at the following recursive representation of bond prices.

$$
\begin{aligned}
A_{n+1} = & \quad -\delta_0 + A_n + B_n\mu_X + \frac{1}{2}e_n'\Sigma_a e_n - e_n'\Sigma_a\Lambda_0 && (1.17) \\
B_{n+1}' = & \quad -\delta_X + B_n'\mathcal{F}H - e_n'\Sigma_a\widehat{\Lambda}_x && (1.18)
\end{aligned}
$$

with

$$
\begin{aligned}
A_1 = & \quad -\delta_0 && (1.19) \\
B_1 = & \quad -\delta_X' && (1.20)
\end{aligned}
$$

The price of a one-period bond, $p_t^1 = -\delta_0 + [\delta_x, \mathbf{0}]X_t = -r_t$. H is a matrix that selects only higher order expectations terms.[10] e_n' is a selection vector that has 1 in the n^{th} position and zeros elsewhere. $\widehat{\Lambda}_X$ is a normalization of Λ_X.

For comparison, I also estimate a full information model, which is essentially that of Ireland (2015). There are no maturity shocks in the baseline model. Equations (1.17) and (1.18) are replaced by

$$
\begin{aligned}
A_{n+1} &= -\delta_0 + A_n + B_n\mu_X - B_n\lambda_0 C + \frac{1}{2}B_n'CC'B_n && (1.21) \\
B_{n+1} &= -\delta_x + B_nF^P - B_nC\lambda_x && (1.22)
\end{aligned}
$$

1.4.2 The macroeconomic environment and prices of risk

This section outlines the evolution of the factors x_t that are sources of priced risk in the empirical model. The parameters of the VAR for the factor dynamics are restricted to allow for structural interpretations of the shocks, ensure the model is identified, and to constrain the estimation to economically relevant areas of the parameter space. The assumptions I make are similar to that of Ireland (2015). Ireland's model has several features that are desirable from an empirical and computational point of view. It is relatively parsimonious, making it straightforward to extend to

[10]More specifically, H is a matrix operator that replaces nth order expectations with $n+1$-th order expectations and annihilates any orders of expectation greater than $\bar{k}$. This is equivalent to writing prices in terms of a (hypothetical) agent whose SDF is equal to the average.

the case of higher order expectations. At the same time, it is rich enough to capture salient features of both the macroeconomy and yields, and thus is useful for understanding the role of information frictions in a more structural macroeconomic model.

Macroeconomic dynamics

Assume short term rates are managed by a central bank that sets an exogenous, time varying, long run inflation target τ_t and then picks a short rate r_t to manage an interest rate gap $g_t^r = r_t - \tau_t$. Define the deviation of inflation from its long run target $g_t^\pi = \pi_t - \tau_t$. Then the evolution of the interest rate "gap" takes the form of a Taylor-type reaction function:

$$g_t^r - g^r = \phi_r(g_{t-1}^r - g^r) + (1 - \phi_r)\left(\phi_\pi g_t^\pi + \phi_y(g_t^y - g^y) + \phi_v v_t\right) + \sigma_r \varepsilon_{rt} \tag{1.23}$$

In this expression, g_t^y is the output gap (defined as the log difference between real GDP and its trend).[11] The financial risk factor v_t shifts prices of risk Λ_t^j in a manner specified below. I will assume that all time variation in prices of risk comes through movement in this factor, which is consistent with the empirical results in Cochrane and Piazzesi (2008), Dewachter et al. (2014), and Bauer (2016) who all find that a single factor is responsible for nearly all time variation in bond risk premia.[12] As in Ireland (2015), I treat this factor as latent during the estimation. Including it in the Taylor rule is a simple way to incorporate contemporaneous feedback between financial conditions and the central banks' interest rate's policy stance. I impose prior restrictions on these parameters. First, I assume that ϕ_v is non-negative. While in principle unnecessary to identify the model's key parameters, this restriction is consistent with the idea that the central bank has raised rates in response to an increase in risk premia.[13] Second, I assume ϕ_r, which governs the degree of

[11] I detrend log real GDP using the Hodrick-Prescott filter with a smoothing parameter of 16000, which amounts to assuming an extremely slow-moving trend in output.

[12] Notably these authors arrive at this conclusion from different methods. Cochrane and Piazzesi (2008) show that a single "tent shaped" factor extracted from the yield curve explains nearly all time variation in term premia. Dewachter et al. (2014)'s risk factor is identified by a similar assumption to that of Ireland (2015) and is highly correlated with the Cochrane-Piazzesi factor. Bauer (2016) use Bayesian methods to estimate a Gaussian term structure model and finds evidence for strong zero restrictions which imply only changes in the "slope" factor affect term premia.

[13] McCallum (2005) suggests a Taylor rule with smoothing and a reaction to the term spread - itself affected by a possibly time varying term premium - is consistent with a negative slope coefficient in Campbell and Shiller (1991)

interest rate smoothing, is restricted to fall between zero and 1. Finally, I assume ϕ_π and ϕ_y are both positive.

The long-run inflation target is assumed to follow a stationary $AR(1)$ process:[14]

$$\tau_t = (1-\rho_{\tau\tau})\tau + \rho_{\tau\tau}\tau_{t-1} + \sigma_\tau \varepsilon_{\tau t} \tag{1.24}$$

with $\rho_{\tau\tau} \in (0,1)$.

For the remaining equations, I follow Ireland (2015) in imposing restrictions similar to the structural VAR literature. Namely, I impose impact restrictions that shocks to monetary policy (ε_t^r) do not affect the inflation or output gaps on impact, and shocks to the output gap do not affect inflation immediately (implicitly assuming that prices are initially sticky in response to "real" shocks). Moreover, I assume shocks to τ_t only affect the output and inflation gaps on impact; otherwise the effect of changes in the inflation target are limited to its changes to the interest rate and inflation gaps (and their attendant effects on output). Hence, lags of τ_t do not appear separately in the other equations, consistent with a notion of long-run monetary neutrality. Finally, I assume that the latent risk factor v_t can affect the inflation and output gaps (that is, allowing for linkages between the financial system and the macroeconomy), and that shocks to other variables also affect v_t on impact. However, I assume that v_t is stationary and only affected by its own lags. To summarize:

$$\begin{bmatrix} g_t^\pi \\ g_t^y \\ v_t \end{bmatrix} = \begin{bmatrix} -\rho_{\pi r} g_r - \rho_{\pi y} g^y \\ g_y - \rho_{yr} g^r - \rho_{yy} g^y \\ 0 \end{bmatrix} + \begin{bmatrix} \rho_{\pi r} & \rho_{\pi\pi} & \rho_{\pi y} & 0 & \rho_{\pi v} \\ \rho_{yr} & \rho_{y\pi} & \rho_{yy} & 0 & \rho_{yv} \\ 0 & 0 & 0 & 0 & \rho_{vv} \end{bmatrix} \begin{bmatrix} g_t^r \\ g_t^\pi \\ g_t^y \\ \tau_t \\ v_t \end{bmatrix} + \begin{bmatrix} 0 & \sigma_\pi & 0 & \sigma_{\pi\tau}\sigma_\tau & 0 \\ 0 & \sigma_{y\pi}\sigma_\pi & \sigma_y & \sigma_{y\tau}\sigma_\tau & 0 \\ \sigma_{vr} & \sigma_{v\pi} & \sigma_{vy} & \sigma_{v\tau} & \sigma_v \end{bmatrix} \begin{bmatrix} \varepsilon_{rt} \\ \varepsilon_{\pi t} \\ \varepsilon_{yt} \\ \varepsilon_{\tau t} \\ \varepsilon_{vt} \end{bmatrix} \tag{1.25}$$

Collecting the factors in x_t:

regressions.

[14]Stationarity is assumed for two, related, technical reasons. The first is that as Ireland (2015) notes, interest rate processes that contain a unit root will leave long-run yields undefined. The second, related issue, is the stationarity of asset prices helps ensure that approximation error caused by truncating $\bar{k}$ can be made arbitrarily small (Nimark (2007)).

$$x_t = \begin{bmatrix} g_t^r \\ g_t^\pi \\ g_t^y \\ \tau_t \\ v_t \end{bmatrix} \tag{1.26}$$

they can be written in matrix form:

$$P_0 x_t = \mu_x + P_1 x_{t-1} + \Sigma_0 \varepsilon_t \tag{1.27}$$

Exact expressions for $P_0, \mu_x, P_1, \Sigma_0$ are shown in appendix 1.B.4. Left multiplying by P_0^{-1} yields (1.7). After a normalization of one shock covariance matrix parameter, the VAR is exactly identified. Like Ireland, I set $\sigma_v = 0.01$.

Restrictions on prices of risk. The matrices governing the mapping of factors into prices of risk shown in (1.13) and (1.15) are high-dimensional. Without additional restrictions, it is unlikely that these matrices would be identified. Moreover, as Bauer (2016) notes, absent restrictions on the prices of risk, the estimation does not take into account cross-sectional information in the yield curve. Moreover, Bauer's results suggest that the data prefer strong zero restrictions (albeit in a setting where yields are driven by the "level, slope curvature" factors rather than macroeconomic shocks). Accordingly, I incorporate two sets of restrictions. First, I follow Ireland (2015) in imposing that, under full information, changes in prices of risk are driven by entirely by changes in v_t, and that v_t is not itself a source of priced risk.[15] Second, I follow Barillas and Nimark (2015) in restricting Λ_t^j to nest a full information version of the model without maturity shocks. This means that the same number of parameters govern prices of risk in the full and dispersed information models.[16] Details of these restrictions are shown in appendix 1.B.5.

[15]This is consistent with the notion that movements in a single variable explain time variation in risk premia, and that variable is related to the state of the macroeconomy.

[16]I also assume the maturity specific shocks shocks have the the same standard deviation across yields, although the shocks to each yield are independent.

1.4.3 Signals

The last step is to specify agents' agents' idiosyncratic signal structure. I do not formally model the information choice of traders as in the rational inattention literature, but impose an exogenous information structure. I assume prices are observed without error, but individuals' observations of the non-price factors driving prices of risk are subject to idiosyncratic noise that is uncorrelated across variables. Recalling (1.8), I assume bond traders observe the short rate and separate signals about inflation and the long-run inflation target. To summarize:

$$x_t^j = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} x_t + \begin{bmatrix} 0 & 0 & 0 & 0 \\ \widetilde{\sigma}_\pi & 0 & 0 & 0 \\ 0 & \widetilde{\sigma}_y & 0 & 0 \\ 0 & 0 & \widetilde{\sigma}_\tau & 0 \\ 0 & 0 & 0 & \widetilde{\sigma}_v \end{bmatrix} \begin{bmatrix} \widetilde{e_t}^\pi \\ \widetilde{e_t}^y \\ \widetilde{e_t}^\tau \\ \widetilde{e_t}^v \end{bmatrix} \tag{1.28}$$

Exogenous information keeps the model tractable enough to allow for likelihood based estimation. The downside is vulnerability to a Lucas-critique-like argument that the allocation of attention is not invariant to policy changes, and the model does not let the precision of signals vary over the business cycle, as it might in a model where agents optimally (re)allocate attention. The advantage is this allows estimation of the precision of traders' information that is consistent with asset price movements over the sample.

1.4.4 Bond price decompositions

Given the expression for prices and a model for inference, we can characterize what portion of bond yields are driven by higher-order beliefs - that is, the portion of yields driven directly by dispersed information. To do this, note that common knowledge of rationality and the VAR describing X_t implies two facts: (1) bond prices are pinned down by the *current* state and thus agents' forecasts of *future* states determine their forecasts of future bond prices, and (2) all information about future X_t is summarized in today's state (Barillas and Nimark (2015)). This implies two agents who agree about X_t today also have the same belief about X_{t+1}, X_{t+2}, etc, and hence prices at future dates.

Intuitively, the difference between actual prices and the price that would obtain if all agents counterfactually held the same beliefs is the direct contribution of dispersed information to the bond price.[17] Like Barillas and Nimark (2015), I use the wedge between the counterfactual price with common beliefs and actual prices to quantify the extent to which dispersed information about particular factors affects bond yields. Moreover, because the risks priced in bonds have a macroeconomic interpretation, the wedge can be decomposed in order to understand what set of higher-order beliefs are particularly important for determining yields at different maturities.

Following Barillas and Nimark (2015), we can define a matrix operator $\bar{H}$ that replaces all higher order expectations with first order expectations, that is:

$$\begin{bmatrix} x_t \\ x_t^{(1)} \\ \vdots \\ x_t^{(1)} \end{bmatrix} = \bar{H} \begin{bmatrix} x_t \\ x_t^{(1)} \\ \vdots \\ x_t^{(\bar{k})} \end{bmatrix} \tag{1.29}$$

The price that would obtain if all higher order expectations coincided with the first order expectation - the "counterfactual consensus price" - is

$$\bar{p}_t^n = A_n + B_n' \bar{H} X_t + \nu_t^n \tag{1.30}$$

We can use this to decompose prices into the component that depends on first order "average" expectations and the component that depends on dispersion of information and the resulting divergence of expectations about expectations. The wedge can be written as:

$$p_t^n - \bar{p}_t^n = B_n' X_t - B_n' \bar{H} X_t = B_n'(I - \bar{H}) X_t \tag{1.31}$$

The counterfactual consensus price, which contains only the effect of average expectations in yields, can be decomposed into short rate expectations and "classical" risk premia - that is, the part of yields that depends on first-order average beliefs net of average rate expectations.

[17] Allen et al. (2006) show in a similar setting how prices of long-lived assets will not generally reflect average expectations when there is private information. Barillas and Nimark (2015) refer to the difference between actual prices and the counterfactual consensus price as the "speculative component"; Bacchetta and Van Wincoop (2006) refer to it as the "higher order wedge." The preferred interpretation of Bacchetta and Van Wincoop is that it is the present value of deviations of higher-order beliefs from first-order beliefs.

$$p_t^n = A_n^{\text{prem}} + B_n^{\text{prem}\prime} X_t + A_n^{\text{rate}} + B_n^{\text{rate}\prime} X_t + \underbrace{B_n'(I - \bar{H})X_t}_{\text{higher order wedge}} + \nu_t^n \tag{1.32}$$

Where $A_n^{\text{prem}} = A_n - A_n^{\text{rate}}, B_n^{\text{prem}\prime} = B_n'\bar{H} - B_n^{\text{rate}\prime}$. To make this decomposition operative, we need the model-implied future expected short rates. For the hypothetical average agent,

$$E_{t|t} r_{t+1} = -\delta_0 - \delta_X H X_{t+1|t} = -\delta_0 - \delta_X(\mu_X + \mathcal{F} H X_t)$$

and so on for further ahead future short rates:

$$\begin{aligned} A_n^{\text{rate}} &= -\, n(\delta_0 + \delta_X \mu_X) - \delta_X \sum_{s=0}^{n-1} \mathcal{F}^s \mu_X \\ B_n^{\text{rate}\prime} &= -\, \delta_X \sum_{s=0}^{n-1} \mathcal{F}^s H \end{aligned} \tag{1.33}$$

The decomposition of the wedge is a straightforward selection of different elements. For example, the portion of the higher-order wedge attributable to higher-order beliefs about the long-run inflation target τ_t is

$$\begin{aligned} &B_n'(I - \bar{H})X_t^\tau \equiv \\ &\qquad B_n'(I - \bar{H}) \cdot \operatorname{diag}\begin{bmatrix} 0 & 0 & 0 & 1 & 0 & \cdots & 0 & 0 & 0 & 1 & 0 \end{bmatrix} X_t \end{aligned} \tag{1.34}$$

Note that this depends on both the level of the (higher order) expectations (i.e., $\tau_t^{(2)}, \tau_t^{(3)}$) and so on), and the appropriate elements of B_n'.

1.5 Solution and estimation

1.5.1 Solution

The solution to the model is a fixed point of the bond pricing terms A_n, B_n and agents' beliefs, summarized in the law of motion and covariance of the vector of higher-order expectations. In particular, we need to find a fixed point between the price recursions, (1.17) and (1.18), the mean-square error matrix for state forecast error ((1.40) in the appendix), and the law of motion for the hierarchy of average higher-order expectations ((1.42) in the appendix). The precise numerical procedure for finding a fixed point is detailed in appendix 1.C.

1.5.2 Econometric model and data

The model period is a quarter and the estimation runs from Q4:1971-Q4:2007. The end date is chosen prior to the zero lower bound period because the linear model does not respect the ZLB constraint. I take data on (non-annualized) zero coupon yields from the yield curve estimates in Gurkaynak et al. (2007), averaged over the quarter. In the econometric model, I use the the short rate (assumed to be the Federal Funds Rate, as in Piazzesi et al. (2013)), and rates on 1,2,3,4, 5 year and 10 year bonds.[18]

To identify agents' beliefs and the macroeconomic dynamics, I use data on the output gap (calculated as the log difference between real GDP and its HP filtered trend using a smoothing parameter of 16,000)), inflation and inflation forecasts (as measured by log changes in the GDP deflator), and treat τ_t and v_t as latent process.[19] I treat individual forecasts as (noisy) observations of the average expectation of financial market participants, and use the cross-sectional variance of forecasts to estimate the cross-sectional dispersion of signals.

Individual survey responses are treated as a noisy indicator of the average expectation, where the extent of the noise is pinned down by the model-implied cross-section of expectation around the first-order average expectation. This matrix can be calculated using the Kalman filtering problem of individual agents (see appendix 1.B.2). Because the number of respondents to the SPF has varied over time, the number of observables at different times is time varying. This is not difficult to incorporate into the Kalman filter used to calculate the log-likelihood of the data. Assuming there are m_t^1 respondents to the 1-period ahead question and m_t^4 to the four-period ahead question in the

[18]Note that I assume agents in the model observe the whole yield curve, not just this subset.

[19]I use the cross-section of one and four quarter ahead forecasts from the Survey of Professional Forecasters in the estimation. The advantage of inflation forecasts is that they are available for the entire sample period with a relatively high response rate. Moreover, inflation forecasts in the SPF are quite accurate on average, which means my choice of data does not automatically favor sizable information frictions.

SPF at time t, the state space system for estimation is

$$\begin{aligned} X_t &= \mu_X + \mathcal{F}X_{t-1} + [\mathcal{C}, \mathbf{0}_{d(\bar{k}+1)\times m_t^1+m_t^4}]\bar{u}_t \\ \bar{u}_t \sim N(0,1) &\text{ with dimension } (d + (\bar{n}-1) + m_t^1 + m_t^4) \times 1 \\ \bar{z}_t &= \mu_{\bar{z},t} + \bar{D}_t X_t + \bar{R}_t \bar{u}_t \end{aligned} \tag{1.35}$$

where in particular $\mu_{\bar{z},t}, \bar{D}_t$ and $\bar{R}_t$ vary in size to account for missing observations. The matrices are reported in appendix 1.D.

For the full information version of the model, I treat forecasts as if they are observations of the rational expectations forecast with *i.i.d.* error. I allow each forecast horizon to have a different error variance. I also treat each individual bond yield as if it were observed with maturity-specific econometric error. Conceptually, these errors are distinct from the errors in the dispersed information version. In the dispersed information model, the "noise" in forecasts is pinned down by the model-consistent state mean square error matrix. As discussed earlier, the maturity specific shocks are a risk faced by traders in the model, rather than being econometric noise in the empirical model.

I estimate the model via Bayesian methods. The advantage of Bayesian methods in this context are the use of prior information to constrain parameters in line with economic theory (in a transparent way), and the characterization of posterior credible sets to illustrate the distributions of parameters. In particular, I use a Metropolis-Hastings Markov Chain Monte Carlo procedure to estimate the model parameters.

Because the model has a large number of parameters and is computationally burdensome to solve, I use somewhat informative priors on macroeconomic variables to focus on reasonable areas of the parameter space. In particular, as noted earlier, I restrict $\phi_v \geq 0$. I also place some informative prior restrictions on VAR parameters. I impose that ρ_{yv} is non-positive, which implies that all else equal, greater risk premia are contractionary. This is consistent with most general equilibrium models with financial frictions. For similar reasons, I impose a slightly informative prior that for ρ_{yr} that is centered around -1, while still allowing the estimation to explore regions of the parameter space where this restriction does not hold. These restrictions, while technically unnecessary for identification, help discipline the estimation. Finally, I follow Ireland in imposing that $\rho_{\tau\tau} = 0.999$. The prior distributions are reported in appendix 1.F.

I follow Ireland (2015) in imposing that $\lambda_\pi^x < 0$ to identify the way changes in v_t affect risk premia, and that that long-run bond prices are well defined by assuming that the parameters governing the physical and risk-neutral dynamics of bonds are stationary under full information. This implies only accepting parameter draws such that the maximum eigenvalues of F^P and $F^P - C\lambda_x$ are less than one in modulus.

I run separate MCMC chains in parallel for each model. For the full information model, each chain is of length 400,000; I discard the first 10% of each chain and subsequently analyze every 1000th draw. The DI model is much more computationally intensive; the results reported here are based on 5 chains of length $23,000$ each. I drop the first 50% of each chain (because it takes longer to stabilize) and use every 100th draw.

1.6 Parameter estimates and impulse responses

Here I report the results of the estimation for the dispersed information model. Analogous figures for the full information model are shown in appendix 1.H.

Parameter estimates

Parameter estimates across chains, and posterior credible sets are reported in table 1.1.

In terms of the macroeconomic dynamics, the full and dispersed information models have relatively similar estimates. This is, perhaps, unsurprising, as the model does not allow for direct feedback from the inference problem of agents to the macroeconomy and the VAR is exactly identified without reference to bond yields. Most of the macro-VAR parameter estimates are basically in line with the results in Ireland (2015). My estimates of the prices of risk in both models differ from those in Ireland (2015). While some of this is likely attributable to differences in samples, the full information estimates of those parameters have a high degree of dispersion, as do the parameters governing the covariance of the non-financial factors with the risk factor v. Taking the full and dispersed information parameters together, it appears that it is difficult to separately identify the prices of risk terms, and the covariances that govern changes of risk.

The key parameters of interest are the parameters that govern the informational quality of agents (the bottom five rows). The relatively small value for σ_ν indicates that prices are not especially noisy. This implies, all else equal, that prices move mostly due to (higher order beliefs about) fundamentals rather than large direct shocks to prices. Agents' noisy signals about fundamentals are the last four rows of the table.

Although there is some imprecision in the estimates, agents' idiosyncratic noise has a fairly large standard deviation. However, this does not necessarily imply that agents' beliefs are inaccurate or dispersed, because agents understand the structure of the economy. For example, traders know an unanticipated increase in inflation is correlated with unanticipated increases in output ($\sigma_{y\pi} > 0$), and that higher inflation today usually depresses growth in the future ($\rho_{y\pi} < 0$). Moreover, agents are allowed to learn from prices, which aggregate the idiosyncratic beliefs of different agents. Hence, we can infer that the quality of agents' idiosyncratic signals is somewhat poor, but we cannot conclude simply from the parameter estimates that agents have incorrect beliefs. All else equal, noisier private signals will simply receive less weight.

Impulse responses

To demonstrate some of the (informational) mechanisms at play in the model, I plot impulse responses for fundamentals, the first three orders of expectation about fundamentals, inflation forecasts, and prices for a subset of the model shocks. The impulse responses discussed in this section are shown for the posterior mode for clarity. The complete set of impulse responses, including posterior credible sets, shown in appendix 1.G.1. Analogous full information impulse responses are shown in

	Mode	Mean	Median	5%	95%	Std.
ϕ_r	0.5357	0.5339	0.5355	0.5324	0.5370	0.0224
ϕ_π	0.1835	0.1851	0.1869	0.1768	0.1908	0.0090
ϕ_y	0.1016	0.1010	0.1010	0.0928	0.1101	0.0066
ϕ_v	0.0349	0.0343	0.0348	0.0305	0.0365	0.0024
σ_r	0.0029	0.0033	0.0032	0.0029	0.0038	0.0003
ρ_{yr}	-0.9329	-0.9954	-0.9994	-1.0440	-0.9312	0.0539
$\rho_{y\pi}$	-0.3871	-0.4088	-0.4134	-0.4361	-0.3730	0.0254
ρ_{yy}	0.9164	0.8988	0.9031	0.8794	0.9136	0.0389
ρ_{yv}	-0.0007	-0.0008	-0.0006	-0.0022	-0.0000	0.0007
$\sigma_{y\pi}$	0.3329	0.2873	0.2711	0.2443	0.3355	0.0362
$\sigma_{y\tau}$	2.6888	2.6788	2.6878	2.6491	2.7163	0.1140
σ_y	0.0062	0.0068	0.0068	0.0063	0.0071	0.0004
σ_τ	0.0008	0.0009	0.0009	0.0008	0.0010	0.0001
$\rho_{\pi r}$	0.9229	0.8949	0.8882	0.8650	0.9323	0.0453
$\rho_{\pi\pi}$	0.4331	0.4368	0.4372	0.4223	0.4564	0.0208
$\rho_{\pi y}$	-0.1998	-0.1870	-0.1872	-0.2027	-0.1724	0.0117
$\rho_{\pi v}$	-0.1139	-0.1105	-0.1114	-0.1143	-0.1062	0.0055
$\sigma_{\pi\tau}$	-0.1380	-0.1483	-0.1452	-0.1993	-0.1123	0.0309
σ_π	0.0031	0.0034	0.0033	0.0031	0.0037	0.0002
ρ_{vv}	0.8633	0.8600	0.8613	0.8589	0.8640	0.0360
σ_{vr}	9.7831	9.8318	9.8834	9.7764	9.8987	0.4138
$\sigma_{v\pi}$	2.1485	2.1574	2.1672	2.1123	2.1919	0.0932
σ_{vy}	-2.1067	-2.1144	-2.1111	-2.1939	-2.0638	0.0972
$\sigma_{v\tau}$	0.8502	0.8581	0.8577	0.8215	0.9125	0.0466
λ_r	1.3216	1.0637	1.1132	0.7289	1.3235	0.1872
λ_π	-4.6156	-4.6789	-4.6655	-4.9244	-4.4789	0.2289
λ_y	0.0223	-0.5522	-0.5930	-1.0610	-0.0260	0.3227
λ_τ	-0.1725	-0.2202	-0.1318	-0.7908	0.1247	0.2885
λ_r^x	18.0439	18.3595	18.5122	17.9321	18.8519	0.8392
λ_π^x	-76.2162	-76.1999	-76.2163	-76.6236	-76.1171	3.1893
λ_y^x	-17.4621	-17.4327	-17.4515	-17.9745	-17.0402	0.7816
λ_τ^x	0.0126	0.1011	0.0987	-0.0640	0.3192	0.1232
σ_ν	0.0025	0.0032	0.0031	0.0025	0.0041	0.0005
$\widetilde{\sigma}_\pi$	0.3418	0.4817	0.5082	0.3303	0.6348	0.1106
$\widetilde{\sigma}_y$	0.7511	0.6247	0.6539	0.4601	0.7715	0.1073
$\widetilde{\sigma}_\tau$	0.5445	0.5860	0.5825	0.4727	0.7032	0.0849
$\widetilde{\sigma}_v$	0.4116	0.4153	0.4175	0.3039	0.5054	0.0609

Table 1.1: Parameter estimates for dispersed information model

appendix 1.H.

The non-asset price impulse responses to one-standard deviation shock to the monetary policy rule are shown in figure 1.5. The top row displays the responses of the fundamental factors, while the subsequent rows show increasing higher-order beliefs about those variables (and inflation expectations in the far right column). For inflation and interest rates, the responses are in terms of annualized percentage points; the output gap is in percentage points, and "risk" is scaled up by 100. As expected, a shock to the short rate shock causes inflation and output to fall over the course of several years; output eventually recovers, while inflation remains depressed for some time. The impulse

Rate shock: Macroeconomic Responses

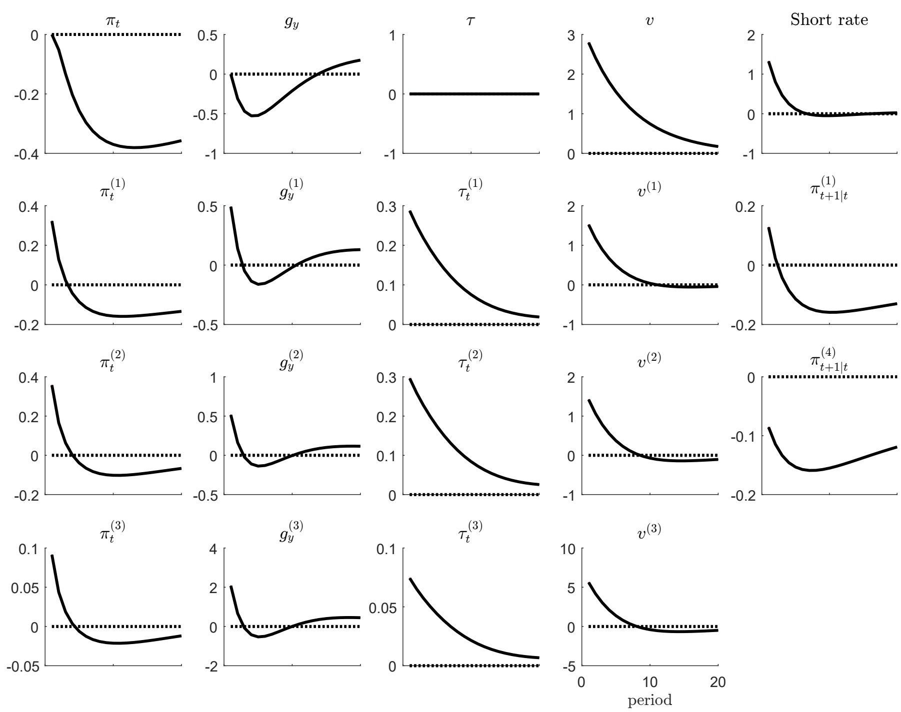


Figure 1.5: Response of non-financial variables to monetary policy rule shock, dispersed information

responses illustrate the identification problem faced by agents in the model. Agents observe that the short rate has risen. However, they do not know precisely why it increased; it might have been a shock to the rate directly, but also could be attributed to a change in inflation, the output gap, risk, or the inflation target. Because they are unable to discern the origin of the change, they place some posterior weight on the possibility that both inflation and the inflation target, are above average.[20] Hence, although inflation and the output gap have fallen, agents persistently mis-attribute the cause of the increase in short rates to changes in the inflation target. They further believe that others believe the inflation target has risen by about the same amount (second order expectations are similar to first order), but third order beliefs increase by less. This implies that traders' beliefs, in addition to being imperfect, are *dispersed* - on average, that others do not believe the same thing as them. Over time, as they observe the evolution of prices and their noisy signals about macroeconomic dynamics, their beliefs approach the "true" impulse responses (top row). Still, even two years after the shock has occurred, they believe that the inflation target is above its long-run level. It is worth emphasizing that agents are Bayesian learners and are doing the best they can; optimal inference in the model is characterized by mistaken beliefs about the origins of shocks and divergence of average beliefs with higher order beliefs.

Interestingly, despite not knowing the fundamental reason rates have risen, *dispersion* of beliefs after an interest rate shock does not have a large direct effect on yields. The response of yields are shown in figure 1.6. The overall response of yields to the shock are shown in the first row. Subsequent rows show the decomposition into rate expectations, "classical" risk premia, and the higher order wedge. Rate expectations (row 2) ,on average, are elevated as a result of the shock and those changes in (first-order) rate expectation explain nearly all of the increase in yields, even at the long end of the yield curve. In other words, while agents may not know *why* rates have increased, it is common belief that the path of short rates will be persistently elevated and the model attributes the increase in yields after a monetary policy shock to rate expectations. This is driven by beliefs

[20]The fact that inflation expectations rise after an identified monetary policy is consistent with the VAR results presented in Melosi (2017) for the response of average inflation expectations to a monetary policy shock identified with sign restrictions.

about the inflation target, which raises the expected path of short rates. Classical risk premia (row 3) and the higher order wedge (row 4) barely move as a result of the shock. The lack of direct

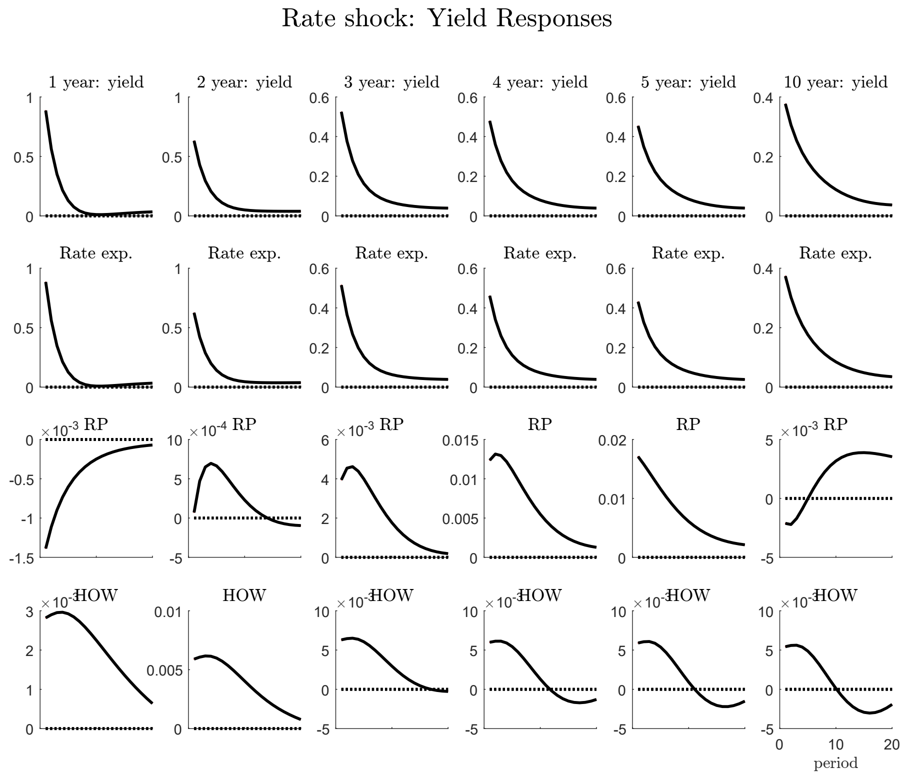


Figure 1.6: Response of financial variables to monetary policy rule shock, dispersed information

influence of higher order beliefs (at least in response to this shock) is reminiscent of Venkateswaran and Hellwig (2009) and Atolia and Chahrour (2013). In both these papers, firms have inaccurate beliefs (in that they incorrectly attribute the sources of the fluctuations they observe). However, in the general equilibrium settings in those papers, it turns out to not matter, in the sense that actions actions based on incorrect beliefs are virtually indistinguishable from ones when they had correct beliefs. In the context of this asset pricing model, however, the latter does not hold. The critical difference here is that rate expectations here are not the same as full information rational expectations. Because the path of expected short rates does not decline as slowly as it would under full information, long term yields rise more after a rate shock and stay elevated for a longer period of time. In other words, the inference problem of agents in and of itself matters for our assessment of financial fluctuations, even without appealing to the higher-order wedge, at least in the case of monetary shocks. As will be shown in section 1.8, this is true in general; once we account for the slow adjustment of expectations, yields are mostly attributed to the path of average short rate beliefs rather than "classical" risk premia or higher-order beliefs.

A similar set of impulse responses for a one standard deviation increase in the inflation target τ are shown in figures 1.7 and 1.8. Movements in the inflation target cause level shifts in the yield curve by persistently raising short rates.[21] Agents are slow to adjust the shock, so the level shift is gradual, rather than immediate, but the shock to the inflation target raises yields across the board by approximately the same amount over the course of several years. This is similar to the role it plays in the full information model, as shown in the appendix, and also to its role in Ireland (2015) and Gurkaynak et al. (2005).

What is interesting is the difference between fundamentals (the top line in figure 1.7) and higher order beliefs about those fundamentals, particularly the output gap. As in the full information model, a higher inflation target - essentially a more dovish policy stance, tolerating a higher rate of inflation in the long run - is associated with a temporary expansion in output. However, agents observing higher rates, accompanied by upward movements in inflation and risk, actually believe that output rises initially, falls over the medium term, and rises again. Higher order believes follow this pattern, although third-order expectations move more dramatically than first or second order beliefs.

"Risk shocks" (shocks to v) are shown in figures 1.9 and 1.10. The effect of risk shocks on macroeconomic variables and asset prices stand in contrast to the risk shocks in Ireland (2015). In his paper, the co-movements brought on by risk shocks are qualitatively similar to those of a

[21] Recall that the inflation target is the most persistent shock, with its autoregressive component calibrated to $\rho_{\tau\tau} = 0.999$.

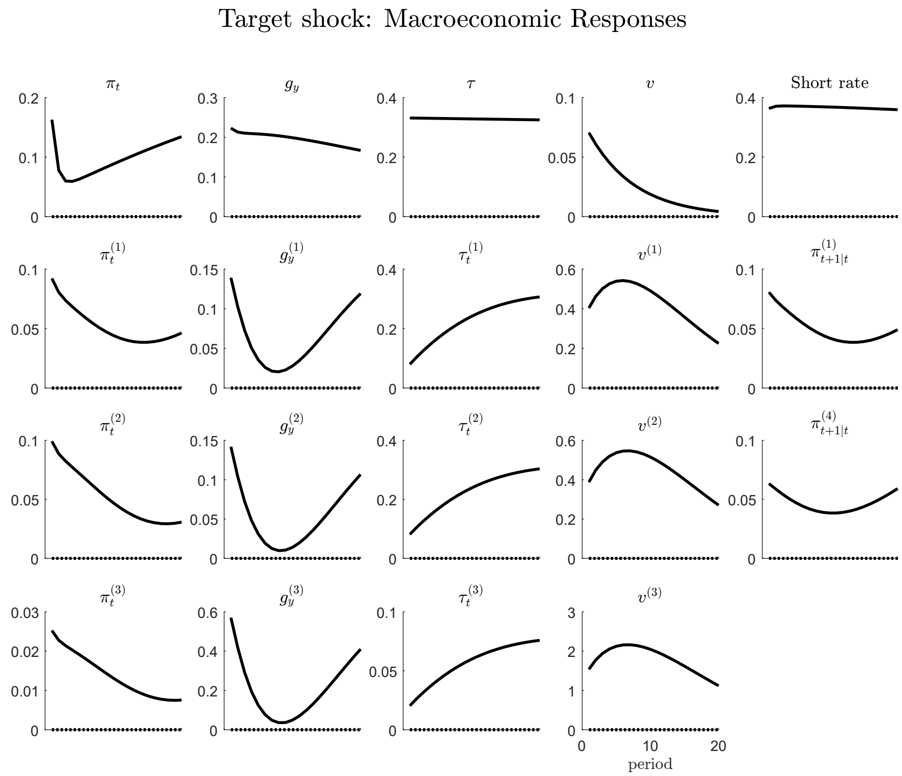


Figure 1.7: Response of non-financial variables to inflation target shock, dispersed information

monetary policy shock, albeit without a "price puzzle." By contrast, impulses to v_t in both the full- and dispersed-information versions of the model estimated here have nearly no direct effect on output, but depress inflation, and the reduction in inflation leads output to grow over time. Since this holds for both the full- and dispersed-information models, it is not a result of the information assumption.

Two possibilities for differences in the dynamic behavior for v_t between these results and those of Ireland (2015) present themselves. One is that this is driven by differences in sample, particularly, the difference in sample period.[22] The second possibility, alluded to earlier, is that the risk parameters

[22]One other sample difference between my full information results and Ireland's is that I include inflation forecasts. Removing those forecasts from the dataset does not qualitatively change the impulse responses at the posterior mode of the full information model.

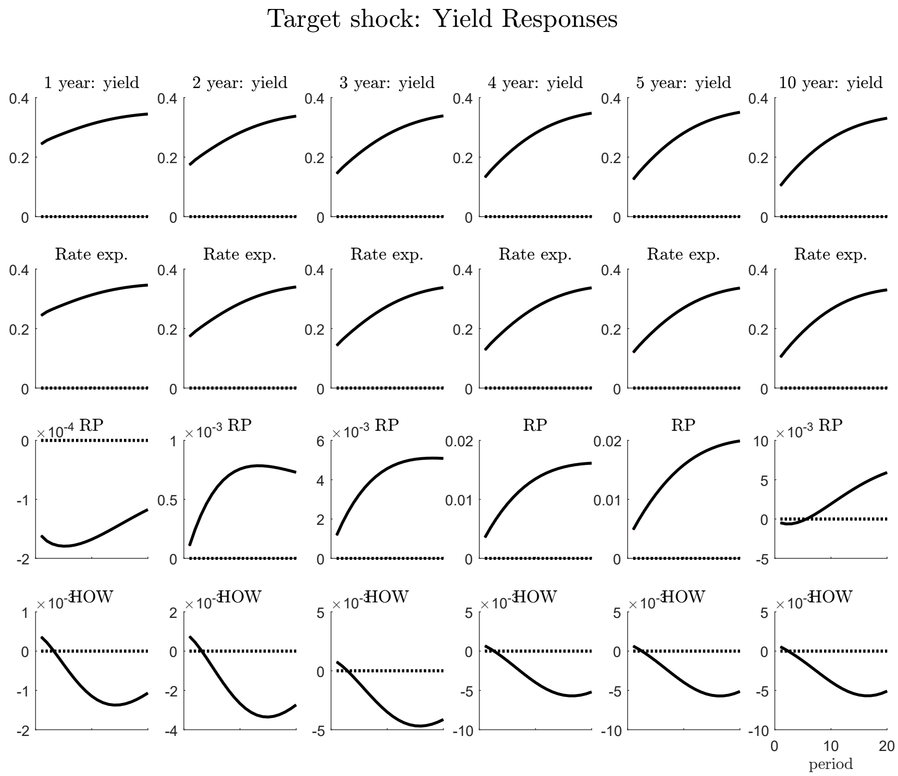


Figure 1.8: Response of financial variables to inflation target shock, dispersed information

and the prices of risk are not particularly well identified. Both of these explanations suggest that stronger prior information might "smooth out" the posterior and make the impulse responses to risk shocks more strongly resemble those of Ireland. In the absence of strong priors, these dynamics appear to be what the data prefers.

In terms of asset prices, the dynamics of Ireland's estimates and the full information estimates are more similar; risk shocks increase risk premia more than monetary policy shocks. But it is worth noting that the size of the change in risk premia is smaller between Ireland's estimates and the full information model. Once again, this is likely due to differences in sample composition and the imprecision of the estimates of parameters governing risk premia. More importantly, "risk premia" move much less in the dispersed information model, even following risk shocks. This is because

Risk shock: Macroeconomic Responses

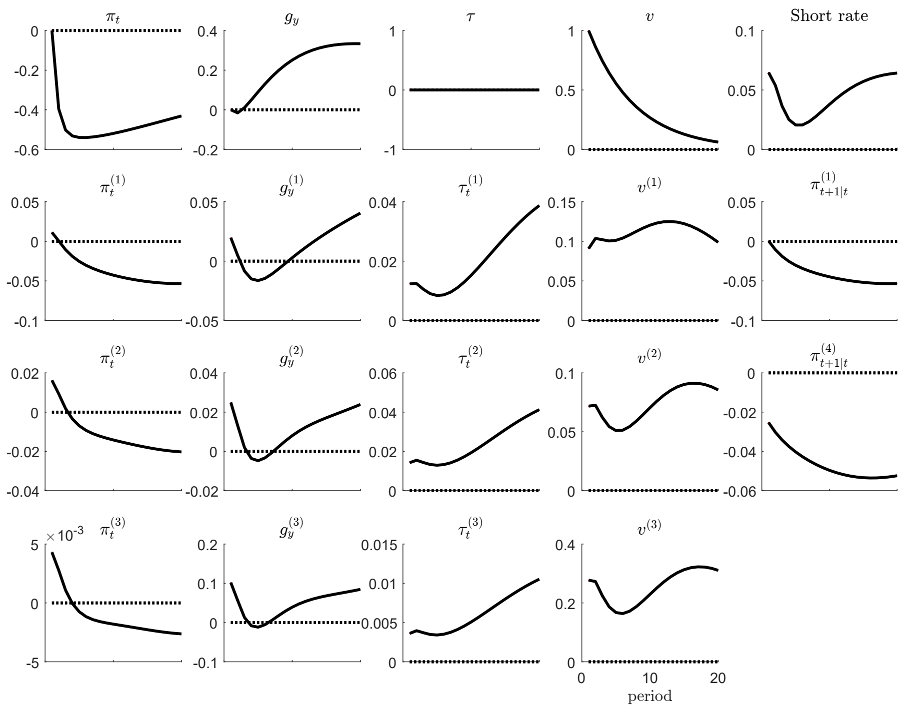


Figure 1.9: Response of non-financial variables to risk shock, dispersed information

risk premia in Ireland's model are the difference between yields and average rational expectations of short rates. From the point of view of the dispersed information model, the rational expectations risk premia conflates the difference between rational expectations and agents' actual average expectations, the effect of higher-order expectations, and "classical" risk premia, which are the part of the counterfactual consensus price unexplained by average rate expectations. Since the models differ quite a bit on what the right specification for average expected short rates is, the full information risk premia appears to be largely driven by the auxiliary assumption of full information rational expectations.

Other impulse responses are shown in appendix 1.G.1. The macroeconomic implications of impulse responses in both the full and dispersed information models are reasonable similar. Shocks to the

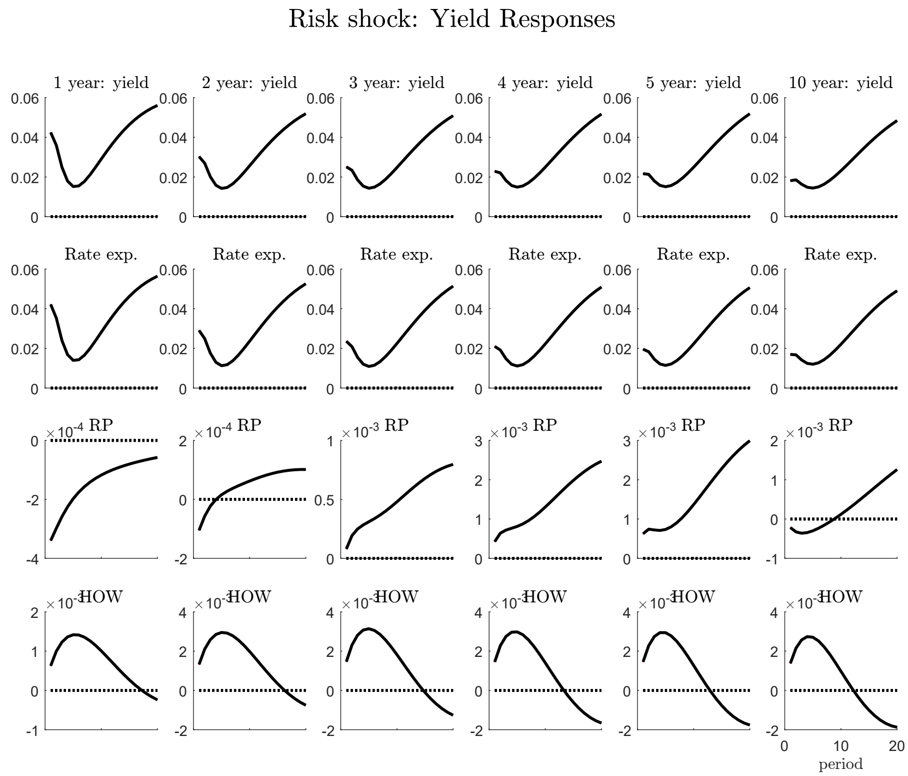


Figure 1.10: Response of financial variables to risk shock, dispersed information

output gap induce positive comovement in inflation, output, and interest rates - in a sense they are similar to demand shocks in DSGE models. Inflation shocks raise output on impact but lower it over the medium term. Apart from the initial (positive) change in output, they somewhat resemble cost-push shocks to the Phillips curve in New Keynesian models.

1.7 What do traders learn from?

Despite the abundance of commonly observed signals, the estimated results imply agents' information is imperfect and dispersed. In this section, I characterize the informativeness of agents' signals. As a preview, agents' private signals are sufficiently noisy that they are not especially informative about fundamentals. They learn about half of what they know about the output and inflation gaps from their private signals, and learn much less about policy or the financial risk factor. The rest of their information about fundamentals comes from prices. Moreover, the majority of traders' information about the beliefs of others also comes from observing prices. The most important price signal appears to be the policy rate of the central bank, which is also the yield on a bond that matures in one period. The short rate is informative about fundamentals, and because everyone knows that everyone learns about fundamentals from this particular signal, it is also informative about higher-order beliefs.

In most models of dispersed information, agents are assumed to learn only from idiosyncratic signals about fundamentals. While this simplification is justified by a desire for tractability, it is worth asking whether belief continues to be dispersed when agents have access to a wide range of price signals. Since agents' noisy signals are, on average, the true realization, it is possible that asset prices clean out idiosyncratic noise and agents are able to determine the true realization of fundamentals. Moreover, it is possible that yields are informative about higher order beliefs, which agents do not observe any direct signals about. On the other hand, it may be the case that prices are sufficiently noisy (due to maturity specific shocks or because they also reflect agents' idiosyncratic noise) that they are not particularly informative for agents. And, because prices reflect the beliefs of agents, it may be possible that yields do not contain any information that agents do not already know.

The approach I use to understand the informativeness of prices is drawn from information theory.[23] In particular, the posterior uncertainty of agents about particular variables (calculated during the agents' Kalman filtering problem) can be characterized in terms of "entropy," which can be thought of as the average number of binary signals needed to fully describe the outcome of a random variable. We can characterize how much agents learn from signals about a particular variable in terms of the reduction in entropy after observing those signals (see appendix 1.E for details).

[23] The entropy-based measure of signal informativeness I use is also used in the rational inattention literature initiated by Sims (2003) to describe the constraint on agents' information processing capacity.

Table 1.2: Reduction in uncertainty about fundamentals (columns) coming from observing a single signal (rows).

Signal(↓), fundamental →	π	g^y	τ	v
r_t	0.42	0.45	0.97	0.64
π^j	0.03	0.03	0.12	0.00
g^{y^j}	0.03	0.07	0.04	0.00
τ^j	0.00	0.00	0.09	0.00
v^j	0.37	0.33	0.03	0.24

Adapting a measure used in Melosi (2017), I examine how informed agents are after viewing a limited subset of signals *relative* to how informed they would be if I let them use the total set of signals outlined in section 1.4.3. In other words, I calculate how informed they are after observing all of their private signals and the yield curve. I then can calculate how informed they would have been under a "counterfactual" subset of signals. Since on average additional information must (weakly) reduce uncertainty, we can think of the reduction in entropy coming from the counterfactual subset of signals as the fraction of total information that could have come from that set; if the reduction in entropy were zero, that would imply that there was no information in that particular signal, while if that number were one, all of the information traders have about that variable is contained in that signal. The advantage of this measure is that it respects both that agents' inference is optimal (by assuming that they do the best they can with whatever signals they are endowed with) and also respects the fact that information may be redundant between signals.

Table 1.2 shows the relative reduction in uncertainty about particular macroeconomic variables (columns) from observing the short rate (first row) or a single private signal (remaining rows). This represents an extreme constraint on the information available to agents. The second row, for instance, suggests, that very little (around 3 %) of the information traders have about inflation comes from their inflation signal in particular (second row, first column). Effectively none of their information about the risk variable could come from their signal about inflation (second row, last column).[24] Three features of the table stand out. First, individual private signals are not terribly

[24]Note that the columns will not generally sum to 1 because some information is redundant between signals and because yields are also informative.

Table 1.3: Reduction in posterior uncertainty about about fundamentals and higher order beliefs from observing only private signals

	g^r	g^π	g^y	τ	v
x_t	0.02	0.41	0.41	0.16	0.24
$x_t^{(1)}$	0.01	0.09	0.08	0.09	0.02
$x_t^{(2)}$	0.02	0.12	0.11	0.08	0.03
$x_t^{(3)}$	0.04	0.12	0.12	0.08	0.04

informative in general. This is unsurprising given the sizable noise estimates. Second, as to be expected from the fact that agents understand the structure of the model, signals are informative not just about their own realization but about the realizations of other variables; for example, knowing something about risk tells you something about the output gap. This feature of the world is ignored in most exogenous information models because they typically assume agents learn about independent exogenous processes instead of learning about endogenous variables. Third, the short rate is very informative about fundamentals. Indeed, observing only the short rate would give you more information about fundamentals than observing any *individual* noisy signal. This is likely for two related reasons. First, the short rate depends directly on the contemporaneous realization of all of the fundamentals. Hence, it encodes the current state of the world. Second, it is observed totally without error. Despite the fact that agents are unable to perfectly identify which fundamental moved the short rate, they do know that noise does not factor into their observation. [25]

A more typical information assumption is that agents have access to several noisy idiosyncratic signals. Table 1.3 shows (relative to the benchmark with price signals) how much agents' posterior uncertainty is reduced by conditioning only on their four private signals. Here, I switch to considering the risks agents face (i.e., leaving rates and inflation in terms of their gaps) rather than the realizations.

As the first row of the table reveals, agents' private signals are most informative about the inflation

[25] One way of breaking this result would perhaps be to consider the informational set up of Melosi (2017), where the "interest rate shock" is actually three shocks; deviations from the Taylor rule and the central bank's forecast errors of inflation and the output gap. Since the interest rate signal is contaminated by the aggregate noise, it is harder to tell whether the interest rate has changed for fundamental reasons and the interest rate would be less reliable as a signal.

and output gaps. Agents get just under half of their information about the macroeconomy from their private signals. They can learn very little about risk and the implicit inflation target from observing their idiosyncratic signals and almost nothing about the rate *gap* $g_t^r = r_t - \tau_t$ (recall they are assumed to not observe the short rate in this counterfactual). The remaining lines show how much of their information about (higher-order) expectations come from private signals. The answer appears to be "not much." Since private signals are about the true realization of variables, rather than higher-order beliefs about those variables, they are more indirect signals about higher order beliefs and thus less informative.

Another way of thinking about the results in table 1.3 is "what are price signals informative about?" It turns out that the majority of information agents have about the financial risk factor (v) and monetary policy (summarized by g^r and τ) comes from observing price signals, including the short rate. Nearly all of their information about the first three orders of expectations is encoded in price signals (rows 2-4). Yield curve variables may not be fully informative about fundamentals or the beliefs of others, but the vast majority of information traders have about the latter seems to come from prices.

This result has two immediate implications. First, it validates thinking of the yield curve as a summary measure of what bond traders believe, which is a common interpretation in the financial press. Indeed, the model implies that the best *bond traders* can do to understand what others believe is by combining their understanding of how expectations are determined with the prices they observe. Since prices depend mostly on higher-order beliefs, prices are useful to bond traders even though they aren't fully informative about fundamentals. Second, the results caution against ignoring the informativeness of prices - agents may have very inaccurate signals on average, but the ability to learn from prices makes that less consequential. This matters directly for models featuring dispersed information. If one were to calibrate the informativeness of private signals by looking at the relative accuracy of a set of forecasts of endogenous variables, for instance, while ignoring the

Table 1.4: Reduction in posterior uncertainty about fundamentals and higher order beliefs from observing private signals and r_t

	g^r	g^π	g^y	τ	v
x_t	0.98	0.85	0.88	0.99	0.93
$x_t^{(1)}$	0.86	0.72	0.72	0.92	0.90
$x_t^{(2)}$	0.85	0.81	0.78	0.90	0.92
$x_t^{(3)}$	0.86	0.85	0.84	0.90	0.94

role of learning from prices, one would incorrectly conclude that private information must be quite accurate. On the other hand, calibrating the distribution of signals directly (for example, using the empirical distribution of productivity to calibrate signals about productivity), one might erroneously conclude agent's information was very bad by ignoring what one can learn from prices.

The estimates imply, however, that most of what traders learn about fundamentals and the first three orders of expectation can be gleaned from a combination of their private signals and the short rate. The results of this counterfactual are shown in table 1.4. Effectively all of what they know about monetary policy and risk comes from their private signals plus the policy rate, and 75% or more of what they know about the first three orders of expectation can be extracted without using bonds with a maturity of greater than one quarter.

There are two, related, implications of this result. The first is that this adds to recent evidence, such as that of Tang (2013) and Melosi (2017) that the Federal Reserve's policy instrument is an important signal. It tells observers a great deal about macroeconomic fundamentals and policy risks. And because it is a public signal that evidently contains a lot of what agents know about fundamentals, it plays an outsized role in market participants' higher order beliefs (along the lines of Morris and Shin (2002)). The second implication is that, assuming agents learn only from private signals and the policy rate - the information assumption of Melosi (2017) and Kohlhas (2015) - is a fair approximation of what bond traders appear to learn from (at least for low orders of expectations).

To emphasize the fact that the policy rate is somewhat special, table 1.5 shows a counterfactual where agents learn from their private signals and ten year yields rather than the federal funds rate.

Table 1.5: Reduction in posterior uncertainty about fundamentals and higher order beliefs from observing private signals and ten year yield

	g^r	g^π	g^y	τ	v
x_t	0.23	0.82	0.76	0.94	0.60
$x_t^{(1)}$	0.10	0.68	0.55	0.70	0.36
$x_t^{(2)}$	0.12	0.77	0.65	0.66	0.32
$x_t^{(3)}$	0.17	0.82	0.72	0.63	0.30

Ten year yields are informative about both fundamentals and higher-order beliefs above and beyond private signals, but they are less informative than the policy rate. This is especially true about the current policy gap (the first column) and the risk factor v (the last column).

Why are ten years yields less informative? The price of a ten year bond is determined not just by fundamentals (the short rate), but also higher order beliefs about the evolution of fundamentals over the next ten years, plus the maturity-specific shock. The fact that the bond is of longer maturity means that (increasingly) higher order beliefs play a greater role in its price. The fact that shocks to fundamentals are transitory, and higher-order beliefs play a bigger role in prices, imply that it will be less informative about current fundamentals.

More broadly, a lesson of this exercise is that different prices may be informative about different things, and some prices are more informative than others. Imperfect-information models with exogenous information involve making choices about what signals it is reasonable for agents to condition their forecasts on. A concern about most dispersed information models is that adding additional sources of information could dramatically affect the predictions of the model. Here, it appears adding additional information in the form of yields of longer-term bonds does not change how much agents learn about fundamentals (or low orders of expectation). This result comes with two caveats. First, in keeping with the majority of the literature, the model is constructed specifically to price a single type of asset. Other types of assets may be informative about a different set of macroeconomic or idiosyncratic risks. Although they are not considered explicitly, information from other asset types would be captured by the precision of private signals. If, for example, stock prices were very informative about the output gap, that implies that agents' signals about the output gap should be more precise. Second, the results of this section are "partial equilibrium" in a sense; the model does not allow for direct feedback from expectations to macroeconomic aggregates, as a more structural business cycle model might. However, from the point of view of an atomistic agent, macroeconomic aggregates are exogenous processes and the precise role of information in generating aggregate fluctuations should not matter.

1.8 Decomposing (higher order expectations in) the yield curve

Despite the abundance of public signals, non-trivial dispersion of higher order beliefs persists in the model. A natural question is what *direct* effect this dispersion of belief has had on prices, and more generally what the model attributes variation in bond yields to. In this section, I use estimates[26] of the underlying higher-order beliefs to decompose prices as outlined in section 1.4.4. I use this to answer two questions: (1) What does the model attribute changes in bond yields to - changes in rate expectations, "classical" risk premia, or higher-order beliefs? (2) Which higher-order beliefs matter for prices? Briefly, the answer to the first question is that (slowly adjusting) rate expectations play the largest role in determining yields at all horizons. Classical risk premia are nearly constant for bonds at all maturities, but the importance of the higher order wedge increases in the maturity of the bond. As for the second, a decomposition suggests that higher order beliefs about monetary policy - the rate gap g^r and the inflation target τ - drive most of the time variation in the wedge.

To think about the decomposition exercise informally, we can think about yields as be being driven by a part that is rate expectations and a residual. The information assumption implies the

[26]The results here are based on Kalman filtered estimates of the state, which can be thought of as inefficient estimates of the underlying hierarchy of higher-order expectations X_t. Kalman smoothing (i.e., the procedure described in Hamilton (1994a)), which takes account of the whole sample to derive estimates, presents numerical problems because the one-step ahead state forecast error matrix is ill-conditioned and inverting it presents numerical difficulties. This manifests in pathologies, such as states that are observed without error being inaccurate. The filtered estimates are most closest to what the Kalman smoother would imply at the end of the sample.

path of rate expectations, and thus determine the "expectations hypothesis" part of bond yields. Combined with our assumption of no-arbitrage and the way risk is priced, assuming agents have full information rational expectations implies the (non-econometric error portion of) the residual term is the risk premium; however, as the decomposition outlined in section 1.4.4 shows, under dispersed information the residual can be interpreted as the present discounted value of deviations from higher-order beliefs from average beliefs, and the gap between average expected short rates and the price that would obtain if agents counterfactually held common beliefs. Of course, the residual in the full information and dispersed-information models will be different because they assume different things about how people perceive short rates evolving over time.

Although one could focus on bonds of any maturity, here I focus on ten year yields.[27] The three-way decomposition is shown in figure 1.11.

Comparing the top two panels, it is clear that the model attributes the majority of movement in bond yields to rate expectations. In other words, accounting for agents' learning problem and their subjective rate expectations makes the premium for investing in long term bonds less volatile. That premium is divided between the "classical" premium and the higher order wedge; they are of roughly equal magnitudes, but the former is close to constant while the wedge varies over time. The reduced importance of compensation for risk in determining bond yields is qualitatively consistent with Piazzesi, Salomao and Schneider (2013), who use a very different methodology to arrive at this conclusion. For a more direct comparison, the full information model results reported in Appendix 1.H a sizable premium for holding ten year bonds, albeit with a great deal of uncertainty attached to the estimate.

This result implies that at least part of the dramatic failure of the expectations hypothesis is attributable to assuming agents' expectations are full information rational expectations. Accounting for the fact that agents' subjective forecasts may be different from the underlying full-information forecast means that volatile time varying risk premia are not needed to explain movements in long

[27]The results for other maturities are found in the appendix 1.G.3.

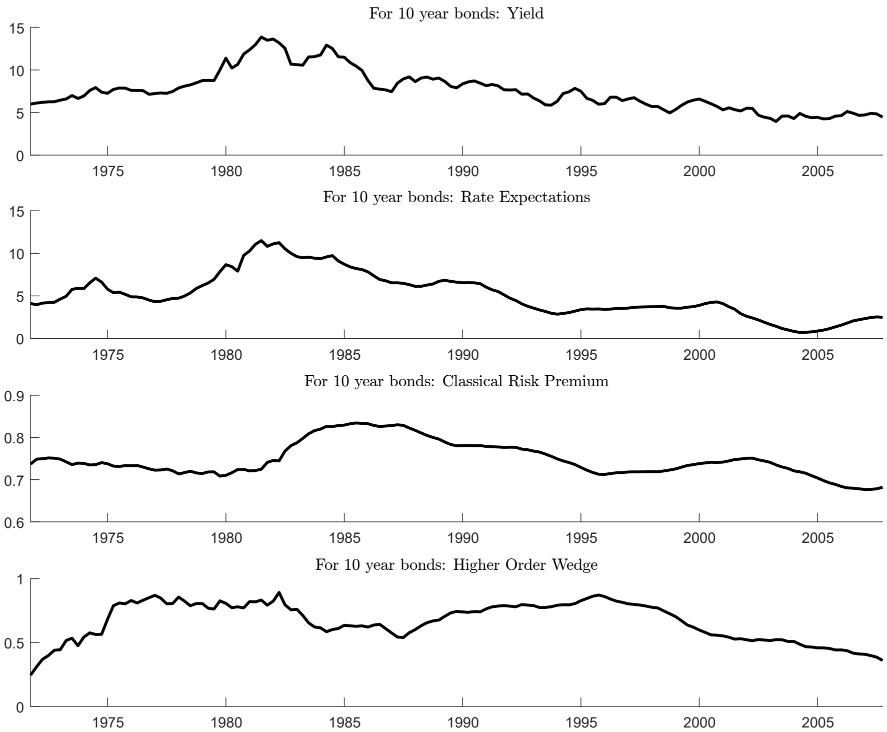


Figure 1.11: Decomposition of 10 year yields, dispersed information

term yields. The remaining premium for holding long term debt is partially about time varying compensation for risk (the "classical" risk premium) but the larger, time varying potion is attributable to the failure of consensus - that is, the fact that agents believe others have different beliefs. Both classical risk premia and the higher order wedge rose during the Volcker disinflation and slowly falling.

Why did risk premia and the higher order wedge decline over this period? Mechanically, a decline in the "classical" risk premium must be attributable to a decline in v_t over time. The inflation target variable τ_t is also falling over this period and since $\sigma_{v\tau} > 0$ the decline in the inflation target appears to have driven the decline in risk. Indeed, in appendix figures 1.20a and 1.20b, the estimated paths of these variables, and the first three orders of expectation about these variables, both steadily declined from the early 1980s to the early 2000s. We can also examine the role of higher order beliefs about these variables in determining yields. In figure 1.12, I show the higher-order wedge decomposed into the contribution from higher-order beliefs about fundamentals. The decomposition reveals that the model attributes growth in the higher order wedge to an increasing role for higher order beliefs about the inflation target. A smaller contribution comes from higher order beliefs about the rate gap $g^r = r_t - \tau_t$. Since r_t is commonly observed, this means that overall *policy uncertainty* contributes the most time variation to the wedge, at least for ten year yields. This is (partially) counterbalanced by higher order beliefs about the risk variable, which grew in the late 1970s and 80s but fell afterwards.

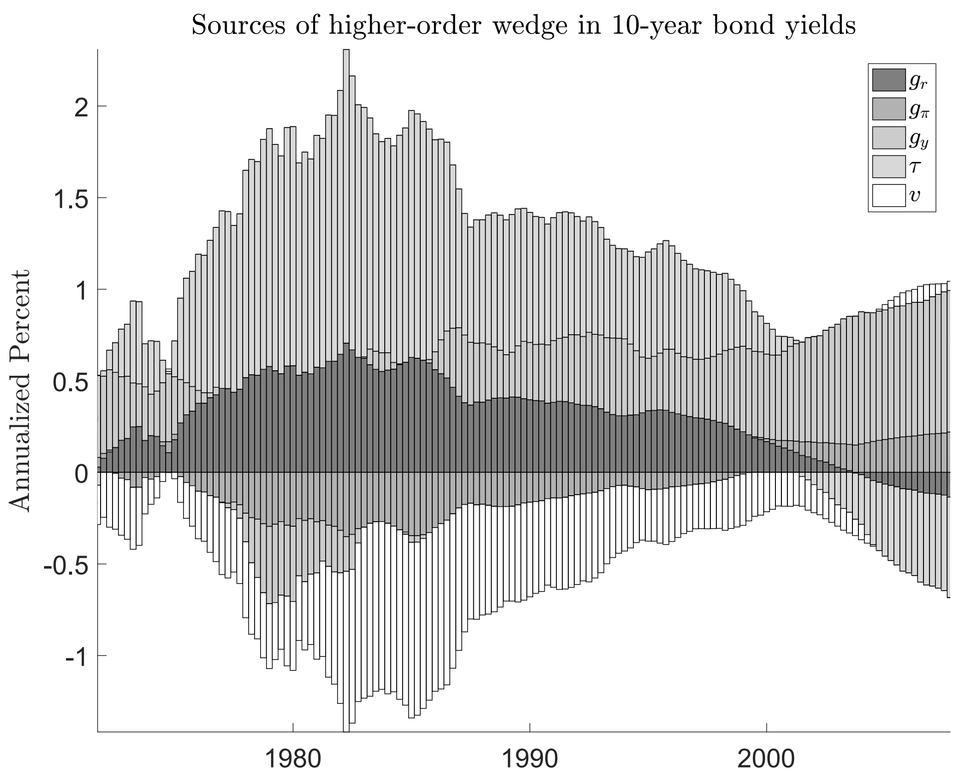


Figure 1.12: Decomposition of higher order wedge, ten year yields. Details of the decomposition are found in section 1.4.4.

How can we interpret this change in the higher order wedge? A plausible explanation for this change is uncertainty about the credibility of the Federal Reserve prior to the Volcker disinflation and a gradual increase in the public's trust in policymakers' commitment to fighting inflation. Uncertainty about the inflation target implies that people may have not only been unsure what the target was, but what others believe the target to be, and what they believe others believe the target was, and so on. Changes in the long-run inflation target are also intuitively more important for long-run bonds because the (nominal) payoff is lower when high unanticipated inflation is sustained. A greater commitment to fighting inflation and greater transparency (for example, announcing Federal funds rate changes and the "bias" of future policy moves) also lead to a gradual consensus about what the Fed's current stance of policy and its implicit inflation target likely was. Greater understanding of what the new monetary policy regime was, in other words, might have lead to the decline in the higher order wedge over time.

The importance of the credibility of the central bank's inflation target is consistent with other studies, such as that of Wright (2011). He argues that changes in the conduct of monetary policy lowered inflation uncertainty (measured using forecast dispersion and the time series of the standard deviation of the"permanent" component of inflation from a time series model with stochastic volatility), and that inflation uncertainty significantly explains the five-to-ten year forward premium across his sample of countries from 1990-2009. Both the results here and in Wright's paper are consistent with the idea that lower inflation uncertainty over time has caused the premium on long-term US government debt to decline. In my model, this is a result of both the relationship between changes in the inflation target and the risk variable - that is, the direct role of the inflation target and risk - as well as the (higher order) uncertainty about monetary policy arising endogenously from the traders' inference problem.

Examining the decomposition also reveals a degree of "canceling out" of the role of higher order beliefs. One way of thinking about this is the fact that different risks are not perfectly correlated with

Table 1.6: Contribution of higher-order wedge to yields at the posterior mode

	1 year	2 year	3 year	4 year	5 year	10 year
Average	0.03	0.15	0.18	0.20	0.44	0.66
Maximum	0.10	0.31	0.40	0.55	0.72	0.89
	Average contributions by source:					
g_r	-0.05	-0.28	-0.41	-0.11	0.11	0.31
g_π	0.03	0.23	0.36	0.19	0.04	-0.09
g_y	0.12	0.59	0.93	0.71	0.48	0.28
τ	-0.14	-0.78	-1.18	-0.50	0.04	0.54
v	0.01	0.09	0.14	-0.08	-0.23	-0.37

each other, and agents' higher-order beliefs are constrained by the macroeconomic environment. The average and maximum contribution of higher-order expectations to yields is shown in table 1.6. The contribution of higher order expectations to the wedge is increasing over the maturity of the bond. This is consistent with the model intuition at the beginning of section 1.4. Longer-maturity bonds are a function of future expectations of future stochastic discount factors. The longer the maturity of the bond, the larger the role of higher-order beliefs in determining the price. Table 1.6 also reveals how higher-order beliefs about different risks play different roles in the wedge across different maturities. This is a result of the expected time path of higher order beliefs and how different risks are priced at different horizons. In particular, as figure 1.7 reveals, higher-order beliefs about the output gap tend to fall over the medium term when the inflation target rises, which (along with the estimated prices of risk) explains why during the period when τ contributes the most to the higher-order wedge for 10 year yields is also when g^y plays such a large role for 3 and 4 year bonds. For bonds of low maturity, the contributions of higher order beliefs are very small in absolute terms and essentially cancel out on average.

The contribution of higher order beliefs, and their time series properties, are somewhat different here than in Barillas and Nimark (2015). They find that higher order beliefs play a larger role in general (with the peak contribution as a fraction of yields in the early 1990s) and also find a large negative role for the higher order wedge during the early 2000s. Part of the difference is likely due to the macroeconomic structure as opposed to the three-variable latent factor model they estimate.

Agents' beliefs about pricing factors and the role of those factors in prices are constrained by the covariances between asset prices and macroeconomic yields in the data. The latent factor model is more flexible. A second important difference is the choice of data. Barillas and Nimark directly use data on interest rate expectations to discipline belief formation, whereas I use inflation forecasts. Inflation forecasts in the SPF are generally regarded as being high quality - in fact, survey-based forecast measures generally perform better than most forecasting models (Faust and Wright (2013)). This feature of the data will imply agents have better average forecasts of inflation, which may mean the choice of data generates a more conservative role for higher order beliefs. Moreover, since zero coupon yields are constructed based on estimates from prices of different kinds of outstanding Treasury debt, there may be a concern that the "model" concept of Treasury yields is different from the concept that the SPF forecasters had in mind, which might exaggerate deviations of yields from rate expectations. This could influence estimates of the higher order wedge. Furthermore, the quarterly time series for interest rate forecasts in the SPF is much shorter than the inflation forecast data and has fewer responses in general. Inflation forecasts are available for the whole sample period. By contrast, the higher order wedge appears to play a greater influence in the Barillas and Nimark (2015) results once rate forecasts become available.

1.9 Conclusion

Survey evidence suggests professional forecasters have dispersed beliefs about future prices of Treasury bonds and macroeconomic variables. Motivated by this fact, I construct and estimate a macro-asset pricing model with dispersed information about macroeconomic fundamentals. The model allows for bond prices to be affected directly by policy, macroeconomic, and financial conditions; agents in the model are slow to identify fundamentals, and must learn them using both private information and commonly observed asset prices. Moreover, dispersed information, and the attendant gap between average beliefs and average beliefs about average beliefs, introduces a direct wedge into prices.

I use this model to understand the informativeness of prices for agents who lack knowledge about fundamentals and the beliefs of others, and to assess what role dispersion of macroeconomic belief may have played in determining Treasury yields prior to the Great Recession. The estimates imply that the direct role of belief dispersion is somewhat modest, but that most of the time variation in the higher order wedge is caused by policy-related factors. In particular, the wedge grew during the 1970s and early 1980s, along with the central bank's implicit inflation target, and fell over the course of the Great Moderation. This is consistent with gradual learning by agents about a new monetary policy regime and the emergence of a consensus about the conduct of monetary policy, perhaps arising from greater transparency and credibility.

I also provide new estimates of the quality of agents' private information and how much they learn from prices. I find individual private signals are quite noisy. By contrast, a great deal of agents' information about fundamentals comes from public prices, and prices are especially informative about the beliefs of others. Absent any of the public signals in the model, agents are about half as informed about macroeconomic fundamentals and know only about a fifth as much about the long-run inflation target of the central bank and financial risk. The most important signal appears to be the policy rate set by the central bank. By assumption, it is driven solely by fundamentals, rather than higher order beliefs, and thus agents attach a great deal of weight to it when forecasting those fundamentals. But since everyone does this, it is also informative about the beliefs of others. This role of public information was noted by Morris and Shin (2002); Tang (2013) and Melosi (2017) both emphasize the importance of the signaling channel of monetary policy. However, my paper is the first to measure the importance of the policy rate as a signal of fundamentals in an asset pricing setting where agents are not artificially constrained from learning from other prices.

The results here add to the body of evidence that deviations from full information are an important feature of the world. Accounting for agents' inference dramatically affects the size and interpretation of term premia. Moreover, dispersion of information does not disappear despite a large number of public signals, and it plays a direct role in prices. These results are instructive for what sorts of signals agents appear to learn from - in particular, asset prices are an important source of information. This suggests macroeconomic models with dispersed information should account for learning from prices when examining the importance of these frictions for macroeconomic outcomes or when assessing normative questions about optimality of policies that have implications for asset markets. It also suggests, at least for asset prices, market consistent information is not enough for aggregate irrelevance of information frictions. This is true in two senses: Dispersed information directly affects prices and the behavior of endogenous variables is quite different than under full information. This stands in contrast to some results in the macro-dispersed information literature (such as Venkateswaran and Hellwig (2009)). The model setting is different, but at a minimum my results emphasize the tension in this literature. Understanding the source of this tension can be important for future research.

There are a number of interesting and important extensions to this paper that would be worth pursuing. In this paper, I have focused on the informational content of a single class of assets - U.S. government debt. However, other assets may have different information implications worth exploring - for example, stocks may be informative about aggregate and sectoral shocks, and exchange rates may be informative about foreign and domestic shocks. Fully exploring the information that traders learn from different classes of assets would be a worthwhile extension. Second, extending the analysis to debt of different countries - along the lines of Wright (2011) - may also be informative about how changes in the monetary policy framework are associated with changes in the importance of higher-order beliefs. Third, throughout the paper I have taken advantage of the fact that yields are affine. This makes characterizing the higher order wedge and informativeness of signals straightforward. However, in the aftermath of the financial crisis, there were nonlinearities in yields introduced by the zero lower bound which may have affected prices' information content. Although technically challenging, extending the analysis to nonlinear filters (such as in the “shadow rate” literature i.e. Wu and Xia (2014)) could be worthwhile.

Appendix

1.A Additional forecast distributions

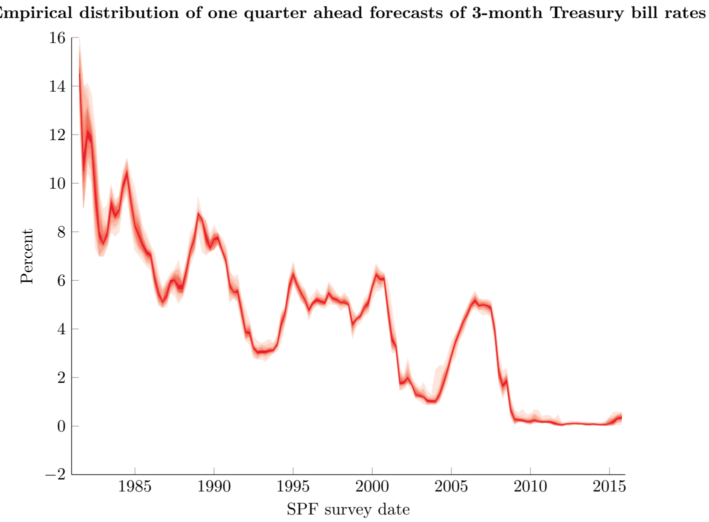


Figure 1.13: Distribution of SPF forecasts of next-quarter average rate on 3-month Treasury bill.

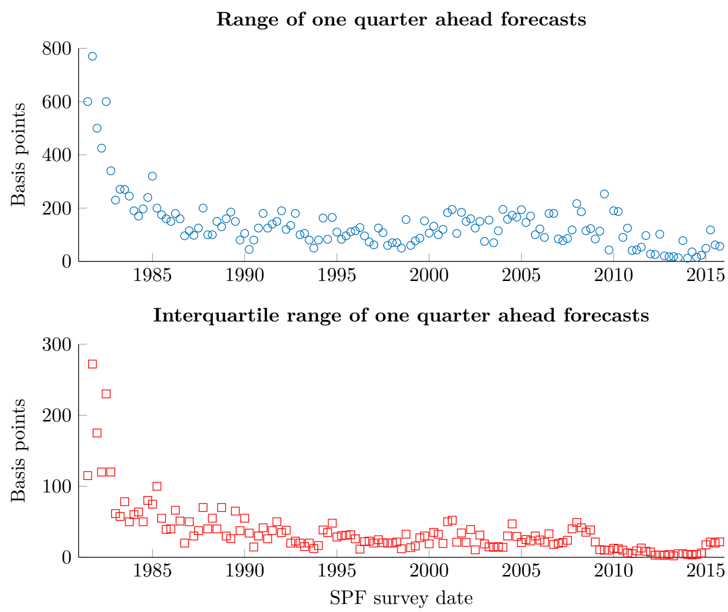


Figure 1.14: Full range (top) and interquartile range (bottom) of next-quarter average rate on 3-month Treasury bill.

1.B Model derivations

1.B.1 Intuition

Beginning with

$$P_t^n = E_t^j \left[M_{t+1}^j P_{t+1}^{n-1} \right]$$

Joint lognormality implies:

$$p_t^n = E_t^j[m_{t+1}^j] + E_t^j[p_{t+1}^{n-1}] + \frac{1}{2}\text{Var}(m_{t+1}^j + p_{t+1}^{n-1})$$

Iterating ahead for another agent (an arbitrary k that agent j will sell the bond to)

$$p_{t+1}^{n-1} = E_{t+1}^k[m_{t+2}] + E_{t+1}^k[p_{t+2}^{n-2}] + \frac{1}{2}\text{Var}(m_{t+2}^k + p_{t+2}^{n-2})$$

Then substituting this into the price expectation term:

$$\begin{aligned} p_t^n &= E_t^j[m_{t+1}^j] \\ &+ E_t^j[E_{t+1}^k(m_{t+2}^k)] + E_t^j[E_{t+1}^k p_{t+1}^{n-2}] + E_t^j[E_{t+1}^k p_{t+1}^{n-2}] \\ &+ \frac{1}{2}\text{Var}(m_{t+1}^j + p_{t+1}^{n-1}) + \frac{1}{2}\text{Var}(m_{t+2}^k + p_{t+2}^{n-2}) \end{aligned}$$

The fact that information sets are not nested means the law of iterated expectations does not apply. However, because no agent has particular information about other agents, agent j's expectations about k's expectations can be replaced by her expectation of the average expectation. Doing so, and integrating both sides over all agents implies the equation in the text.

1.B.2 The filtering problem

The individual agent's filtering problem, and its aggregation into the vector of average higher order expectations, follows Nimark (2007) and Barillas and Nimark (2015).

Call X_t the underlying state we want to estimate (the vector of higher order expectations, including 0th order expectations). Call $\Sigma_{t|t-1} \equiv E[(X_t - X_{t|t-1})(X_t - X_{t|t-1})']$.

Forecast step. Given information dated time $t-1$, j's forecast of the signal is

$$z_{t|t-1}^j = \mu_Z + DX_{t|t-1} \tag{1.36}$$

The associated covariance matrix of signal forecasting error is

$$\Omega_{t|t-1} \equiv E[(z_t^j - z_{t|t-1}^j)(z_t^j - z_{t|t-1}^j)'] \\ = D\Sigma_{t|t-1}D' + RR' \tag{1.37}$$

Updating step. Projection of $X_t - X_{t|t-1}$ onto $z_t^j - z_{t|t-1}^j$ and rearrangement gives that j's conditional expectation of the state given her time t information is

$$\begin{aligned} X_{t|t}^j &= X_{t|t-1} + \underbrace{\Sigma_{t|t-1}D'\Omega_{t|t-1}^{-1}}_{K_t}(z_t^j - z_{t|t-1}^j) \\ &= X_{t|t-1}^j + K(DX_t + R\begin{bmatrix} u_t \\ \eta_t^j \end{bmatrix} - DX_{t|t-1}^j) \\ &= \mu_X + \mathcal{F}X_{t-1|t-1}^j + K[D(\mu_X + \mathcal{F}X_{t-1} + \mathcal{C}u_t) + R\begin{bmatrix} u_t \\ \eta_t^j \end{bmatrix} - D(\mu_X + \mathcal{F}X_{t-1|t-1}^j)] \end{aligned} \tag{1.38}$$

Deriving the aggregate law of motion. Partition R into a part associated with aggregate shocks and one associated with idiosyncratic shocks, i.e. $R \equiv \begin{bmatrix} R_u & R_\eta \end{bmatrix}'$. Integrating $X_{t|t}^j$ to obtain the vector of *average* higher order expectations "zeros out" the idiosyncratic shocks, and we're left with

$$\begin{aligned} X_{t|t} =& \mu_X + (\mathcal{F} - KD\mathcal{F})\, X_{t-1|t-1} + K[D(\mu_X + \mathcal{F}X_{t-1} + \mathcal{C}u_t) + R_u u_t - D\mu_X + \mathcal{F}X_{t-1|t-1})] \\ =& \mu_X + (\mathcal{F} - KD\mathcal{F})\, X_{t-1|t-1} + KD\mathcal{F}X_{t-1} + K(D\mathcal{C} + R_u)u_t \end{aligned} \tag{1.39}$$

Note that these expressions have been written in terms of the steady state Kalman gain K. To find the steady state Kalman gain, we can derive the following discrete-time algebraic Riccati equation (which follows from some algebra during the updating step)

$$\begin{aligned} \Sigma_{t+1|t} =& E[(X_{t+1} - X_{t+1|t})(X_{t+1} - X_{t+1|t})'] \\ =& \mathcal{F}(\Sigma_{t|t-1} - \Sigma_{t|t-1}D'\Omega_{t|t-1}^{-1}D\Sigma_{t|t-1})\mathcal{F}' + RR' \end{aligned} \tag{1.40}$$

and iterate until convergence. The resulting steady state $\Sigma_{t+1|t}$, combined with (1.37), immediately implies K.

Recall that we conjectured a VAR(1) process for X_t, namely

$$X_t \equiv \begin{bmatrix} x_t \\ X_{t|t} \end{bmatrix} = \mu_X + \mathcal{F}X_{t-1} + \mathcal{C}u_t \tag{1.41}$$

so matching coefficients we can find $\mathcal{C}, \mathcal{F}$ (recall there are d factors and we truncate at order $\bar{k}$)

$$\begin{aligned} \mathcal{F} &= \begin{bmatrix} F^P & \mathbf{0}_{d\times d\bar{k}} \\ \mathbf{0}_{d\bar{k}\times d} & \mathbf{0}_{d\times d\bar{k}} \end{bmatrix} + \begin{bmatrix} \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d\bar{k}} \\ \mathbf{0}_{d\bar{k}\times d} & [\mathcal{F} - KD\mathcal{F}]_{\_} \end{bmatrix} + \begin{bmatrix} \mathbf{0}_{d\times d(\bar{k}+1)} \\ [KD\mathcal{F}]_{\_} \end{bmatrix} \\ \mathcal{C} &= \begin{bmatrix} C & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ [K(D\mathcal{C} + R_u)]_{\_} \end{bmatrix} \end{aligned} \tag{1.42}$$

where _ indicates truncation to ensure conformability and considering with only considering expectations up to $\bar{k}$.

1.B.3 Generating bond price equations

The steps here are identical to Barillas and Nimark (2015).

$$p_t^n = A_n + B_n' X_t + \nu_t^n$$

To arrive at this form, substitute the SDF (1.13) into the (log) arbitrage condition:

$$p_t^n = \ln E\left[\exp\left\{-r_t - \frac{1}{2}\Lambda_t^{j\prime}\Sigma_a\Lambda_t^j - \Lambda_t^{j\prime}a_{t+1}^j + p_{t+1}^{n-1}\right\}|\Omega_t^j\right] \tag{1.43}$$

Here we use the definition of a_{t+1}^j (1.14) to substitute p_{t+1}^{n-1} out for its expectation plus the forecast error for that particular maturity

$$P_{t+1}^{n-1} = E\left[p_{t+1}^{n-1}|\Omega_t^j\right] + e_{n-1}' a_{t+1}^j \tag{1.44}$$

where e_n' is a horizontal selection vector with 1 in the nth element and zeros elsewhere.

Since we assumed agents knew the model equations we can write, we can write

$$E[p_{t+1}^{n-1}|\Omega_t^j] = A_{n-1} + B_{n-1}' \underbrace{\left(\mu_X + \mathcal{F}E[X_t|\Omega_t^j]\right)}_{E[X_{t+1}|\Omega_t^j]} \tag{1.45}$$

Define an operator H that selects just the average higher order expectations from X_t^j (1.16), that is

$$
\begin{aligned}
E[X_t|\Omega_t^j] =&HX_t^j \\
\text{where } H =& \begin{bmatrix} \mathbf{0}_{d\bar{k}\times d} & I_{d\bar{k}} \\ \mathbf{0}_{d\times d} & \mathbf{0}_{d\times d\bar{k}} \end{bmatrix}
\end{aligned} \tag{1.46}
$$

Combining these three expressions gives

$$
E[p_{t+1}^n|\Omega_t^j] = A_{n-1} + B_{n-1}'\mu_X + B_{n-1}'\mathcal{F}HX_t^j \tag{1.47}
$$

substituting this in to the no-arbitrage condition

$$
p_t^n = \ln E\left[\exp\left\{-r_t - \frac{1}{2}\Lambda_t^{j\prime}\Sigma_a\Lambda_t^j - \Lambda_t^{j\prime}a_{t+1}^j + A_{n-1} + B_{n-1}'\mu_X + B_{n-1}'\mathcal{F}HX_t^j + e_{n-1}'a_{t+1}^j\right\}|\Omega_t^j\right] \tag{1.48}
$$

The inner expression consists of constants and lognormal random variables. It can be written in terms of things known to agent j at time t (so the expectation is superfluous):

$$
\begin{aligned}
p_t^n = \ln\exp\bigg\{ & -r_t - \frac{1}{2}\Lambda_t^{j\prime}\Sigma_a\Lambda_t^j - A_{n-1} + B_{n-1}'\mu_X + B_{n-1}'\mathcal{F}HX_t^j \\
& + \frac{1}{2}(e_{n-1}'\Sigma_a e_{n-1} + \Lambda_t^{j\prime}\Sigma_a\Lambda_t^j - 2e_{n-1}'\Sigma_a\Lambda_t^j)\bigg\}
\end{aligned} \tag{1.49}
$$

where the last term is 1/2 times the variance of $(e_{n-1}' - \Lambda_t^j)a_{t+1}^j$. Simplifying:

$$
p_t^n = -r_t + A_{n-1} + B_{n-1}\mu_X + B_{n-1}'\mathcal{F}HX_t^j + \frac{1}{2}e_{n-1}'\Sigma_a e_{n-1} - e_{n-1}'\Sigma_a\Lambda_t^j \tag{1.50}
$$

The price of the n period bond at time t is a function of constants, the current risk-free rate, and j specific terms. By no arbitrage, this expression holds for all j at all times, but, like Barillas and Nimark (2015), I focus on a hypothetical agent whose state coincides with the cross-sectional average state. Then we can substitute X_t for X_t^j in the previous expression, since $X_t \equiv \int X_t^j dj$.

Finally substitute (1.6) and (1.15) into the previous expression:

$$
\begin{aligned}
p_t^n =& -(\delta_0 + \delta_x' x_t) + A_{n-1} + B_{n-1}\mu_X + B_{n-1}'\mathcal{F}HX_t \\
& + \frac{1}{2}e_{n-1}'\Sigma_a e_{n-1} - e_{n-1}'\Sigma_a\left(\Lambda_0 + \Lambda_x^j + \Lambda_\nu\int E[\nu_t|\Omega_t^j]dj\right)
\end{aligned} \tag{1.51}
$$

Define $\delta_X' \equiv \begin{bmatrix}\delta_x' & \mathbf{0}\end{bmatrix}$ and rearrange this

$$
\begin{aligned}
p_t^n =& -\delta_0 + A_{n-1} + B_{n-1}\mu_X + \frac{1}{2}e_{n-1}'\Sigma_a e_{n-1} - e_{n-1}'\Sigma_a\Lambda_0 \\
& -\delta_X' X_t + B_{n-1}'\mathcal{F}HX_t - e_{n-1}'\Sigma_a\Lambda_x X_t \\
& - e_{n-1}'\Sigma_a\Lambda_\nu \int E[\nu_t|\Omega_t^j]dj
\end{aligned} \tag{1.52}
$$

We had guessed

$$
p_t^n = A_n + B_n' X_t + \nu_t^n \tag{1.9}
$$

To arrive at the conjectured form, impose two additional restrictions. First, restrict:

$$
\Lambda_\nu = -\Sigma_a^{-1} \tag{1.53}
$$

which also reduces the number of free parameters in the model. Secondly, we can substitute to replace the remaining $e_{n-1}'\int E[\nu_t|\Omega_t^j]dj$ term via a convenient normalization. Note model consistent expectations and the conjectured bond price equation imply

$$
p_t^n = E[A_n + B_n X_t + \nu_t^n|\Omega_t^j] = A_n + B_n H X_t^j + e_{n-1}' E[\nu_t|\Omega_t^j] \tag{1.54}
$$

Setting this equal to the conjectured bond equation implies

$$
\begin{aligned}
A_n + B_n H X_t + e_{n-1}'\int E[v_t|\Omega_t^j]dj =& \quad A_n + B_n X_t + \nu_t^n \\
\Rightarrow e_{n-1}'\int E[v_t|\Omega_t^j]dj =& \quad B_n(I-H)X_t + \nu_t^n
\end{aligned} \tag{1.55}
$$

Substituting these restrictions:

$$
\begin{aligned}
p_t^n =& -\delta_0 + A_{n-1} + B_{n-1}'\mu_X + \frac{1}{2}e_{n-1}'\Sigma_a e_{n-1} - e_{n-1}'\Sigma_a\Lambda_0 \\
& -\delta_X' X_t + B_{n-1}'\mathcal{F}HX_t - e_{n-1}'\Sigma_a\Lambda_x X_t \\
& + B_n(I-H)X_t + \nu_t^n
\end{aligned} \tag{1.56}
$$

Finally, write $B = \begin{bmatrix} B_2' & \dots & B_{\bar{n}}' \end{bmatrix}$ and note that $B_n = e_{n-1}B$. Normalizing prices of risk:

$$
\Lambda_x = \widehat{\Lambda}_x + B(I-H) \tag{1.57}
$$

and then

$$
\begin{aligned}
p_t^n = & - \delta_0 + A_{n-1} + B'_{n-1}\mu_X + \frac{1}{2}e'_{n-1}\Sigma_a e_{n-1} - e'_{n-1}\Sigma_a \Lambda_0 \\
& - \delta'_X X_t + B'_{n-1}\mathcal{F}HX_t - e'_{n-1}\Sigma_a \widehat{\Lambda}_x X_t \\
& + \nu_t^n
\end{aligned}
\tag{1.58}
$$

This implies the recursive forms for the bond price equations:

$$
A_{n+1} = \quad -\delta_0 + A_n + B_n\mu_X + \frac{1}{2}e'_n\Sigma_a e_n - e'_n\Sigma_a\Lambda_0 \tag{1.59}
$$

$$
B'_{n+1} = \quad -\delta_X + B'_n\mathcal{F}H - e'_{n-1}\Sigma_a\widehat{\Lambda}_x \tag{1.60}
$$

with

$$
A_1 = \quad -\delta_0 \tag{1.61}
$$

$$
B_1 = \quad -\delta'_X \tag{1.62}
$$

which implies $p_t^1 = -\delta_0 + [\delta_x, \mathbf{0}]X_t = -r_t$.

1.B.4 Macroeconomic structure

$$
P_0 = \begin{bmatrix} 1 & -(1-\phi_r)\phi_\pi & -(1-\phi_r)\phi_y & 0 & -(1-\phi_r)\phi_v \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{1.63}
$$

$$
\mu_x = \begin{bmatrix} (1-\phi_r)g^r - (1-\phi_r)g_y \\ -\rho_{\pi r}g_r - \rho_{\pi y}g^y \\ g_y - \rho_{yr}g^r - \rho_{yy}g^y \\ (1-\rho_{\tau\tau})\tau \\ 0 \end{bmatrix} \tag{1.64}
$$

$$
P_1 = \begin{bmatrix} \phi_r & 0 & 0 & 0 & 0 \\ \rho_{\pi r} & \rho_{\pi\pi} & \rho_{\pi y} & 0 & \rho_{\pi v} \\ \rho_{yr} & \rho_{y\pi} & \rho_{yy} & 0 & \rho_{yv} \\ 0 & 0 & 0 & \rho_{\tau\tau} & 0 \\ 0 & 0 & 0 & 0 & \rho_{vv} \end{bmatrix} \tag{1.65}
$$

$$
\Sigma_0 = \begin{bmatrix} \sigma_r & 0 & 0 & 0 & 0 \\ 0 & \sigma_\pi & 0 & \sigma_{\pi\tau}\sigma_\tau & 0 \\ 0 & \sigma_{y\pi}\sigma_\pi & \sigma_{y\pi}\sigma_\pi & \sigma_{y\tau}\sigma_\tau & 0 \\ 0 & 0 & 0 & \sigma_\tau & 0 \\ \sigma_{vr} & \sigma_{v\pi} & \sigma_{vy} & \sigma_{v\tau} & \sigma_v \end{bmatrix} \tag{1.66}
$$

$$r_t = \delta_0 + \delta_x' x_t \tag{1.6}$$

with

$$\delta_0 = \mathbf{0}_{5\times 1} \tag{1.67}$$

$$\delta_x' = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 \end{bmatrix} \tag{1.68}$$

and the matrices governing the evolution of fundamentals (1.7) as

$$\begin{aligned} \mu^P &= P_0^{-1}\mu_0 \\ F^P &= P_0^{-1}P_1 \\ C &= P_0^{-1}\Sigma_0 \end{aligned} \tag{1.69}$$

1.B.5 Restrictions on Prices of Risk

Recall expressions for the stochastic discount factor and prices of risk:

$$m_{t+1}^j = -r_t - \frac{1}{2}\Lambda_t^{j\prime}\Sigma_a\Lambda_t^j - \Lambda_t^{j\prime}a_{t+1}^j \tag{1.13}$$

$$\Lambda_t^j = \Lambda_0 + \Lambda_x X_t^j + \Lambda_\nu E[\nu_t|\Omega_t^j] \tag{1.15}$$

To impose the Ireland (2015) restriction, I set:

$$\lambda_0 = \begin{bmatrix} \lambda^r & \lambda^\pi & \lambda^y & \lambda^\tau & 0 \end{bmatrix}' \tag{1.70}$$

$$\lambda_x = \begin{bmatrix} 0 & 0 & 0 & 0 & \lambda_x^r \\ 0 & 0 & 0 & 0 & \lambda_x^\pi \\ 0 & 0 & 0 & 0 & \lambda_x^y \\ 0 & 0 & 0 & 0 & \lambda_x^\tau \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{1.71}$$

To additionally impose the Barillas and Nimark (2015) restriction, recall that the vector of bond price innovations a_{t+1}^j is a linear combination of forecasting error in the factors X_{t+1} and maturity-specific price shocks ν_{t+1}.

$$a_{t+1}^{j} = \Psi \begin{bmatrix} X_{t+1} - E\left[X_{t+1}|\Omega_t^j\right] \\ \nu_{t+1} \end{bmatrix} \tag{1.72}$$

To see this, write j's one-period ahead bond pricing error for a particular maturity as

$$\begin{aligned} a_{t+1}^{n,j} &= p_{t+1}^{n-1} - p_{t+1|t}^{j} \\ &= B_{n-1}'(X_{t+1} - E_t^j X_t) + \nu_{t+1}^{n-1} \end{aligned} \tag{1.73}$$

So stacking these errors in a vector a_{t+1}^j gives

$$a_{t+1}^{j} = \underbrace{\begin{bmatrix} B_1' & \\ \vdots & I_{\bar{n}-1} \\ B_{\bar{n}-1}' & \end{bmatrix}}_{\Psi} \begin{bmatrix} X_{t+1} - E\left[X_{t+1}|\Omega_t^j\right] \\ \nu_{t+1} \end{bmatrix} \tag{1.74}$$

Left multiplying by $\Lambda_{t+1}^{j\prime}$:

$$\Lambda_{t+1}^{j\prime} a_{t+1}^{j} = \Lambda_{t+1}^{j\prime} \Psi \begin{bmatrix} X_{t+1} - E\left[X_{t+1}|\Omega_t^j\right] \\ \nu_{t+1} \end{bmatrix} \tag{1.75}$$

We want to restrict this so that

$$\Lambda_{t+1}^{j\prime} a_{t+1}^{j} = \left(\begin{bmatrix} \lambda_0 \\ \mathbf{0} \end{bmatrix} + \begin{bmatrix} \lambda_x & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \right)' \begin{bmatrix} X_{t+1} - E\left[X_{t+1}|\Omega_t^j\right] \\ \nu_{t+1} \end{bmatrix} \tag{1.76}$$

If we removed dispersed information or maturity-specific shocks, this restriction would imply only fundamentals matter for bond prices, given the restrictions in (1.70) and (1.71). When maturity specific shocks are equal to zero, these additional restrictions must hold:

$$\begin{aligned} \begin{bmatrix} \lambda_0 \\ \mathbf{0} \end{bmatrix}' \begin{bmatrix} X_{t+1} - E\left[X_{t+1}|\Omega_t^j\right] \\ \nu_{t+1} \end{bmatrix} &= \Lambda_0' a_{t+1}^j \\ \left(\begin{bmatrix} \lambda_x & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} X_t^j \right)' \begin{bmatrix} X_{t+1} - E\left[X_{t+1}|\Omega_t^j\right] \\ \nu_{t+1} \end{bmatrix} &= \widehat{\Lambda}_x' a_{t+1}^j \end{aligned} \tag{1.77}$$

where $\widehat{\Lambda_x}$ is a normalization (see appendix 1.B.3). This can be achieved by setting

$$
\begin{aligned}
\Phi &= \Psi(\Psi'\Psi)^{-1} \\
\Lambda_0 &= \Phi \begin{bmatrix} \lambda_0 \\ \mathbf{0} \end{bmatrix} \\
\widehat{\Lambda}_x &= \Phi \begin{bmatrix} \lambda_x & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \\
\Psi &= \begin{bmatrix} B_1' & \\ \vdots & I_{\bar{n}-1} \\ B_{\bar{n}-1}' & \end{bmatrix}
\end{aligned}
\tag{1.78}
$$

These restrictions are the same as those imposed in Barillas and Nimark (2015).

1.C Fixed point procedure

0. Given a set of parameters, we construct H using (1.46), δ_0 δ_x, μ^P, F^P, C, λ_0 and λ_x using (1.67)-(1.71). We need an initial guess of B (typically starting with the full information B). This implies an initial A_n using (1.17), and thus D for the agents' filtering problem. We must also guess $\mathcal{C}, \mathcal{F}$, typically at the full information solution.

1. The Kalman filtering problem implies steady state $\Sigma_{t+1|t}$ using (1.40).[28] This implies steady state $\Omega_{t+1|t}$ and K. Construct $\mathcal{F}, \mathcal{C}$ from (1.42).

2. We have

$$
\Sigma^a = \Psi \begin{bmatrix} \Sigma_{t+1|t} & \mathbf{0} \\ \mathbf{0} & \Sigma_\nu \end{bmatrix} \Psi' \tag{1.79}
$$

where Σ_ν is the covariance matrix of maturity shocks, a diagonal matrix where the nonzero elements are of the form

$$
\sqrt{\mathrm{Var}(e_t^n)} = n\sigma_\nu \tag{1.80}
$$

[28] In practice, the bulk of time spent on the solution is in this step. Since no closed form exists for the Kalman gain in a general multivariate setting, I must numerically find the Kalman gain by solving the discrete-time algebraic Riccati equation. In this particular setting, the fastest way to solve the equation seems to be through iteration until convergence, with an additional step to ensure that the matrix is symmetric. The latter step is necessary to avoid numerical problems due to round-off which is common in large-dimension Kalman filtering problems.

(this implies the variance of maturity shocks is constant across yields, which reduces the number of free parameters). Recall we had assumed $\Lambda_\nu = -\Sigma_a^{-1}$.

3. Update our guess of B using (1.18) and check for convergence. If $B, \mathcal{C}, \mathcal{F}$ have converged, stop. Else, go to step 1.

1.D Econometric matrices

The model-consistent notion of dispersion of signals around the average comes from agents' Kalman filtering equations. Any dispersion in belief must come from idiosyncratic signals. The idiosyncratic error covariance matrix is the solution to the following Riccati equation:

$$\begin{aligned}\Sigma_j &= E[(X_t^j - X_t^{(1)})(X_t^j - X_t^{(1)})'] \\ &= (\mathcal{F} - KD\mathcal{F})\Sigma_j(\mathcal{F} - KD\mathcal{F})' + KR_\eta R_\eta' K'\end{aligned} \tag{1.81}$$

Hence the cross-sectional variance in average forecasts is just the appropriate element of Σ_j:

$$Var(\pi_{t|t}^j) = \underbrace{[0, 1, 0, 1, 0, \mathbf{0}_{1\times d*(\bar{k})]}]}_{\equiv e^\pi} \Sigma_j e^{\pi\prime} \tag{1.82}$$

The non-constant parts of the econometric matrices in (1.35) are:

$$\bar{D}_t = \begin{bmatrix} I_4 & \mathbf{0}_{4\times 2} & \mathbf{0}_{4\times\bar{k}} \\ & -\frac{1}{4}B_4' & \\ & -\frac{1}{8}B_8' & \\ & -\frac{1}{12}B_{12}' & \\ & -\frac{1}{16}B_{16}' & \\ & -\frac{1}{20}B_{20}' & \\ & -\frac{1}{40}B_{40}' & \\ & e^\pi\mathcal{F} \times I_{m_t^1} & \\ & e^\pi \cdot \mathcal{F}^4 \times I_{m_t^4 s} & \end{bmatrix} \tag{1.83}$$

$$\bar{R}_t = \begin{bmatrix} 0_{3\times d+\bar{n}-1+m_t^1} \\ \sigma_\nu \begin{bmatrix} & e_3 \\ & e_7 \\ & e_{11} \\ & e_{15} \\ \mathbf{0}_{6\times d} & e_{19} \\ & e_{39} \end{bmatrix} \\ \sqrt{e^\pi \mathcal{F}\Sigma^j \mathcal{F}' e^{\pi\prime}} \times I_{m_t^1} \\ \sqrt{e^\pi \mathcal{F}^4 \Sigma^j (\mathcal{F}^4)' e^{\pi\prime}} \times I_{m_t^4} \end{bmatrix} \tag{1.84}$$

For the full information model, the equations are the same. However, instead of the σ_ν terms, the observed bond yields are assumed to be observed with yield-specific error.[29], and the cross-sectional estimation error terms for the forecasts are replaced with horizon-specific error terms $\widetilde{\sigma_\pi}^h, h = 1, 4$.

1.E Information-theoretic concepts

In the discussion of the share of information coming from private signals in section 1.7, I refer to a number of concepts from information theory, which I detail here without proof; more details are found in Veldkamp (2011) and Cover and Thomas (2006). As described in section 1.7, I characterize the extent to which variables are informative using the notion of entropy - the amount of information required to describe a random variable (Cover and Thomas (2006)). Entropy is typically expressed in terms of "bits," i.e., in terms of log base 2 units, which is convenient because the entropy of a fair coin toss is 1 bit. Intuitively, the entropy of a random variable in bits is the number of $0-1$ binary signals required on average to describe its realization.

The entropy of a normally distributed variable. If x is a normally distributed variable with variance σ^2, its entropy is $\frac{1}{2}\log_2(2\pi e\sigma^2)$ (Cover and Thomas, 2006, Chapter 8).

Conditional entropy. Conditional entropy $H(x|y)$ is a measure of how much information it takes to describe x given that y is known (Veldkamp, 2011, Chapter 3.2). It is defined as the joint entropy

[29] A common practice to avoid a stochastic singularity problem, used by Ireland (2015) among others, is to assume that only certain yields are observed with error. However, as Piazzesi (2009) points out, which set of yields to treat as viewed with error is essentially arbitrary, and assuming all of them are viewed with error does not pose any computational difficulty in this setting.

of x, y minus the entropy of y, that is $H(x|y) = H(x, y) - H(y)$. The calculation of the conditional entropy of a normal variable is analogous to the unconditional case, replacing the variance with the conditional variance (Veldkamp (2011)).

Mutual information. The mutual information of two variables x and y, $I(x; y)$ is the measure of the amount of information one contains about the other. It can be calculated in terms of entropies ((Cover and Thomas, 2006, Theorem 2.4.1))):

$$I(x, y) = H(x) - H(x|y) = H(y) - H(y|x)$$

Measure of signal use. Similar to Melosi (2017, 2014), I use the "share" of mutual information as my characterization of how much information about a variable comes from (a particular subset) of signals ω_{red}. In particular, the "share" of information about a variable x used by an agent is:

$$Share_x = I(x; \omega_{red})/I(x; \omega_{full)}$$

where ω_{red} is the reduced set of signals (for example, only private signals without the use of bond prices) and ω_{full} is the complete set of signals detailed in section 1.4.3.

In practice, conditional variances needed to calculate mutual information are taken as particular entries from agents' state nowcasting error matrix ($\Sigma_{t|t}$) (see appendix 1.B.2). The conditional variance of the subset of signals is calculated by solving the filtering problem of the agent assuming they have a "counterfactual" subset of signals (just as described in appendix 1.B.2, using $A, B, \mathcal{F}, \mathcal{C}$ from the actual model solution.

Note that this share is bounded between 0 and 1 because, on average, conditioning must reduce entropy (Cover and Thomas, 2006, Theorem 2.6.5).

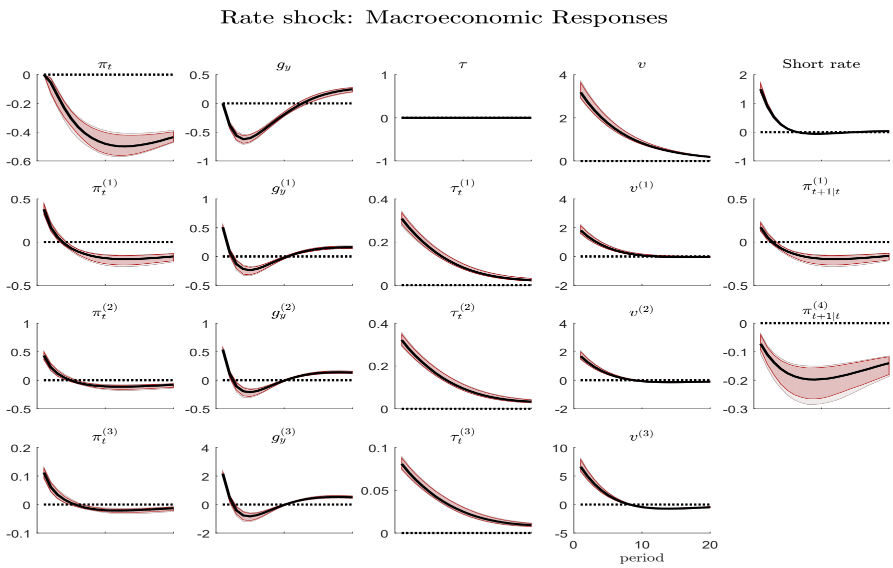
1.F Priors

Table 1.7: Prior distribution of model parameters

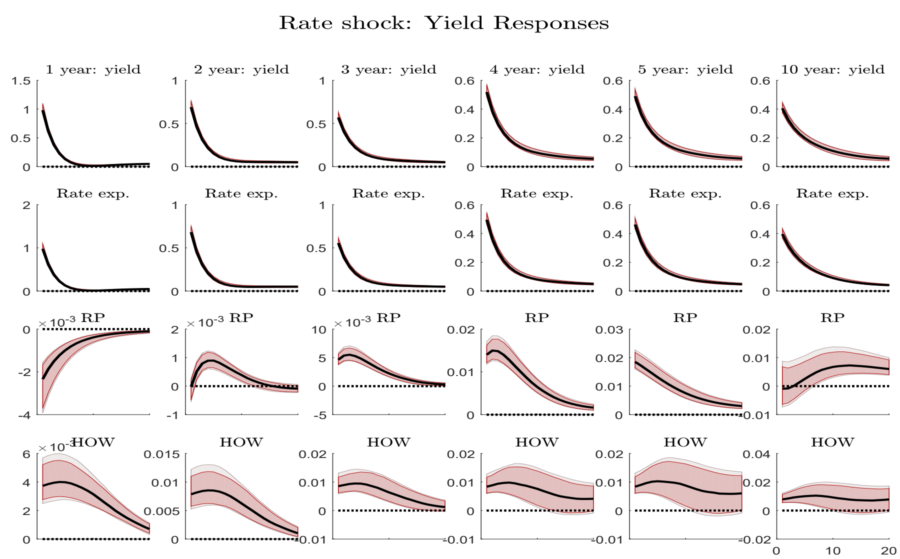
Parameter	Prior distribution	Prior mean	Prior s.d.	Model
ϕ_r	Beta	0.5000	0.0500	
ϕ_π	Gamma	0.5000	0.3000	
ϕ_y	Gamma	0.5000	0.3000	
ϕ_v	Trunc. Normal	0.0000	0.5000	
σ_r	Inverse Gamma	0.0050	0.2000	
ρ_{yr}	Normal	-1.0000	0.5000	
$\rho_{y\pi}$	Normal	0.0000	0.5000	
ρ_{yy}	Inverse Gamma	0.9000	0.2000	
ρ_{yv}	Trunc. Normal	0.0000	2.0000	
$\sigma_{y\pi}$	Normal	0.0000	1.0000	
$\sigma_{y\tau}$	Normal	0.0000	1.0000	
σ_y	Inverse Gamma	0.1000	3.0000	
σ_τ	Inverse Gamma	0.0050	0.3000	
$\rho_{\pi r}$	Normal	0.0000	2.0000	
$\rho_{\pi\pi}$	Inverse Gamma	0.9000	0.2000	
$\rho_{\pi y}$	Normal	0.0000	0.5000	
$\rho_{\pi v}$	Normal	0.0000	2.0000	
$\sigma_{\pi\tau}$	Normal	0.0000	3.0000	
σ_π	Inverse Gamma	0.0050	0.3000	
ρ_{vv}	Beta	0.8000	0.1000	
σ_{vr}	Normal	0.0000	3.0000	
$\sigma_{v\pi}$	Normal	0.0000	3.0000	
σ_{vy}	Normal	0.0000	3.0000	
$\sigma_{v\tau}$	Normal	0.0000	3.0000	
λ_r	Uniform(-100,100)			
λ_π	Uniform(-100,100)			
λ_y	Uniform(-100,100)			
λ_τ	Uniform(-100,100)			
λ_r^x	Uniform(-100,100)			
λ_π^x	Uniform(-100,-0.001)			
λ_y^x	Uniform(-100,100)			
λ_τ^x	Uniform(-100,100)			
σ_ν	Uniform(0.001,0.02)			DI
$\widetilde{\sigma}_\pi$	Uniform(0.001,100)			DI
$\widetilde{\sigma}_y$	Uniform(0.001,100)			DI
$\widetilde{\sigma}_\tau$	Uniform(0.001,100)			DI
$\widetilde{\sigma}_v$	Uniform(0.001,100)			DI
$\widetilde{\sigma}_4$	Inverse Gamma	1.0000	3.0000	FI
$\widetilde{\sigma}_8$	Inverse Gamma	1.0000	3.0000	FI
$\widetilde{\sigma}_{12}$	Inverse Gamma	1.0000	3.0000	FI
$\widetilde{\sigma}_{16}$	Inverse Gamma	1.0000	3.0000	FI
$\widetilde{\sigma}_{20}$	Inverse Gamma	1.0000	3.0000	FI
$\widetilde{\sigma}_{40}$	Inverse Gamma	1.0000	3.0000	FI
$\widetilde{\sigma}_\pi^1$	Inverse Gamma	1.0000	3.0000	FI
$\widetilde{\sigma}_\pi^4$	Inverse Gamma	1.0000	3.0000	FI

1.G Additional Results, dispersed information model

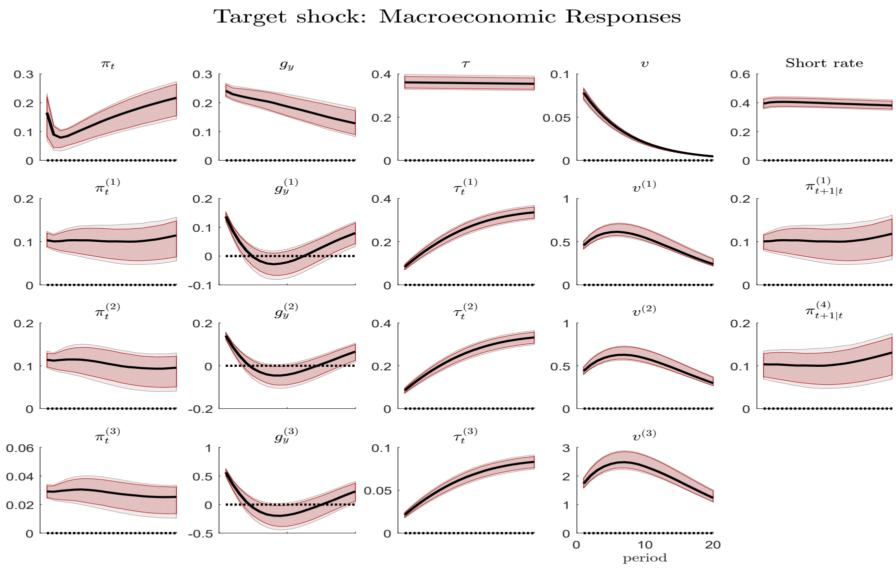
1.G.1 Impulse Responses



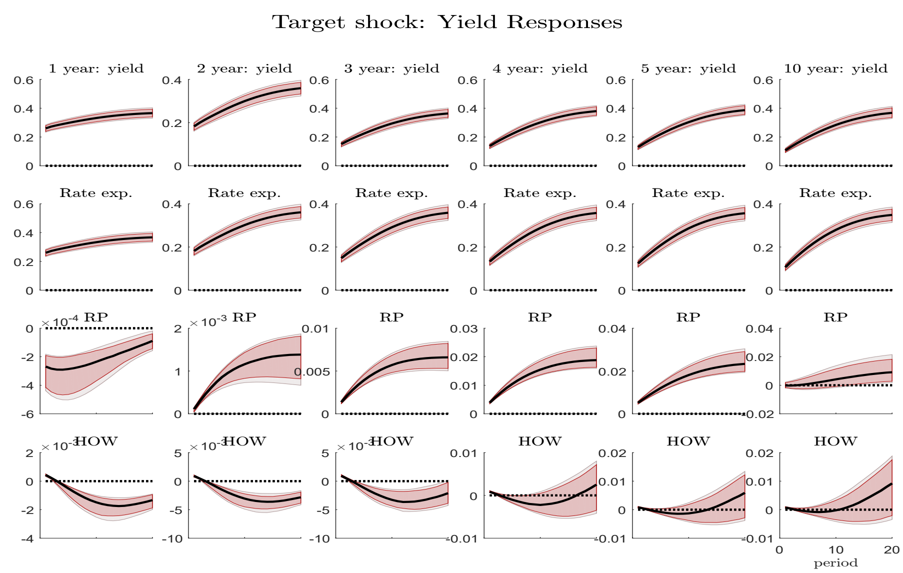
(a) Response of non-financial variables to monetary policy rule shock, dispersed information. Black line represents the median response. Dark red indicates 80% posterior credible set. Lighter bands indicate the 90% posterior credible set. Inflation and interest rates are shown in annualized percentage points. The output gap is shown in percentage points. v is scaled by 100.



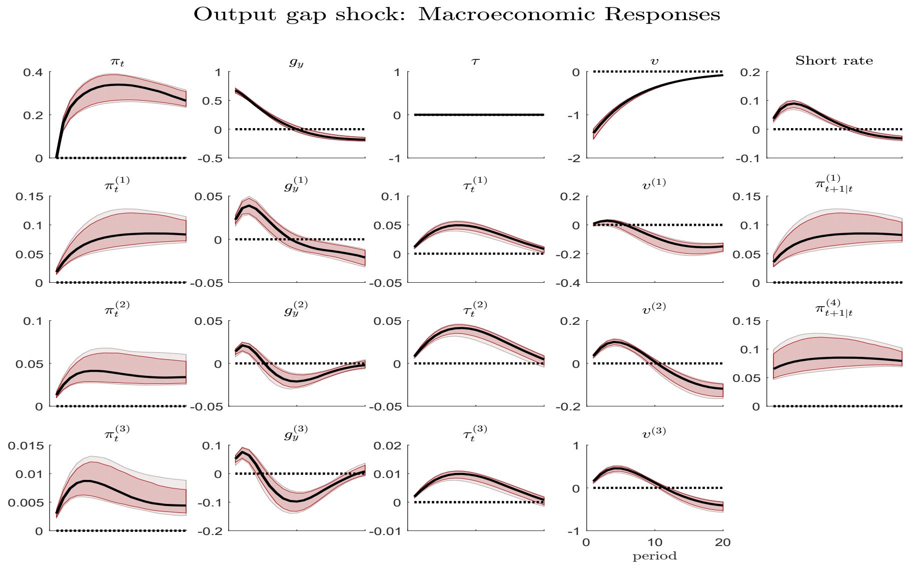
(b) Response of financial variables to monetary policy rule shock, dispersed information. Black line represents the median response. Dark red indicates 80% posterior credible set. Lighter bands indicate the 90% posterior credible set. Inflation and interest rates are shown in annualized percentage points. The output gap is shown in percentage points. v is scaled by 100.



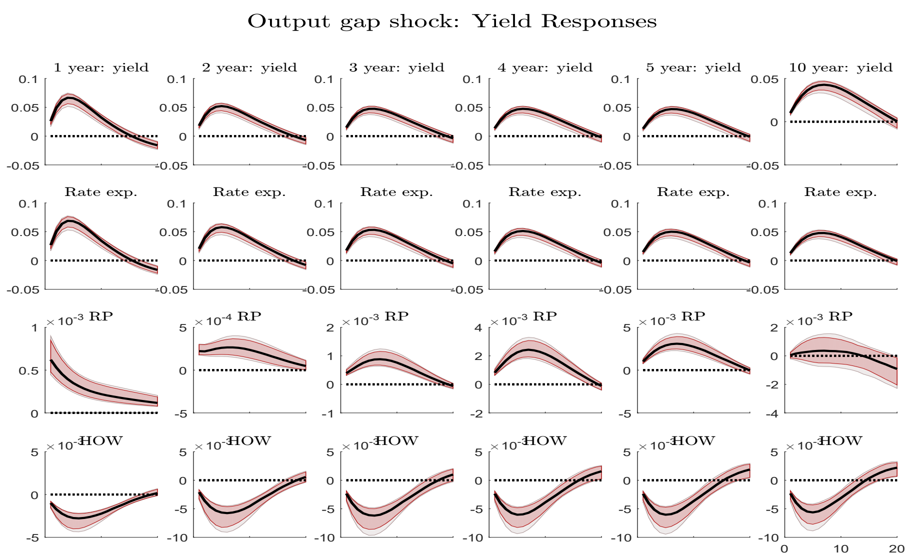
(a) Response of non-financial variables to inflation target shock, dispersed information. Black line represents the median response. Dark red indicates 80% posterior credible set. Lighter bands indicate the 90% posterior credible set. Inflation and interest rates are shown in annualized percentage points. The output gap is shown in percentage points. v is scaled by 100.



(b) Response of financial variables to inflation target shock, dispersed information. Black line represents the median response. Dark red indicates 80% posterior credible set. Lighter bands indicate the 90% posterior credible set. Inflation and interest rates are shown in annualized percentage points. The output gap is shown in percentage points. v is scaled by 100.

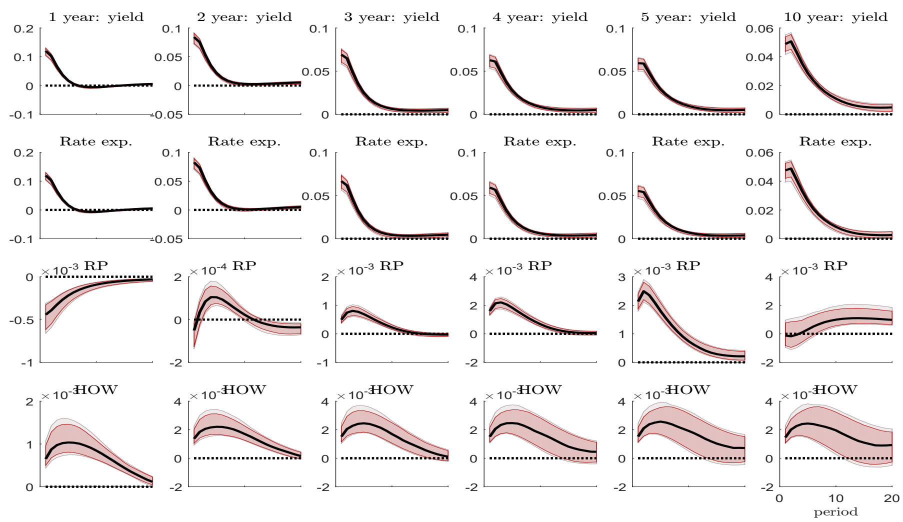


(a) Response of non-financial variables to output gap shock, dispersed information. Black line represents the median response. Dark red indicates 80% posterior credible set. Lighter bands indicate the 90% posterior credible set. Inflation and interest rates are shown in annualized percentage points. The output gap is shown in percentage points. v is scaled by 100.



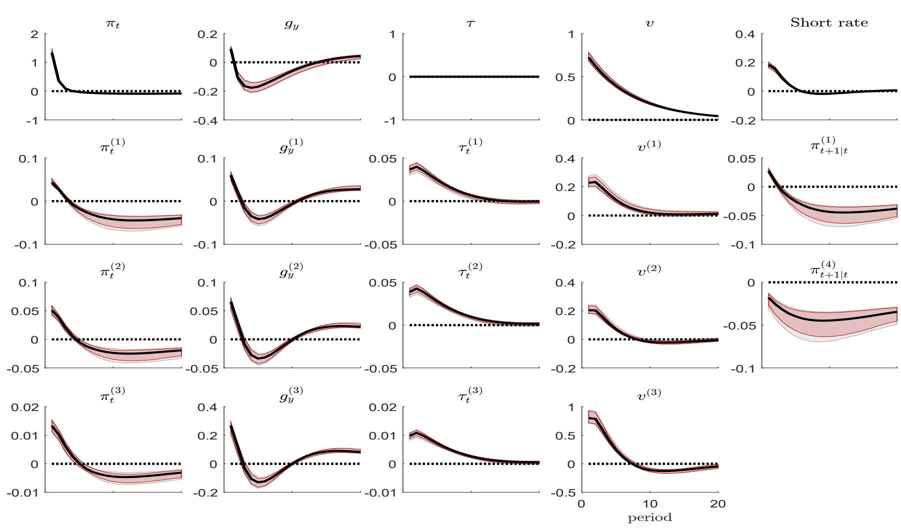
(b) Response of financial variables to output gap shock, dispersed information. Black line represents the median response. Dark red indicates 80% posterior credible set. Lighter bands indicate the 90% posterior credible set. Interest rates are shown in annualized percentage points.

Inflation shock: Yield Responses



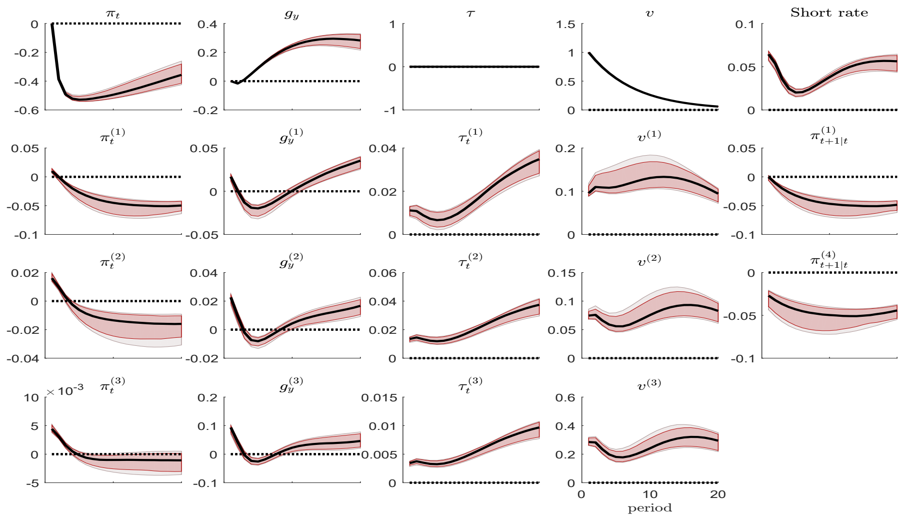
(a) Response of non-financial variables to inflation shock, dispersed information. Black line represents the median response. Dark red indicates 80% posterior credible set. Lighter bands indicate the 90% posterior credible set. Inflation and interest rates are shown in annualized percentage points. The output gap is shown in percentage points. v is scaled by 100.

Inflation shock: Macroeconomic Responses



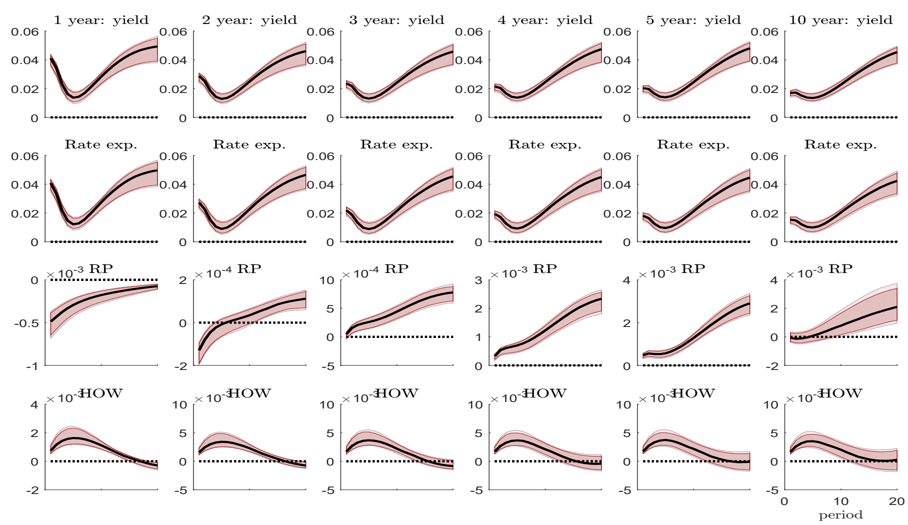
(b) Response of financial variables to inflation shock, dispersed information. Black line represents the median response. Dark red indicates 80% posterior credible set. Lighter bands indicate the 90% posterior credible set. Interest rates are shown in annualized percentage points.

Risk shock: Macroeconomic Responses



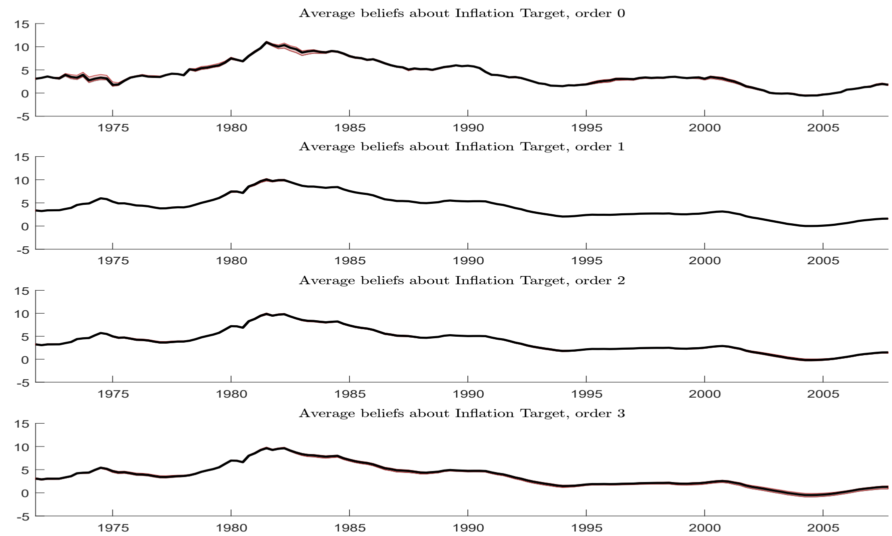
(a) Response of non-financial variables to risk shock, dispersed information. Black line represents the median response. Dark red indicates 80% posterior credible set. Lighter bands indicate the 90% posterior credible set. Inflation and interest rates are shown in annualized percentage points. The output gap is shown in percentage points. v is scaled by 100.

Risk shock: Yield Responses

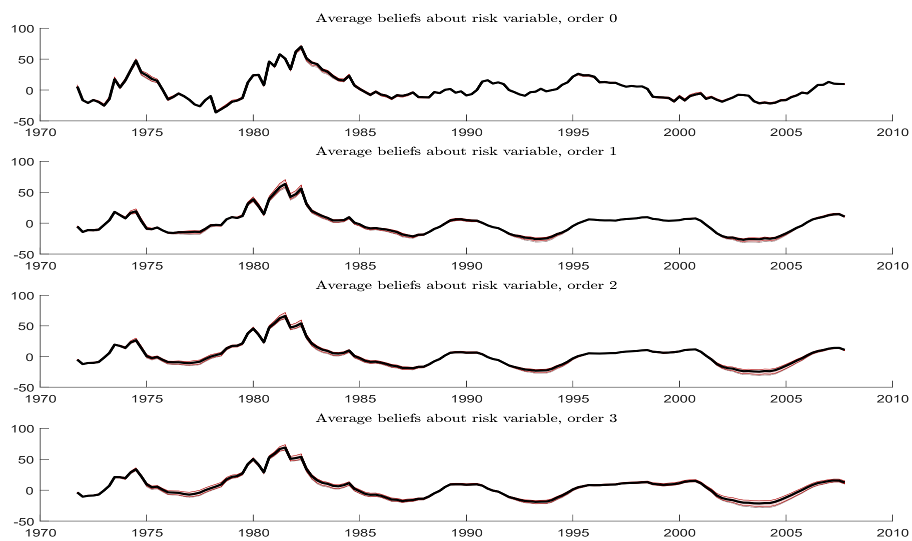


(b) Response of financial variables to risk shock, dispersed information. Black line represents the median response. Dark red indicates 80% posterior credible set. Lighter bands indicate the 90% posterior credible set. Interest rates are shown in annualized percentage points.

1.G.2 State Estimates and Yield Decompositions



(a) Filtered estimate of inflation target and first three orders of expectation an annualized percent, dispersed information model.



(b) Filtered estimate of risk variable and first three orders of expectation, dispersed information model.

1.G.3 Yield Decompositions at Posterior Mode

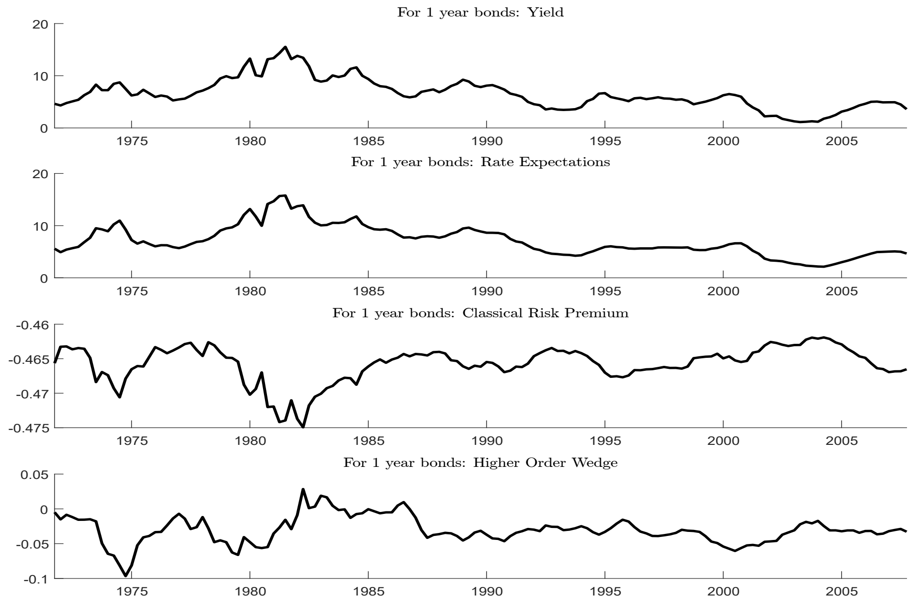


Figure 1.21: Decomposition of 1 year yields, posterior mode of dispersed information model

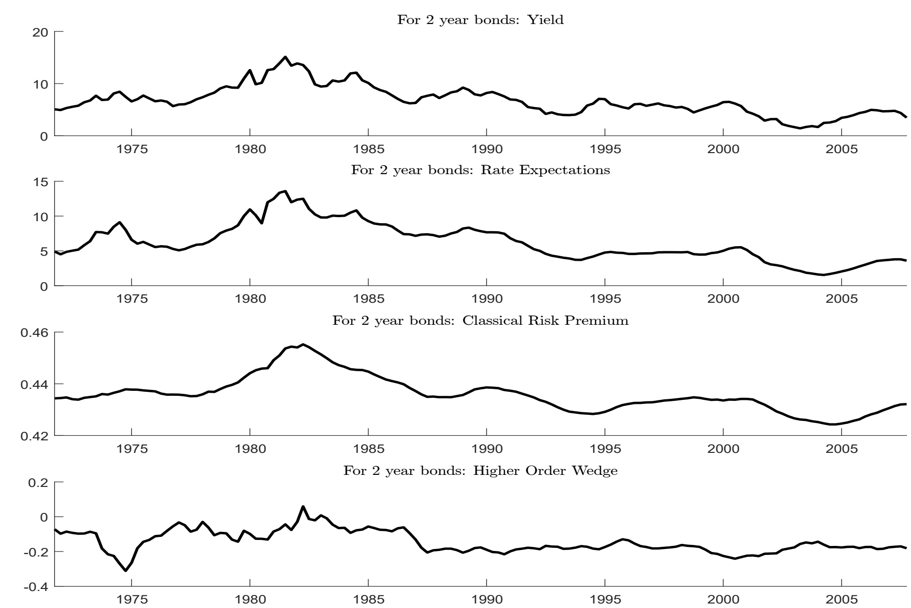


Figure 1.22: Decomposition of 2 year yields, posterior mode of dispersed information model

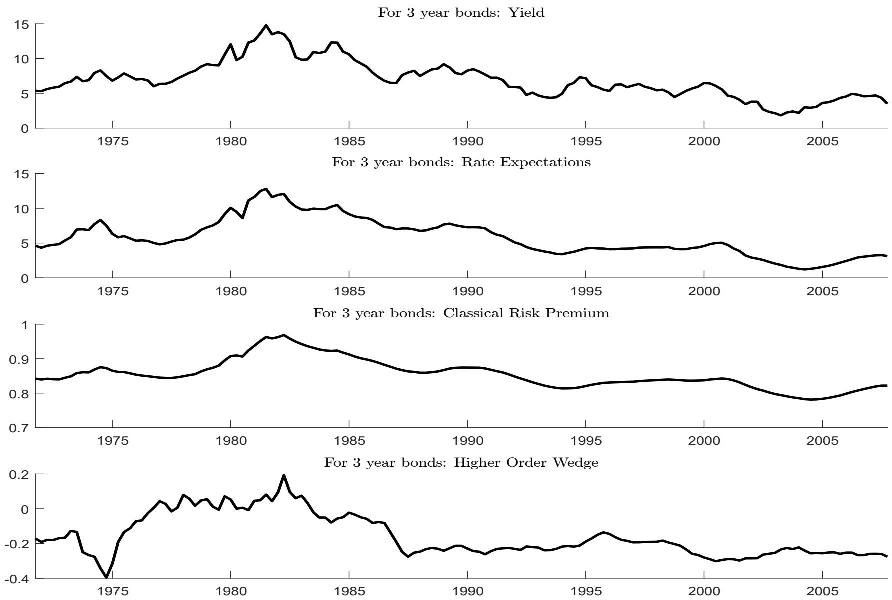


Figure 1.23: Decomposition of 3 year yields, posterior mode of dispersed information model

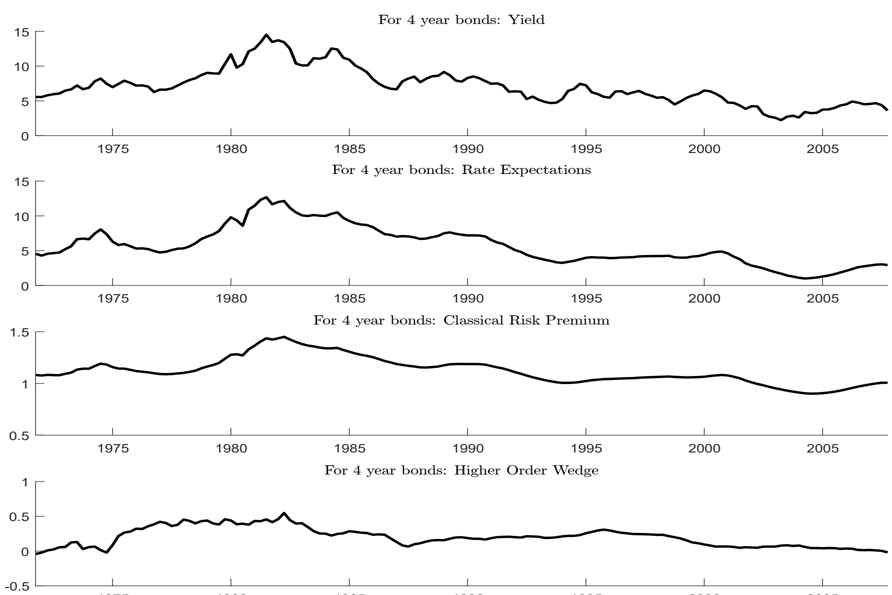


Figure 1.24: Decomposition of 4 year yields,posterior mode of dispersed information model

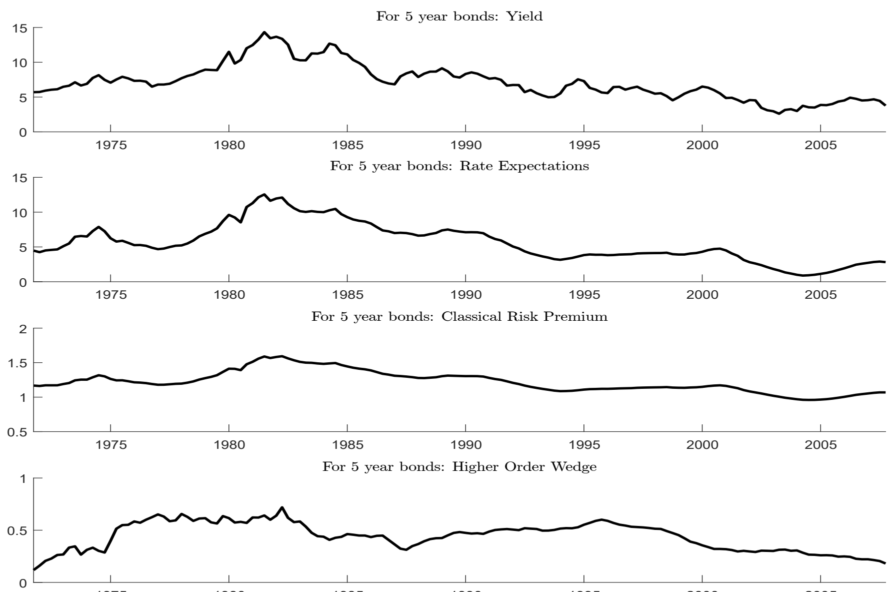
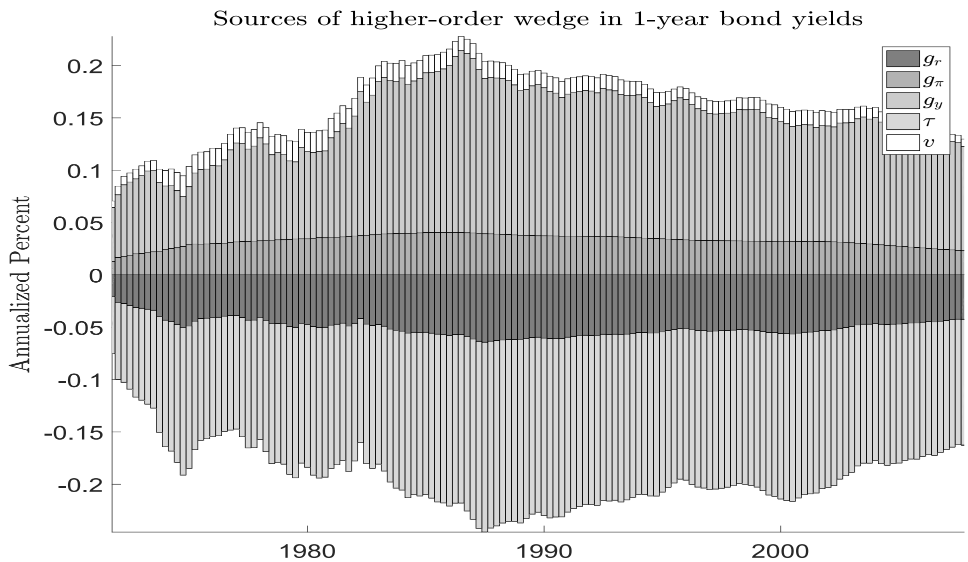
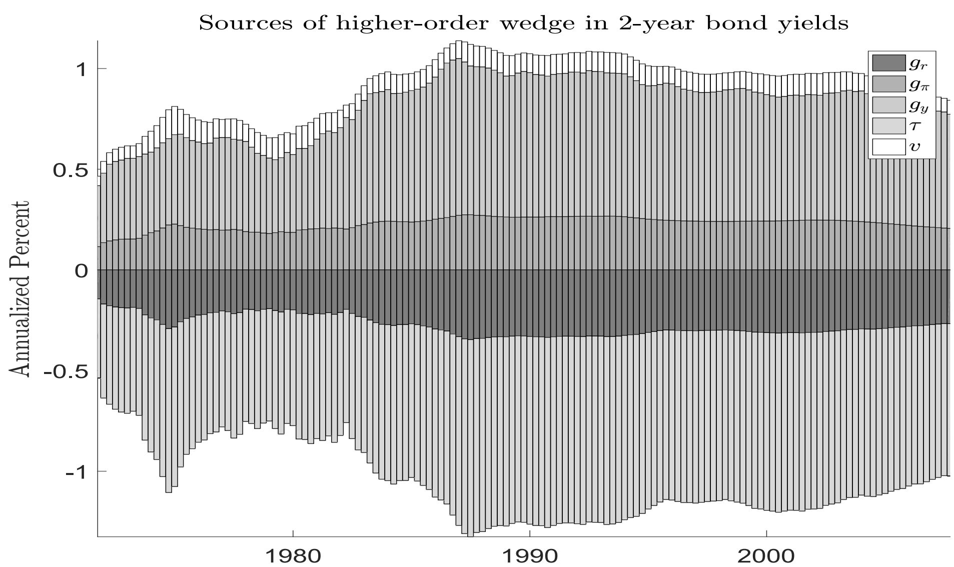


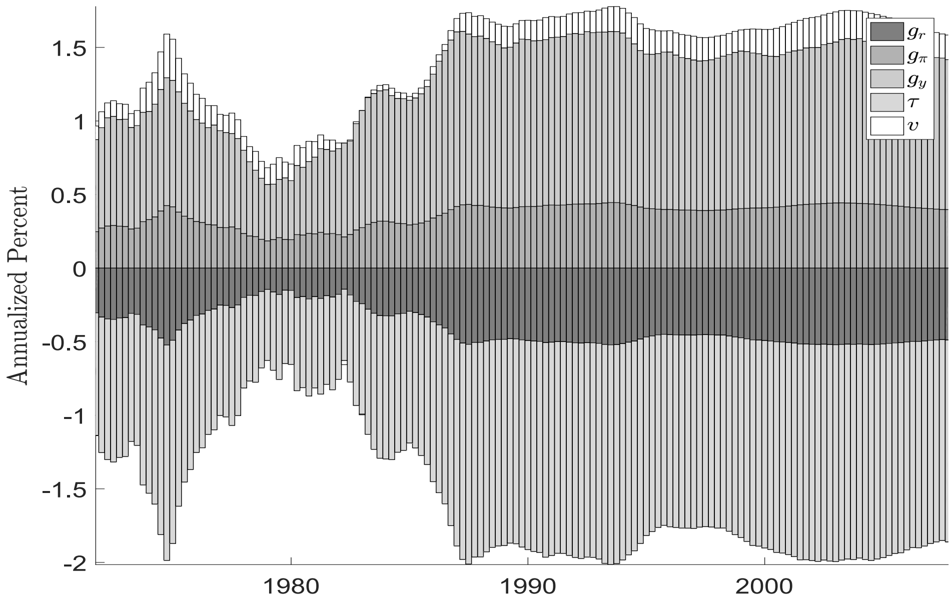
Figure 1.25: Decomposition of 5 year yields, posterior mode of dispersed information model

1.G.4 Wedge Decompositions at Posterior Mode

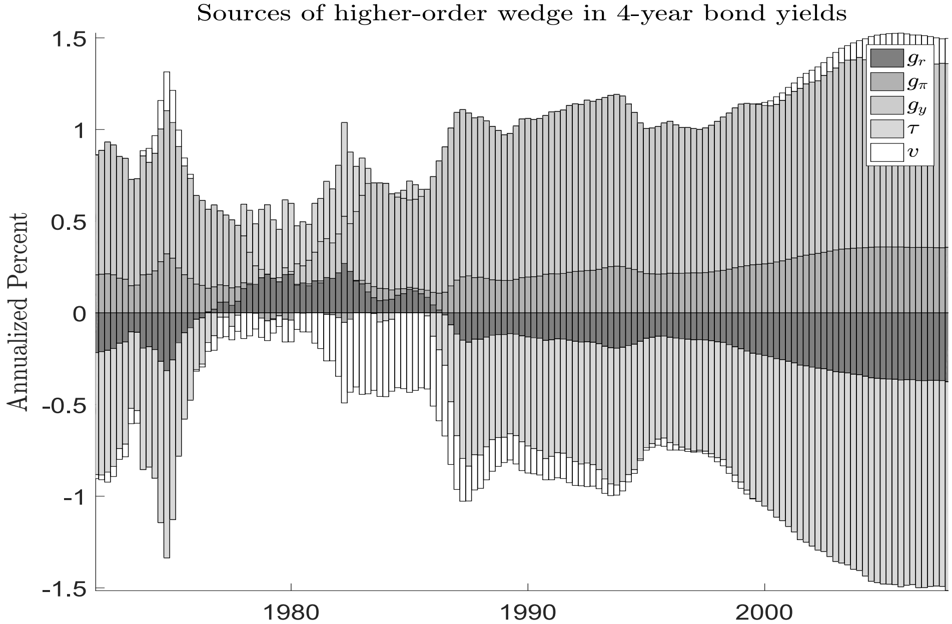


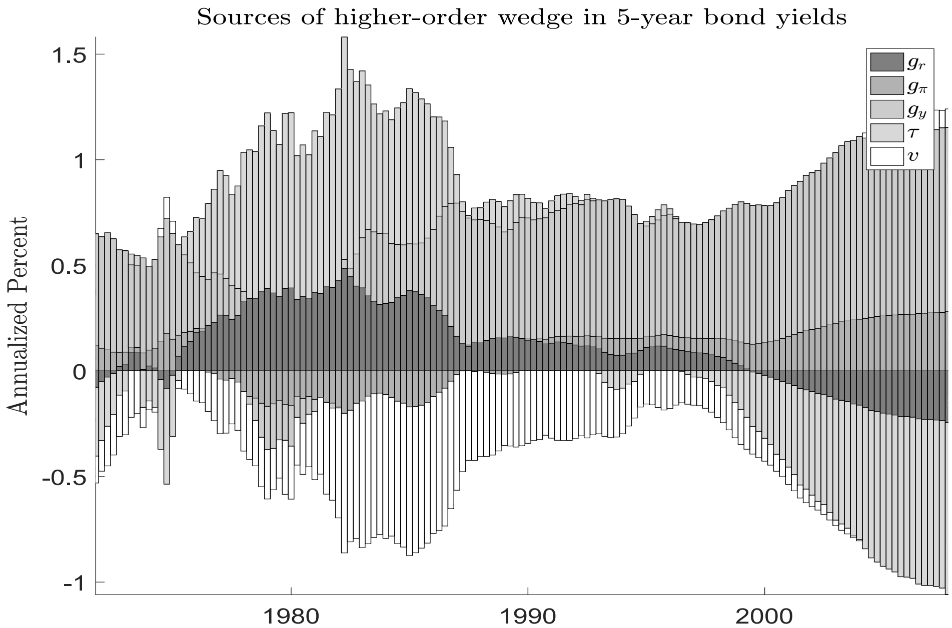


Sources of higher-order wedge in 3-year bond yields



Sources of higher-order wedge in 4-year bond yields





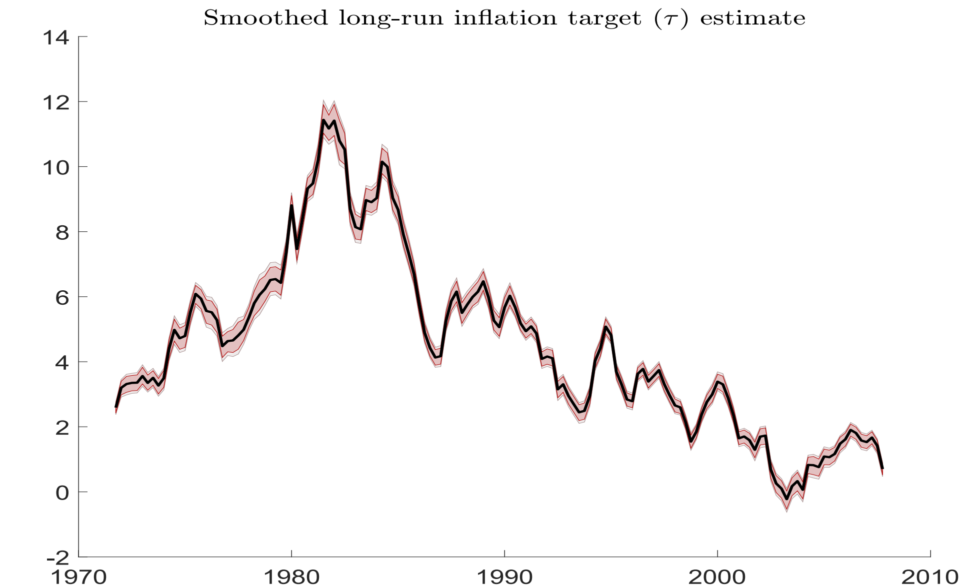
Sources of higher-order wedge in 5-year bond yields
Annualized Percent
1.5
1
0.5
0
-0.5
-1
1980
1990
2000
g_r
g_π
g_y
τ
v

1.H Results, full information model

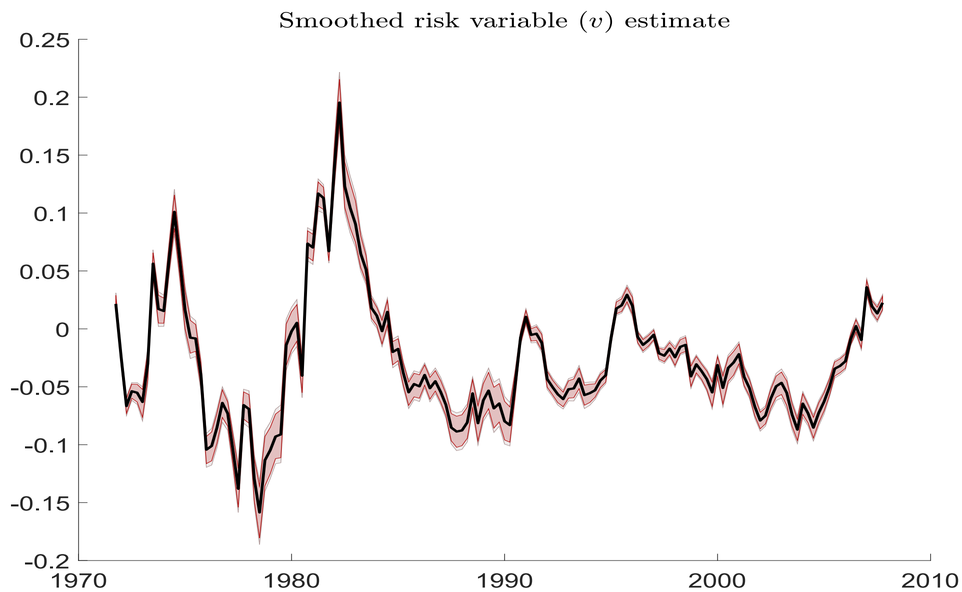
Table 1.8: Posterior estimates, full information Model

	Mode	Mean	Median	5th percentile	95th percentile	Std. Dev
ϕ_r	0.5349	0.5373	0.5350	0.5154	0.5727	0.0170
ϕ_π	0.1771	0.1760	0.1702	0.1098	0.2522	0.0435
ϕ_y	0.1178	0.1114	0.1106	0.0908	0.1353	0.0133
ϕ_v	0.0283	0.0221	0.0221	0.0128	0.0314	0.0058
σ_r	0.0020	0.0021	0.0021	0.0019	0.0023	0.0001
ρ_{yr}	-0.9946	-0.9278	-0.9334	-1.1513	-0.6789	0.1454
$\rho_{y\pi}$	-0.3899	-0.4031	-0.4077	-0.5526	-0.2503	0.0943
ρ_{yy}	0.9525	0.9277	0.9266	0.8835	0.9730	0.0272
ρ_{yv}	-0.0013	-0.0149	-0.0131	-0.0368	-0.0011	0.0109
$\sigma_{y\pi}$	0.2903	0.3863	0.3868	0.1419	0.6309	0.1509
$\sigma_{y\tau}$	2.6851	2.6387	2.6360	2.2813	2.9877	0.2174
σ_y	0.0066	0.0068	0.0068	0.0061	0.0075	0.0004
σ_τ	0.0012	0.0012	0.0012	0.0010	0.0013	0.0001
$\rho_{\pi r}$	0.8326	0.8344	0.8280	0.7050	1.0022	0.0831
$\rho_{\pi\pi}$	0.4024	0.4059	0.4057	0.3579	0.4558	0.0299
$\rho_{\pi y}$	-0.1750	-0.1787	-0.1774	-0.2191	-0.1440	0.0235
$\rho_{\pi v}$	-0.0842	-0.0832	-0.0822	-0.1005	-0.0706	0.0088
$\sigma_{\pi\tau}$	-0.1585	-0.1803	-0.1836	-0.3356	-0.0191	0.0956
σ_π	0.0040	0.0040	0.0040	0.0036	0.0044	0.0003
ρ_{vv}	0.8610	0.8550	0.8533	0.8321	0.8931	0.0177
σ_{vr}	9.8381	9.7069	9.7540	9.2463	9.9673	0.2172
$\sigma_{v\pi}$	2.1456	1.9804	2.0660	1.2514	2.5665	0.4220
σ_{vy}	-2.1201	-2.1669	-2.1473	-2.6129	-1.7968	0.2609
$\sigma_{v\tau}$	0.8711	0.8384	0.7641	0.1380	1.6089	0.4336
λ_r	1.2591	1.1900	1.2892	-0.3004	2.2988	0.7652
λ_π	-4.3221	-4.4279	-4.7589	-7.6234	0.1405	2.3402
λ_y	-0.6214	-0.2257	-0.4712	-2.4933	2.3231	1.5115
λ_τ	-0.1198	-0.1123	-0.1139	-0.2503	0.0229	0.0960
λ_r^x	18.3449	18.5892	18.6880	13.1336	23.8793	3.3348
λ_π^x	-76.2955	-77.1364	-76.8624	-80.6353	-73.8037	2.0341
λ_y^x	-17.4764	-22.7287	-21.9104	-33.4025	-17.1115	4.4454
λ_τ^x	0.1536	0.0564	0.0529	-1.7520	1.8245	1.0744
$\widetilde{\sigma}_4$	0.0012	0.0012	0.0012	0.0011	0.0014	0.0001
$\widetilde{\sigma}_8$	0.0011	0.0011	0.0011	0.0010	0.0012	0.0001
$\widetilde{\sigma}_{12}$	0.0009	0.0010	0.0010	0.0008	0.0011	0.0001
$\widetilde{\sigma}_{16}$	0.0010	0.0009	0.0009	0.0008	0.0011	0.0001
$\widetilde{\sigma}_{20}$	0.0010	0.0010	0.0010	0.0008	0.0011	0.0001
$\widetilde{\sigma}_{40}$	0.0014	0.0014	0.0014	0.0012	0.0016	0.0001
$\widetilde{\sigma}_\pi^1$	0.0032	0.0032	0.0032	0.0031	0.0032	0.0000
$\widetilde{\sigma}_\pi^4$	0.0039	0.0039	0.0039	0.0038	0.0039	0.0000

1.H.1 State estimates and Yield decompositions



(a) Smoothed estimate of inflation target, full information



(b) Smoothed estimate of risk variable, full information

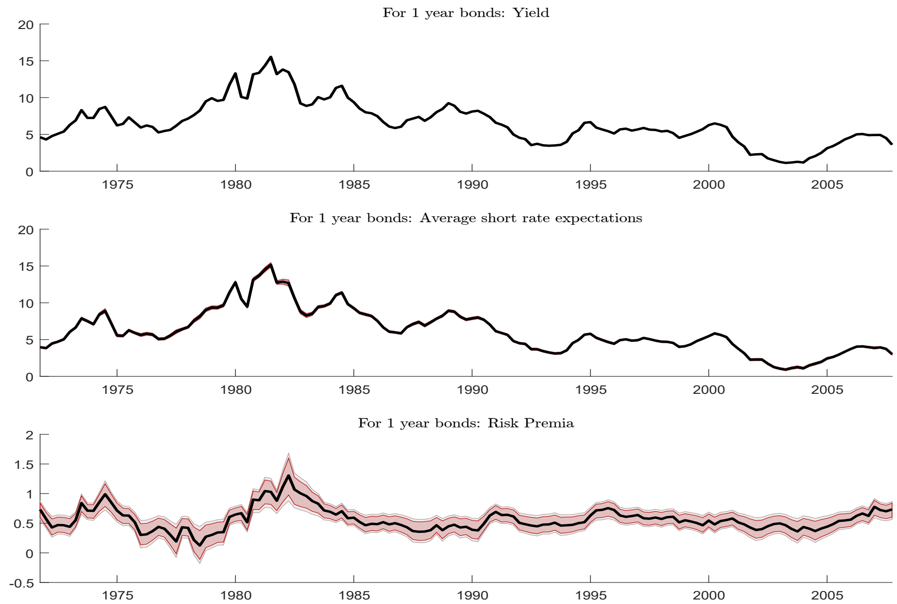


Figure 1.27: Decomposition of 1 year yields, full information

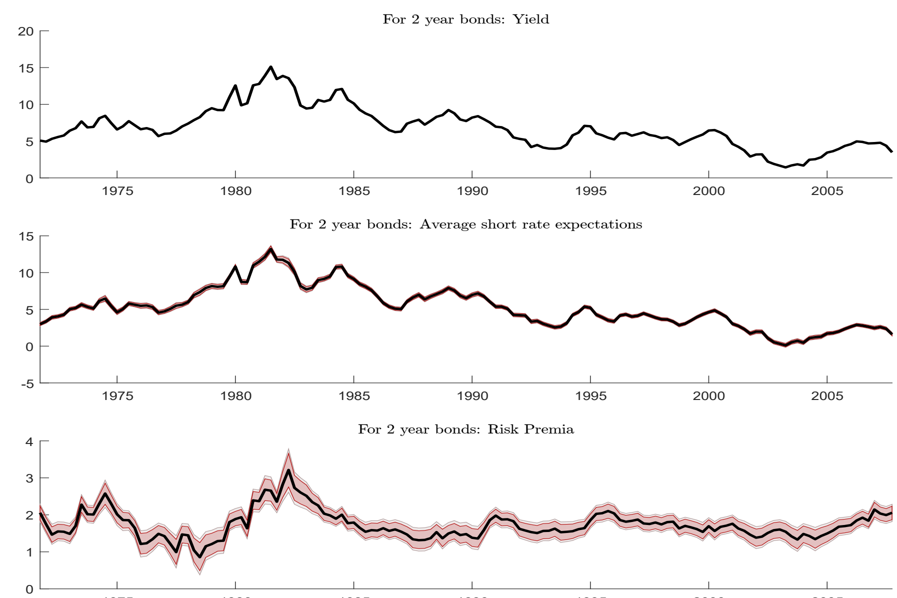


Figure 1.28: Decomposition of 2 year yields, full information

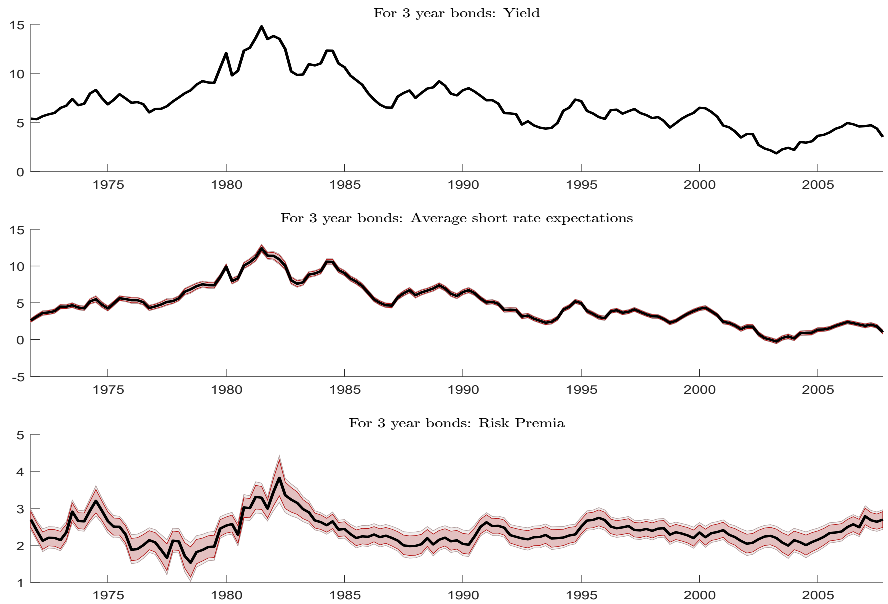


Figure 1.29: Decomposition of 3 year yields, full information

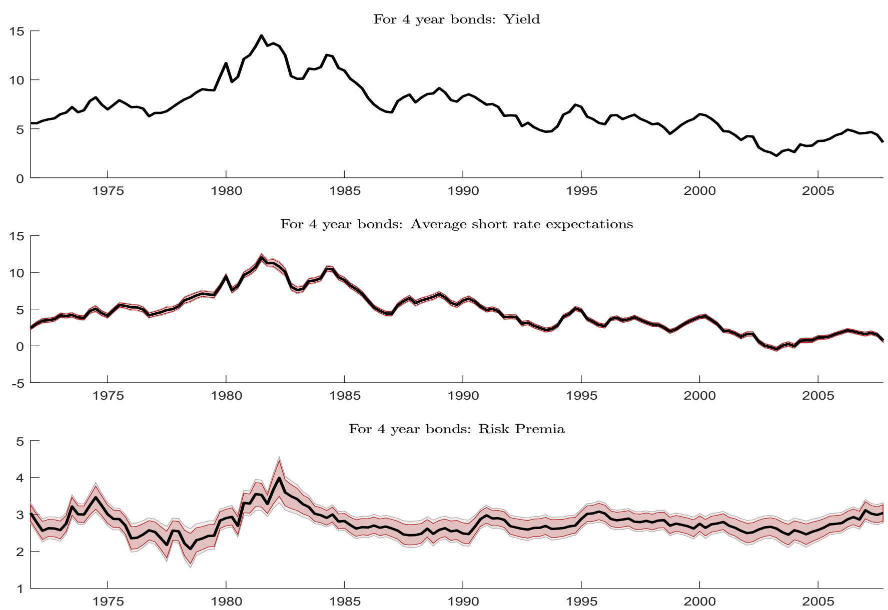


Figure 1.30: Decomposition of 4 year yields, full information

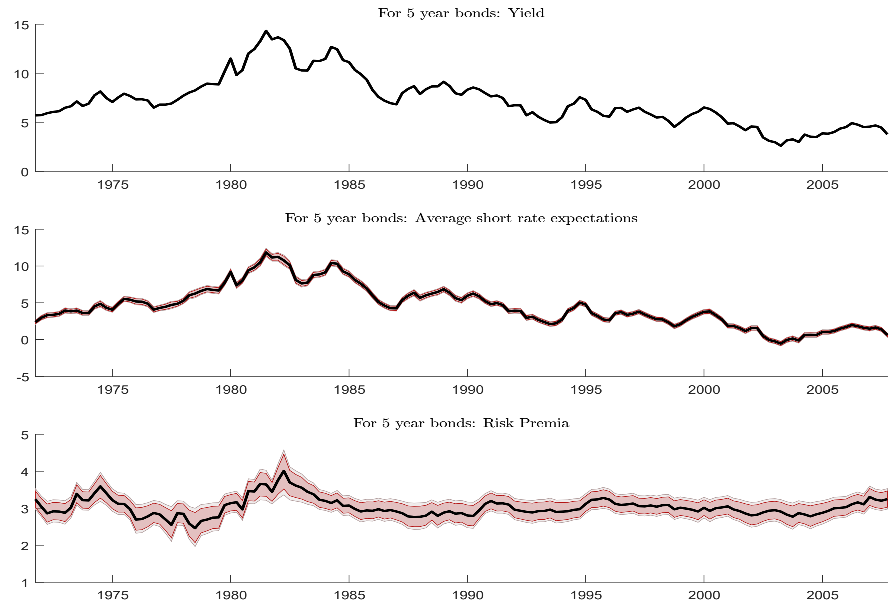


Figure 1.31: Decomposition of 5 year yields, full information

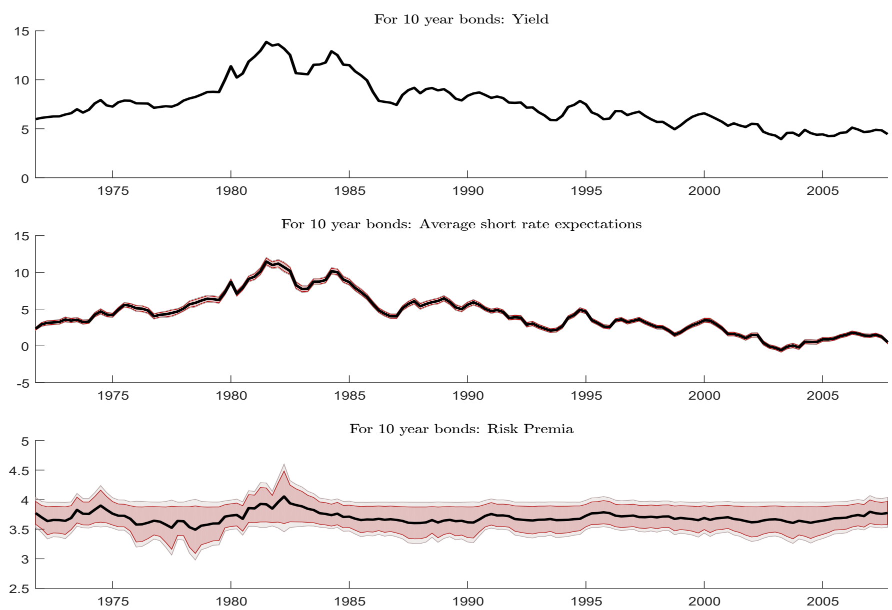
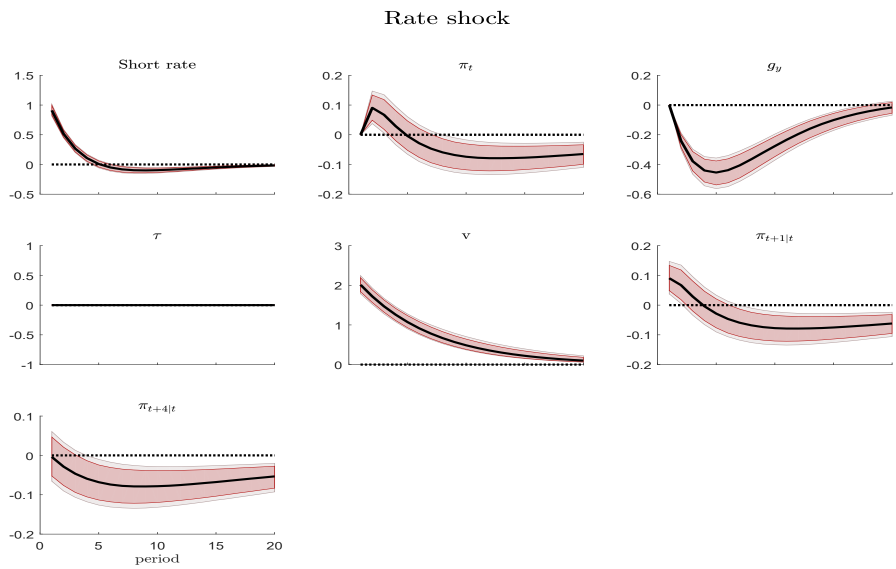
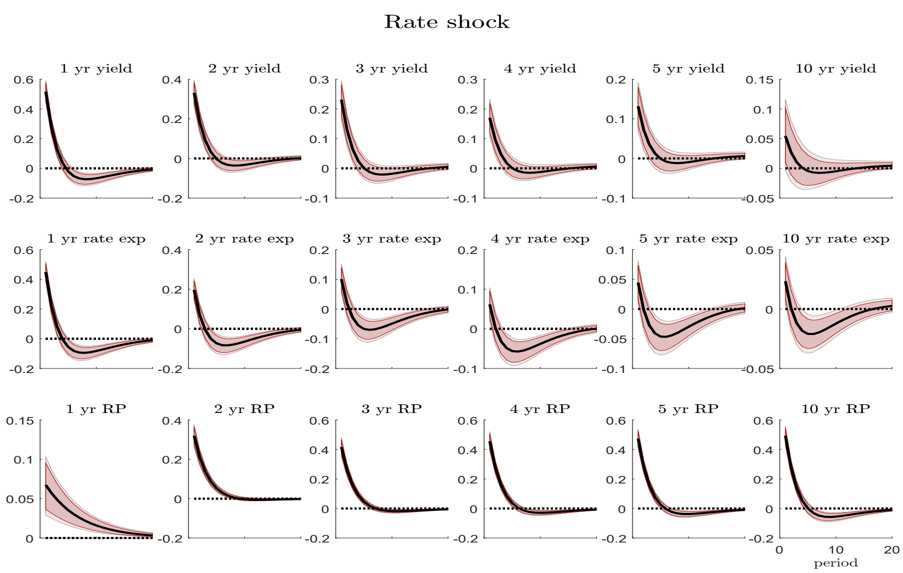


Figure 1.32: Decomposition of 10 year yields, full information

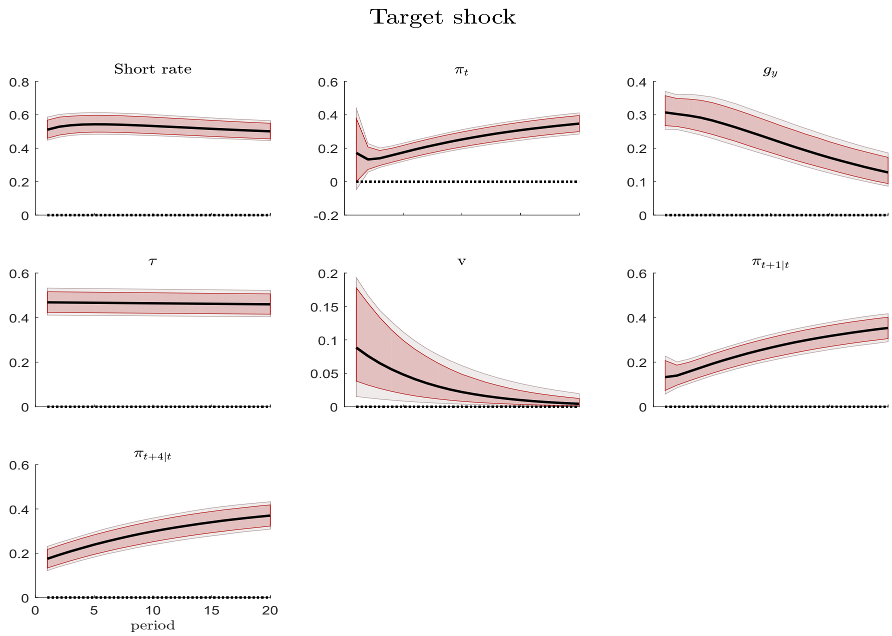
1.H.2 Impulse Responses



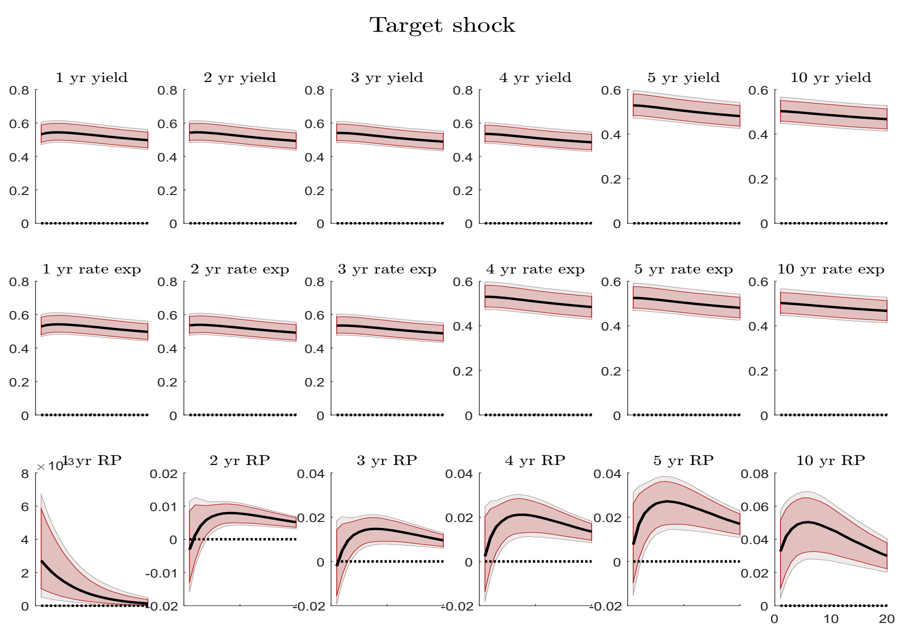
(a) Response of non-financial variables to monetary policy rule shock, full information



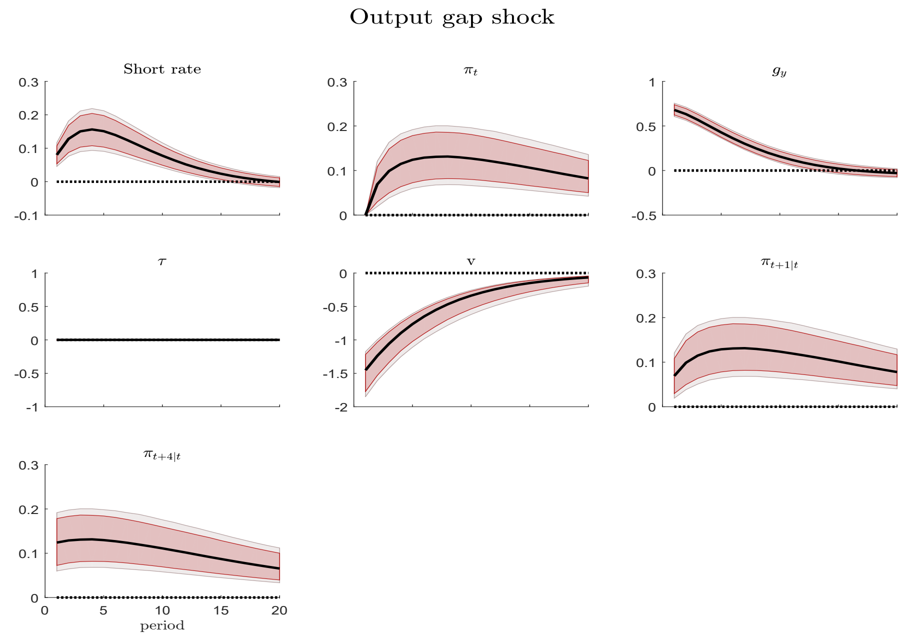
(b) Response of financial variables to monetary policy rule shock, full information



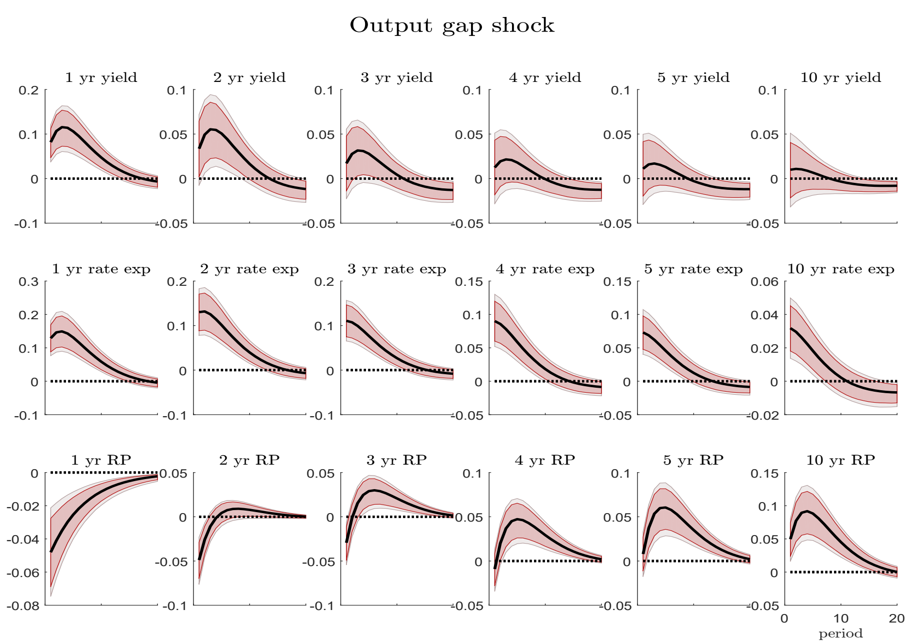
(a) Response of non-financial variables to inflation target shock, full information



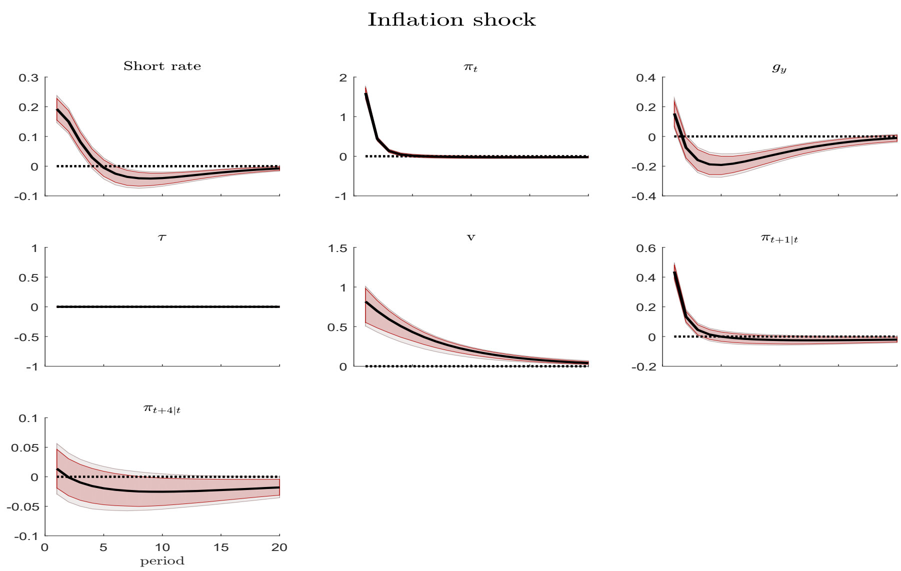
(b) Response of financial variables to inflation target shock, full information



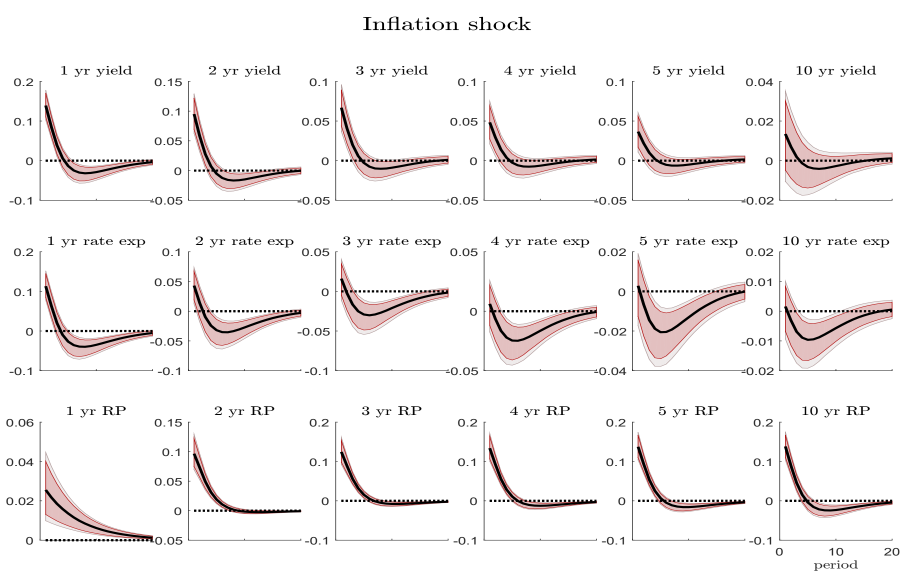
(a) Response of non-financial variables to output gap shock, full information



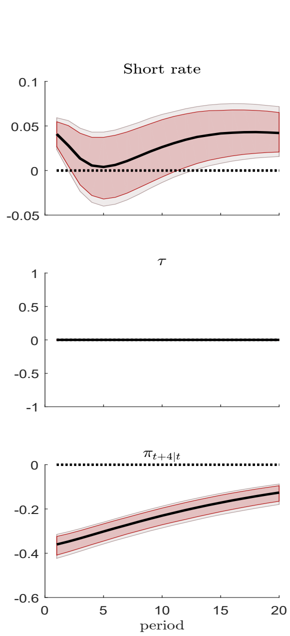
(b) Response of financial variables to output gap shock, full information



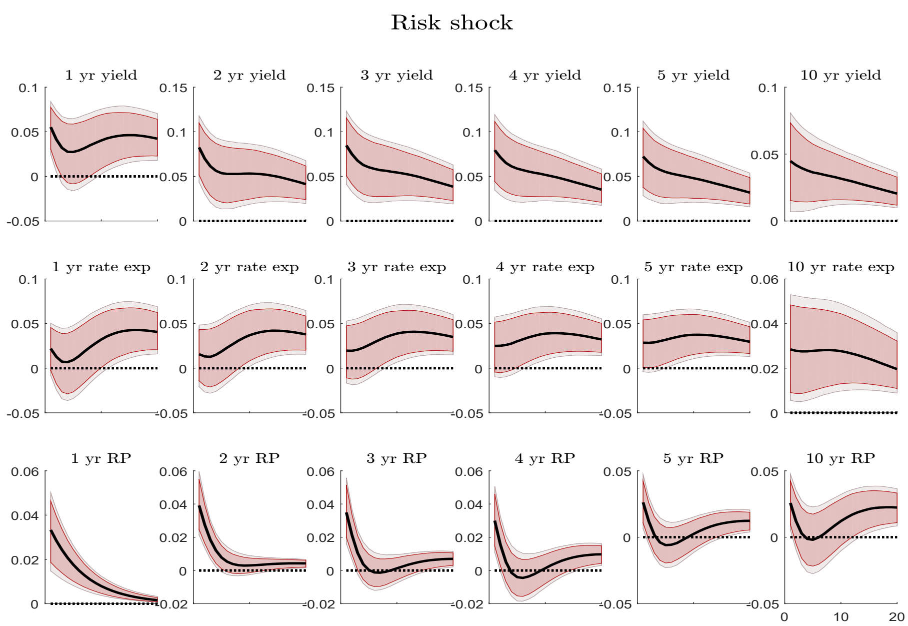
(a) Response of non-financial variables to inflation shock, full information



(b) Response of financial variables to inflation shock, full information



(a) Response of non-financial variables to risk shock, full information



(b) Response of financial variables to risk shock, full information

Chapter 2

Your Guess is as Good as Mine: Central Bank Information and Monetary Policy

Abstract

I document that Federal Reserve staff and professional forecasters have non-overlapping information in their forecasts and that the Federal Reserve staff are not systematically more accurate than the median professional. In light of this evidence, I develop a dynamic New Keynesian model where firms and the central bank have incomplete, non-overlapping information sets and examine its ability to match data on business cycle variables and inflation forecasts simultaneously. In the model, price setting firms learn about persistent aggregate shocks via both idiosyncratic signals and the central bank's policy rate; firms and the central bank must form higher-order expectations about each others' beliefs. The calibrated model is able to capture the auto- and cross-correlation structure of inflation forecasts of central bankers. The degree and type of noise in the central bank's forecasting process matters for the monetary transmission mechanism and the model's ability to qualitatively match business cycle comovements.

I gratefully acknowledge comments from Ryan Chahrour, and suggestions from participants at the Boston College Macroeconomics Lunch and Dissertation Workshops, and helpful conversation with Alexandre Kohlhas.

2.1 Introduction

Many economic environments are characterized by both dispersed information and strategic interdependence. The manager of a firm does not know what actions her competitors will take, nor does she observe their beliefs. But her rivals' beliefs affect the prices they set and thus her optimal price. She can infer their beliefs using her own information; but, knowing her competitors are doing the same thing, she must try to understand what they expect about what she expects, and so on.

In this type of environment, public signals can play an important role: they inform agents about the underlying state of the world and make coordination easier by introducing a common component of beliefs. One important public signal is monetary policy, and a number of recent papers have discussed what firms learn from interest rate changes. In a full-information context, Nakamura and Steinsson (2013) find evidence of a "Fed information effect" - monetary policy changes appear to influence beliefs about both the stance of policy and future economic fundamentals. Tang (2015) shows that, when agents have (common) imperfect information, inflation expectations can respond positively to a contractionary monetary policy shock. This effect depends crucially on the form of the central bank's policy response function. These papers emphasize the information role of public signals, rather than the coordination role; firms in these models all share a common imperfect information set. By contrast, Melosi (2017) estimates a model where price-setting firms observe idiosyncratic signals and also learn from the monetary policy rate. The latter communicates information about the aggregate technology and demand shocks, as well as exogenous deviations from the Taylor rule. Hence, agents observing rate hikes put posterior weight on both that the central bank sees inflation today, but also knows other firms have extracted a positive inflation signal from the policy rate movement. Knowing that others believe prices will rise, the optimal action is to raise prices. He finds that the signaling channel is empirically important for understanding why interest rate increases did not head off inflation in the 1970s.

An potentially important feature of the signaling channel of monetary policy is that, like firms, central bankers must form expectations of current and future economic conditions when setting policy. Hence, the signal drawn from monetary policy actions must depend on the quality of information accessed by the central banker. If central bank information is poor, then firms will choose to ignore it and it will cease to serve either an informational or communication role. In this paper, I document that central bank forecasts contain different information than firms' forecasts, but are not systematically more accurate (at least, when considering forecasts of current-quarter inflation). Based on this finding, I examine the ability of a model with endogenous central bank forecasts and dispersed firm information to explain the data. In Melosi (2017), the central bank's policy rule is set according to the true realizations of the output gap and inflation, plus three exogenous shocks which capture persistent measurement error by the central bank and persistent deviations from the Taylor rule; expectational error is another shock. By contrast, in this paper, I extend Melosi's model by explicitly modeling the inference problem of the central bank. In particular, I endogenize persistent central bank error by explicitly modeling how the central bank forms expectations. In order to forecast inflation, the central banker must form an expectation not only about exogenous shocks, but also what firms belief about fundamentals, the central banker's actions, and the beliefs of other firms. The theoretical extension follows Kohlhas (2015). His focus is on understanding the theoretical properties of central bank information disclosure for aggregate uncertainty and analyzing disclosure of information, and he makes a number of simplifying assumptions to make the analysis more tractable. As a result, his model has difficulty matching facts on the dispersion of private sector forecasts of inflation, their accuracy relative to Federal Reserve forecasts, and he does not examine the ability of the model to match realize data. By contrast, I examine the ability of a model with signal extraction by the central bank to match data on forecasts and the macroeconomy simultaneously. This is important for both understanding the positive relevance of the central bank's expectations formation process and how central bankers' information quality affects their efforts to stabilize the economy. Endogenizing the belief formation process of central banks also has implications for optimal implementable policies (Svensson and Woodford (2004a), Boehm and House (2014)).

I show that while calibrated versions of the model can quite successfully match unconditional second moments of forecasts and (to an extent) inflation, it has a more difficult time matching business cycle co-movements between measured total factor productivity (TFP), output, and interest rates. Based on the estimated results of Melosi (2014, 2017), the calibration features a low degree of price stickiness and relies on information frictions to generate endogenous persistence in inflation, but the model predicts that inflation is overly-correlated with changes in technology relative to the data. Assumptions about household information turn out to be critical for the ability of the model to match data on output; here, households have perfect information and demand shocks essentially result in noninflationary expansions in output. The calibrated results demonstrate that the exact transmission of shocks depends crucially on how the central banker responds, which in turn depends on how noisy her observations are. When the central banks' observations of technology are as noisy or noisier than firms, the model implies a negative correlation between output and TFP rather than the weak, positive correlation observed in the data and has difficulty generating inflation in response to positive demand shocks. In general, the qualitative (and quantitative) reaction of inflation to exogenous shocks depends critically on the central banker's inference problem. Endogenously matching the persistence of central bank beliefs requires noisiness in their observations of macroeconomic fundamentals, but too much or the "wrong kind" of noise hurts the ability of the simple New Keynesian model to match the data. The results suggest that to match the cross- and auto-correlations of output, inflation, interest rates, TFP, and firm and central bank inflation forecasts, central bankers must have no worse information about technology shocks than firms and similar (or perhaps slightly noisier) information about shocks to household demand.[1]

[1] A second, methodological lesson is that models with an "informationally large" agent and dispersed information across firms may find themselves caught between accuracy and the curse of dimensionality. I comment on at the end of the results section.

Relationship to the literature. As discussed above, this paper is most closely related to Melosi (2017) and Kohlhas (2015). The model in this paper also owes a genealogical debt to the literature on public information in games with dispersed information. Morris and Shin (2002) examine the effects of a public signal in a stylized "beauty contest" model. Woodford (2001) considers a similar model with dynamic inference. The approach in this paper is most closely related to the more recent literature on information in business cycle models. Of these, my paper is most close to that of Nimark (2008) and Melosi (2014) who estimate models with dispersed information among firms and show that New Keynesian models with dispersed information can successfully replicate a number of stylized facts about inflation and prices. These papers, however, do not feature a signaling channel. A closely related paper which takes a less parametric approach is Chahrour and Ulbricht (2017). Their interest is estimating information wedges, that is, deviations from full-information actions of firms, households, and the central bank, and then deriving an information structure that implements those wedges. Their method is highly flexible, which allows it to match aggregate data well, but they do not quantitatively examine whether model-implied beliefs are consistent with forecasts. I impose a stronger information structure which allows me to examine the implications for difference in information quality across fundamentals for both forecast data and realized macroeconomic data, and also to examine more closely the information contained in monetary policy rates.

A number of other papers also consider the importance of central bank information with less structure on beliefs or in a reduced-form context. For example, Orphanides and Williams (2009) consider the implications of imperfect central bank knowledge for monetary policy in a learning context where the central bank has imperfect knowledge of the model parameters rather than the state. Romer and Romer (2000) examine forecasts of the Federal Reserve Board staff and commercial forecasters. I re-examine their findings in the next section. More recently, Coibion and Gorodnichenko (2012b) document that forecast errors in Romer and Romer's data, as well as in surveys of household and professional forecasters, is inconsistent with full-information rational expectations. This paper complements these by exploring the ability of a textbook model with optimal inference to rationalize both business cycle data and implied beliefs. Svensson and Woodford (2004b) examine optimal policy in an asymmetric information environment, but one where firms have full information and the central bank does not.

The plan for the paper is as follows. In section 2.2, I document some facts about central bank and professional forecasters - in particular, that there appears to be evidence of a lack of information overlap and time-variation in relative forecasting performance. In section 2.3 and 2.4, I derive the model with dispersed firm and central bank information and discuss the solution technique. In section 2.5 I present and discusses quantitative results.

2.2 Central bank and firm information: Evidence from nowcasts

Most rational expectations models assume agents have complete information about the state of the world and the structure of the model up to the current period when forming beliefs. In such models, agents' "nowcasts" - i.e., their forecasts about within-period conditions - are perfect. Moreover, even imperfect information models often assume common information abd thus a degenerate forecast distribution. By contrast, a recent literature on information frictions emphasizes differences in beliefs, either because of infrequent and staggered updating ("sticky information" as in Mankiw and Reis (2006)) or because agents observe private idiosyncratic signals. The typical focus in this literature is on dispersion at the level of firms. In this section I document evidence that the information of Federal Reserve policymakers is imperfect, and is different from, but not systematically more accurate than, the median forecasters' information. Using forecasts to examine the information available to policymakers and firms is the approach of a number of papers following Romer and Romer (2000), and I use a similar methodology to theirs with an updated data set.[2] I focus on inflation forecasts from the published estimates of the Federal Reserve's Summary of Economic and Financial Conditions - the "Greenbook" - as well as data from the Survey of Professional Forecasters (SPF).

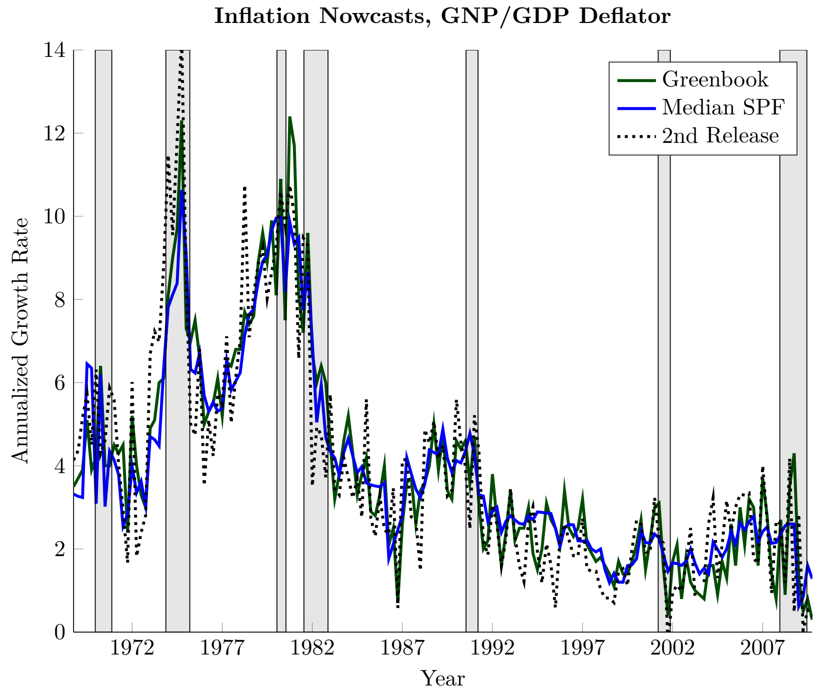
Figure 2.1a plots the time-series behavior of the median SPF and Greenbook nowcasts.[3] The median SPF and Greenbook forecasts generally are highly correlated, but often quite different from one another and from the data. Some of this may reflect the fact that taking the median forecast smooths out some of the variation. In figure 2.1b I replace the median forecast with the percentiles of the individual forecasts from the SPF. The Greenbook forecast usually falls within the center of the distribution from the SPF, but sometimes ventures quite far out of it, especially in the late 1970s. Figure 2.2 re-frames the first figure in terms of real-time forecasting error by the Federal Reserve Board staff (top) and the median SPF forecaster (bottom). Visual inspection suggests neither appears to have a particular advantage, particularly in the early part of the sample. Moreover, residuals appear to be serially correlated, inconsistent with a world where agents have full information rational expectations.

The time series plots do not tell much about whether the two sets of forecasts are informationally redundant. Continuing to focus on inflation, I follow Romer and Romer (2000) in undertaking two simple tests for systematic differences of information. First, I regress real time inflation on the nowcasts in the Greenbook and the median SPF forecaster. Essentially, this asks whether each are separately useful in forecasting current period inflation. While Romer and Romer (2000) find that private nowcasts are generally redundant of the Fed's nowcasts, I find that both are significant and positive, suggesting they both contain information about inflation that does not appear to overlap.[4] Moreover, while the R^2 of this regression is high, it is less than 1; some information about

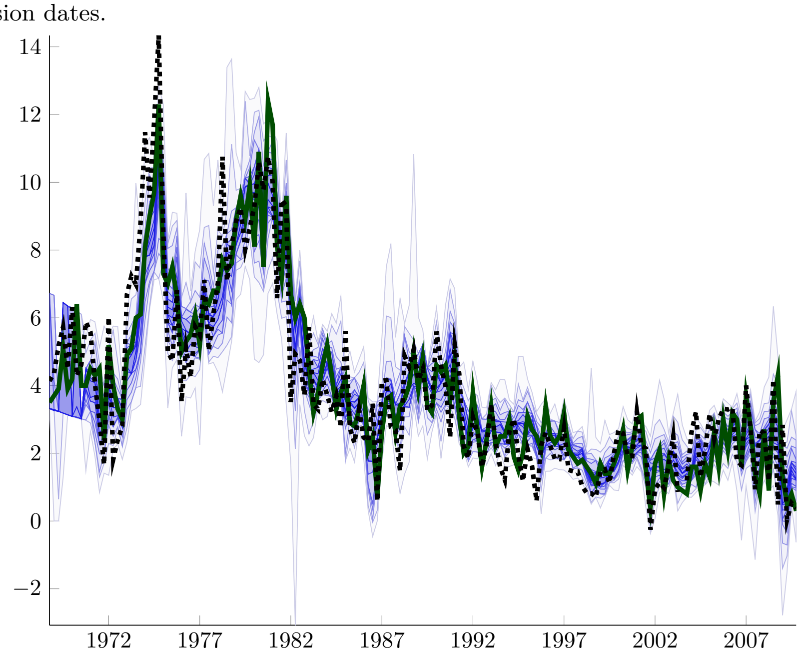
[2] Sims (2002) finds some supportive evidence for the idea that the Federal Reserve has a forecasting advantage that comes from private knowledge about its own actions or superior data collection. Hubert (2014) argues that the Federal Reserve has an informational advantage in forecasting inflation, but not output, owing to "superior information."

[3] Inflation is measured as the annualized percent change in the GNP/GDP deflator. As Romer and Romer (2000) note, the second revision ensures conceptual consistency between the "real time" data used in forecasting evaluation while avoiding the errors associated with first estimates; arguably, the more accurate first revision is the series actually targeted by forecasters.

[4] One reason for the differences between these results and that of Romer and Romer (2000) could be the duration of



(a) GDP/GNP annualized inflation "nowcasts" versus second release of NIPA data. Gray bands indicate NBER recession dates.



(b) GDP/GNP annualized inflation "nowcasts" versus second release of NIPA data. The dotted black line indicates inflation as reported in the second release of the National Income and Product Accounts. The solid green line indicates the Federal Reserve "Greenbook" forecast of inflation in the same quarter, while the blue bands indicate the distribution of individual forecasts for current-quarter inflation in the Survey of Professional Forecasts.

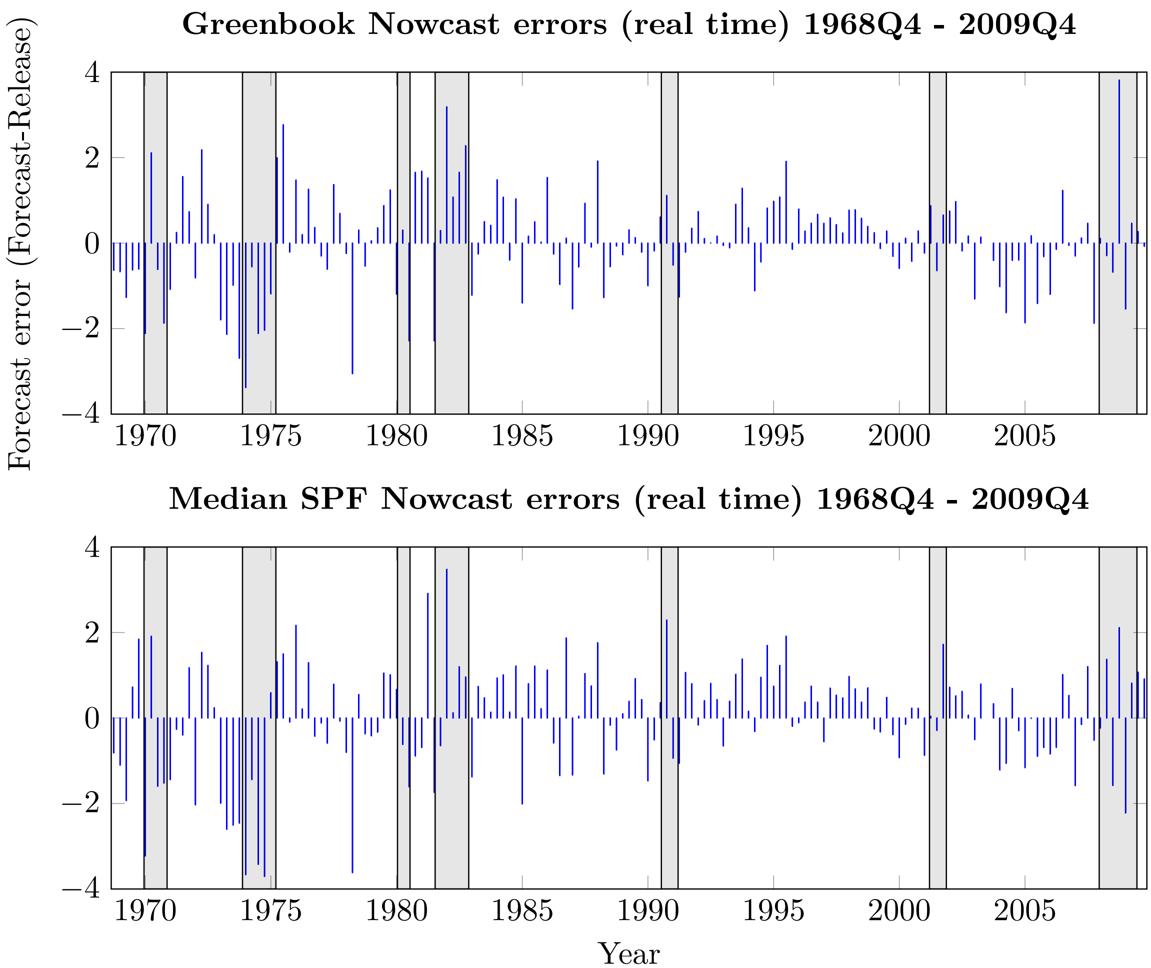


Figure 2.2: Real-time nowcasting error in the Greenbook and Survey of Professional Forecasters.

Table 2.1: Regression results for $\pi_t^{2nd} = \alpha + \beta_1 \pi_{t|t}^{GB} + \beta_2 \pi_{t|t}^{SPF} + \varepsilon_t$

α	-0.140
	(0.177)
β_1	0.705
	(0.115)
β_2	0.327
	(0.127)
Adjusted R^2	0.831

Standard errors in parentheses.

Table 2.2: Regression results for $SE_t^{GB} - SE_t^{SPF} = \gamma + \varepsilon_t$

γ	-0.268
	(0.151)

Standard error in parentheses

current inflation is not present in the forecasts of either the Federal Reserve staff or the median SPF forecaster.

Moreover, it is not the case that the Federal Reserve staff appear to be systematically better at nowcasting inflation than their SPF counterparts. Again, following Romer and Romer (2000), I regress the difference in squared errors in each quarter against a constant. A significant constant suggests a difference in the mean accuracy of the forecasts. As the results in table 2.2 show, this does not appear to be the case; although the point estimate is negative (implying greater accuracy of the Greenbook forecast), the result is not significant.

To generate some intuition behind this result, I re-run the regressions from table 2.2 using 10-year rolling windows. Perhaps surprisingly, I find that the only detectable forecasting advantage is in the first few quarters of the sample (i.e., in the late 1960s) and beginning in the early 1980s following the Volker disinflation. While the point estimate is generally negative, it is rarely statistically different from zero. This suggests that at best the Federal Reserve staff have an only somewhat better nowcasting track record than the median professional.

The results presented above suggest (1) the Federal Reserve staff's forecasts are not systematically better than those of the (median) respondents to the SPF and (2) The information that the SPF

the sample - their sample ends in 1991, while mine extends through the beginning of the Great Recession.

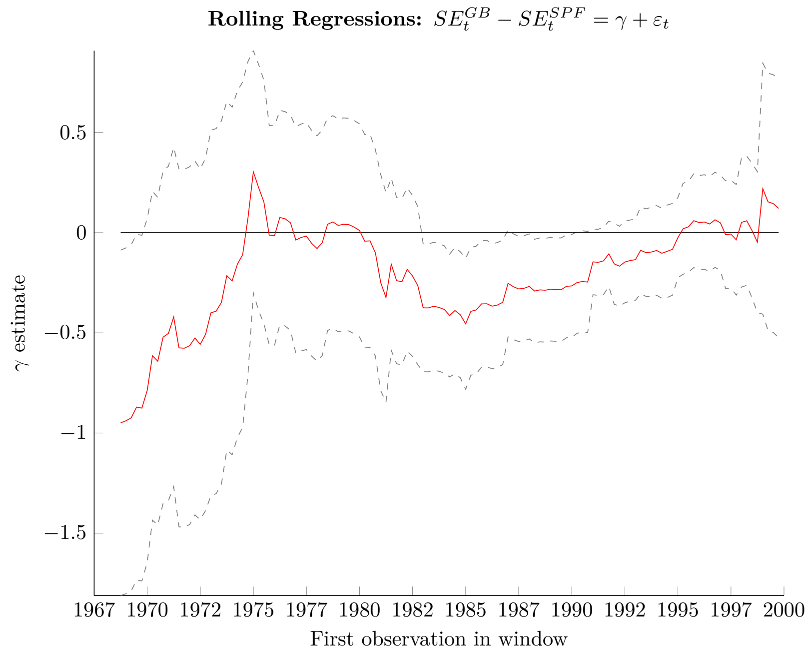


Figure 2.3: Estimated difference in MSE of Greenbook and Median SPF forecast, rolling (ten year window) regressions. The red line indicates the estimated γ while the dotted line indicates two-standard-deviation confidence bands.

respondents have is not redundant of the central bank. However, there are a number of *prima facie* reasons central bank information may matter in ways that are different from firm information and may be important for understanding business cycles. First, if monetary policy has real effects, than the informational process that leads to monetary policymaking matters. If a firm makes a pricing "mistake" due to a bad forecast, it has no aggregate implications. But if the central bank forecasts badly and sets an interest rate "too high" or "too low", its effect on household savings and consumption decisions may have business cycle implications, as emphasized by Orphanides and van Norden (2002). Second, if monetary policy has a signaling role, as in Melosi (2017), then to the extent that agents internalize that the central bank's forecasts are of higher or lower quality may impact how much weight is placed on the interest rate as a signal. Third, as Kohlhas (2015) emphasizes, monetary policy changes may have implications for the informativeness of other signals, to the extent that monetary policy stabilization could "obscure" underlying changes in the economy, but may also be useful to firms who are trying to assess the central banks' private information.

2.3 A model with cross-higher order expectations

In this section, I take a simple dynamic New Keynesian model and augment it to include dispersed information across firms and the central bank. The model is similar to that of Melosi (2017). Firms produce goods using only labor and set prices subject to a Calvo friction. Firms and central bankers have different information sets and receive noisy signals about persistent exogenous aggregate demand and technology shocks. Monetary policy is set according to a Taylor rule, but also subject to persistent shocks, which firms do not observe directly. The model rationalizes differences in forecasts through differences in signal quality; consistent with the evidence cited in the previous two sections, the Federal Reserve has a strict informational advantage because it knows its own actions, but firms may learn from what the Federal Reserve is doing.

2.3.1 Timing

Time in the model is discrete. At the beginning of the period, exogenous shocks are realized, and the central bank observes signals about the state. Given those signals, the central banker infers the state of the world and sets the interest rate on government debt according to a Taylor (1993) rule. In the second stage, firms observe their idiosyncratic technology, a private signal about demand, and the interest rate set by the central banker. They use this information to update their estimate of the state and set prices if able. In the final stage, households perfectly observe all shocks, and decide their consumption, saving, and labor supply decisions. Firms hire labor to produce and deliver the demanded quantity of their good at the price set at the previous stage. The fiscal authority issues bonds and collects taxes (makes transfers) so that its budget constraint is satisfied; and markets clear.

2.3.2 Households

Households have identical (full) information sets and preferences and I use a representative household to stand in for all households.[5] The representative household maximizes utility from a composite consumption good (aggregated using a Dixit-Stiglitz technology) and leisure, subject to a period budget constraint and no-Ponzi condition. Their optimization problem is:

$$\begin{aligned}
&\max_{C_{t+s},B_{t+s},N_{t+s}} E_t \sum_{s=0}^{\infty} \beta^{t+s} g_{t+s} \left[\ln C_{t+s} - \chi_n N_{t+s}\right] \\
&\text{subject to} \\
&P_{t+s}C_{t+s} + B_{t+s} = W_{t+s}N_{t+s} + R_{t+s-1}B_{t+s-1} + \Pi_{t+s} - T_{t+s} \\
&\text{for all } s
\end{aligned}$$

In the household's problem β is the deterministic discount factor, g_t is a stochastic preference shifter, ν is the elasticity of substitution between varieties, χ_n is the parameter governing the marginal disutility of labor, and C_t is the amount of the composite consumption good:

$$C_t = \left(\int_0^1 C_{j,t}^{(\nu-1)/\nu}\right)^{\nu/(\nu-1)}$$

where $C_{j,t}$ is the differentiated variety produced by firm j. B_t are 1-period government bonds that pay a (known in advance) gross nominal rate R_t next period. Workers are hired in spot markets at (common) wage rate W_t. Π_t are lump-sum profits from the disaggregated goods firms and T_t are lump sum taxes or transfers from the government. P_t is the aggregate price index.

Given the Dixit-Stiglitz structure of the composite good, demand for individual differentiated goods and the price level of the composite good take the familiar form:

[5]The decision to model households as having full information may see somewhat counterintuitive, given that other agents have partial information. This decision is made for several reasons. First, it avoids complications about the distribution of wealth. Second, it keeps the structure of the model close to that of Melosi (2017), which facilitates comparison of my results to his. However, it has some complications for the results, which I discuss at the end of section 2.5.

$$
\begin{aligned}
C_{j,t} &= \left(\frac{P_{j,t}}{P_t}\right)^{-\nu} C_t \\
P_t &= \left(\int_0^1 P_{j,t}^{1-\nu}\right)^{\frac{1}{1-\nu}}
\end{aligned}
$$

The household's optimization problem yields the following intratemporal optimality condition

$$\frac{W_t}{\chi_n} = P_t C_t$$

and the intertemporal optimality condition (the household Euler equation)

$$E_t \beta \left[\frac{g_{t+1}}{g_t} \frac{C_t}{C_{t+1}} \frac{R_t}{\pi_{t+1}}\right] = 1$$

where gross inflation is $\pi_t = P_{t+1}/P_t$.

Finally, I denote the nondeterministic part of the household stochastic discount factor as $\Xi_{t+s|t}$ and the SDF as

$$
\begin{aligned}
\Xi_{t|t+s} &= \frac{1/C_{t+s}}{1/C_t} \\
\mathrm{SDF}_{t|t+s} &\equiv \beta^s \Xi_{t+s|t}
\end{aligned}
$$

respectively.

2.3.3 Firms

Firms produce output $Y_{j,t}$ with a linear production technology:

$$Y_{j,t} = A_{j,t} N_{j,t}$$

where $N_{j,t}$ is the labor employed by a particular firm j and where $A_{j,t}$ is the firm's idiosyncratic productivity. I assume that labor markets are perfectly competitive, so firms take the wage rate as given.

Firms' price setting is subject to a Calvo friction. A fraction θ are unable to reoptimize in a given period, and instead index their prices to the steady state inflation rate π_*. Moreover, firms can not observe the true state of the world, and instead must infer it given their information set. I assume they observe an infinite history of their idiosyncratic productivity, their idiosyncratic signal about demand conditions $g_{j,t}$, and the central bank's interest rate. Their information set at time t is:

$$\mathcal{I}_{j,t} \equiv \{\ln A_{j,\tau}, \ln g_{j,\tau}, \mathrm{R}_{\tau} : \tau \leq t\} \tag{2.1}$$

Denote nominal marginal cost

$$MC_{j,t} = W_t / A_{j,t}$$

Assume firms who are allowed to change prices in period t maximize their lifetime discounted perceived profits given their technology and a commitment to meet demand at the price they set. Because households own firms, profits are discounted using the household stochastic discount factor. All output is consumed, so $Y_{j,t} = C_{j,t}$.

Let $P^*_{j,t}$ be the optimal price selected by firm j that gets to reoptimize at t. The firm's optimality condition is:

$$E_{j,t}\left\{\sum_{s=0}^{\infty}(\beta\theta)^s \Xi_{t|t+s}\left[(1-v)\pi_*^s + v\frac{MC_{j,t+s}}{P^*_{j,t}}\right]Y_{j,t+s}\right\} \quad = 0 \tag{2.2}$$

2.3.4 Fiscal policy and the central bank

The fiscal authority is passive and does not consume. It uses lumps-sum taxes and transfers to balance its budget:

$$R_{t-1}B_{t-1} = B_t + T_t$$

The central bank sets interest rates according to its policy rule:

$$R_t = R_* \pi_* E_t\left[\left(\frac{\pi_t}{\pi_*}\right)^{\phi_\pi}\left(\frac{Y_t}{Y_t^f}\right)^{\phi_y}\middle| \mathcal{I}_t^{CB}\right]\eta_{r,t} \tag{2.3}$$

where R_* is the steady state real interest rate, π_t is the gross inflation rate, and Y_t^f is the output that would obtain if there were no information frictions and prices were fully flexible. Policy is conditional on the information set of the central bank at time t, $\mathcal{I}_t^{CB}$. $\eta_{r,t}$, which I refer to as "the policy shock" captures reasons other than information imperfections the central bank might deviate from the Taylor rule prescribed rate.

2.3.5 Exogenous processes

Aggregate technology shocks follow a stationary AR(1) process in logs:

$$\begin{aligned} \ln a_t = &\quad \rho_a \ln a_{t-1} + \sigma_a \varepsilon_{a,t} \\ \text{with} &\quad \varepsilon_{a,t} \sim \text{i.i.d.} N(0,1) \end{aligned} \tag{2.4}$$

The monetary policy rule is subject to shocks $\eta_{r,t}$ which also follow an AR(1) process in logs:

$$\begin{aligned} \ln \eta_{r,t} = &\quad \rho_r \ln \eta_{r,t-1} + \sigma_r \varepsilon_{r,t} \\ \text{with} &\quad \varepsilon_{r,t} \sim \text{i.i.d.} N(0,1) \end{aligned} \tag{2.5}$$

Finally, demand shocks (the preference shifter in household utility) takes a similar form

$$\begin{aligned} \ln g_t = &\quad \rho_g \ln g_{t-1} + \sigma_g \varepsilon_{g,t} \\ \text{with} &\quad \varepsilon_{g,t} \sim \text{i.i.d.} N(0,1) \end{aligned} \tag{2.6}$$

Firm and central bank signals about technology and demand conditions take the form of the truth plus normally distributed *i.i.d* noise, where firms share a common noise variance, which may differ that of the central bank.

2.4 Solving the model and information assumptions

To make the analysis more tractable, I make a number of assumptions about the information available to firms and the central bank. In particular, the model is analyzed linearly approximating the first order conditions. I assume that agents know the linear approximation to the model (its structure and parameters) and use it to conduct inference. The linearity of their inference problem, combined with assuming that all of the shocks are Gaussian, means that firms will optimally use the Kalman filter for analysis.

I also solve the model under the assumption of "common knowledge of rationality"(Nimark (2007)). Informally, it is common knowledge that agents form model-consistent expectations. This will result in the formation of higher order expectations, but given stationarity of the fundamental disturbances and discounting of the future, additional higher orders of expectation eventually converge and we can find the approximate solution to the infinite-order model by truncating at a finite order $\bar{k}$ (Kohlhas, 2014, Appendix B).

2.4.1 The linearized model

After linearizing around the nonstochastic steady state, the model resembles the canonical three-equation dynamic New Keynesian model:

$$\pi_t = (1-\theta)\bar{E}_t\pi_t + (1-\theta)(1-\beta\theta)\bar{E}_t y_t - (1-\theta)(1-\beta)a_t + \beta\bar{E}_t\pi_{t+1} \tag{2.7}$$
$$y_t = g_t - E_t g_{t+1} + E_t y_{t+1} + E_t\pi_{t+1} - R_t \tag{2.8}$$
$$R_t = \phi_\pi E_t^{CB}\pi_t + \phi_y E_t^{CB}(y_t - a_t) + \eta_t \tag{2.9}$$

where $\bar{E}_t$ indicates the *average* firm expectation conditional on time t information and E_t^{CB} indicates the expectation of the central bank conditional on their information up to time t. a_t appears in (2.7) because real marginal costs are approximately equal to $y_t - a_t^j$ and $\int_j a_t^j = a_t$. a_t appears in (2.9) because up to a first order approximation, deviations from steady state in the natural rate of output - the level of output that would obtain in the absence of information or nominal frictions - are equal to deviations in technology.[6] Higher order expectations - i.e., expectations of expectations - arise naturally in this setting. Recursively substituting (2.7) reveals current inflation depends on firms' higher order average beliefs about future inflation, output, and technology. Higher order expectations across the firm and central bank also arise. To see why this is the case, assume

[6] The derivation of these conditions is found in Appendix 2.A and is essentially the same as Nimark (2008) and Melosi (2017).

(for now) that $\phi_y = 0$. The central bank, in order to set its interest rate, must form expectations about current period inflation. Since we assumed that expectations were model-consistent, we have that

$$R_t = \quad \phi_\pi E_t^{CB}((1-\theta)\bar{E}_t\pi_t + (1-\theta)(1-\beta\theta)\bar{E}_t y_t - (1-\theta)(1-\beta)a_t + \beta\bar{E}_t\pi_{t+1}) + \eta_t$$

i.e., current period interest rates depend on the central bank's expectation of firms' average expectation of inflation, for example. But then note that by (2.8), current realized output depends on R_t and by (2.7), inflation depends on average expectations of output. So we have that

$$\pi_t = f(\bar{E}_t y_t) = f(\bar{E}_t(E_t^{CB}\bar{E}_t\pi_t))$$

i.e., inflation is a function of average beliefs about output, which is a function of the central bank's expectation of average beliefs about inflation, so firms' beliefs about inflation depend on their beliefs about the central bank's beliefs about their beliefs, and so on.

Even if the central bank had full information, the model would imply an infinite regress of higher-order expectations about other firms' beliefs. Now there are two infinite regress problems - one among firms and one between firms and the central bank.[7]

I solve the model numerically by matching coefficients. I suppose (and later confirm) that the the solution can be written in terms of a (truncated) state vector $X_t^{(0:\bar{k})}$ which contains all of the higher-order expectations in the model. I assume the model solution takes the form

$$\begin{aligned} \begin{bmatrix} \pi_t \\ y_t \end{bmatrix} &= AX_t^{(0:\bar{k})} + BR_t \\ R_t &= TX_t^{(0:\bar{k})} \end{aligned} \tag{2.10}$$

And the state follows the law of motion:

[7]Although the notion of forming cross-higher order expectations may seem a little unusual, it is an immediate implication of model-consistent expectations in this context. Moreoever, it appears to have real-world relevance: Barrons.com's Income Investing blog in September 2015 quoted a market analyst as saying: "We believe the U.S. economy is on solid footing. we believe the Fed believes the U.S. economiy is on solid footing [...] it's more difficult to have confidence in that forecast. The Fed's decision on whether to raise rates in September is likely to hing on how confident the FOMC feels about their forecasts." (Stone (2015)).

$$X_{t+1}^{(0:\bar{k})} = MX_t^{(0:\bar{k})} + N \begin{bmatrix} \varepsilon_t^a \\ \varepsilon_t^r \\ \varepsilon_t^g \\ e_t^{CB,a} \\ e_t^{CB,g} \end{bmatrix} \tag{2.11}$$

To arrive at this representation of the economy, we first need to more concretely characterize how agents perform inference.

2.4.2 Higher order expectations

As the discussion in the previous subsection makes clear, agents in the model must form higher-order expectations in order to infer their optimal action. In this section, I characterize those beliefs; the discussion here follows Kohlhas (2015, 2014) closely.

Write the actual realization of a particular variable z as $z^{(0)}$. Collecting the three exogenous states in a vector x_t, we have

$$x_t \equiv x_t^{(0)} \equiv \begin{bmatrix} a_t \\ \eta_t \\ g_t \end{bmatrix}$$

Denote the higher order expectations of *firms* recursively:

$$x_t^{(n)} \equiv \int_0^1 E_t[x_t^{(n-1)}|\Omega_t^j]dj$$

i.e., the average expectation of the realization of the exogenous variables is $x_t^{(}1)$, the average firm expectation of its average expectation is $x_t^{(2)}$ and so on.

Denote the expectations of the central bank with a (CB) superscript. That is, the average expectation of the central bank about the realization of variables is $x_t^{(CB)}$, its expectation of the average expectation is $x_t^{(CB)(1)}$ and so on. Note that because information sets are not nested, the law of iterated expectations does not apply. Thus, we need to keep track of, for example, the average firm expectation of the central bank expectation of the firm's average expectation:

$$
\begin{aligned}
x_t^{(1)(CB)(1)} &= \int_j E_t^j x_t^{(CB)(1)} dj \\
x_t^{(CB)(1)(CB)(1)} &= E_t^{CB} x_t^{(1)(CB)(1)}
\end{aligned}
$$

I refer to an "order" of expectation as a collection the expectations of both firms and the central bank where the expectations operator has been applied the same total numbre of times; i.e., the second order of expectations contains the firms' average expectations of average expectations, its average expectation of the central bank's expectation, and the central bank's expectation of the firm average expectations, as shown in equation (2.12).

$$
X_t^{(0:\infty)} = \left[\begin{array}{c}
a_t \\
\eta_t \\
g_t \\
\hline
x_t^{(1)} \\
\hdashline
x_t^{(CB)} \\
\hline
x_t^{(2)} \\
x_t^{(1)(CB)} \\
\hdashline
x_t^{(CB)(1)} \\
\hline
x_t^{(3)} \\
x_t^{(2)(CB)} \\
x_t^{(1)(CB)(1)} \\
\hdashline
x_t^{(CB)(2)} \\
x_t^{(CB)(1)(CB)} \\
\hline
x_t^{(4)} \\
x_t^{(3)(CB)} \\
x_t^{(2)(CB)(1)} \\
x_t^{(1)(CB)(2)} \\
x_t^{(1)(CB)(1)(CB)} \\
\hdashline
x_t^{(CB)(3)} \\
x_t^{(CB)(2)(CB)} \\
x_t^{(CB)(1)(CB)(1)} \\
\hline
\vdots \\
\vdots
\end{array}\right] \tag{2.12}
$$

To isolate terms that are type-specific, write:

$$
X_t^{(0:\infty)} = H \begin{bmatrix} x_t \\ \bar{E}_t X_t^{(0:\infty)} \\ E_t^{CB} A_n X_t^{(0:\infty)} \end{bmatrix} \tag{2.13}
$$

where A_n is a matrix that deletes the redundant orders of expectations. H is a reordering matrix that reorganizes the vector $\begin{bmatrix} x_t, & \bar{E}_t X_t^{(0:\infty)}, & E_t^{CB} A_n X_t^{(0:\infty)} \end{bmatrix}^T$ into the form of (2.12). Some details about how higher orders of expectation are organized in practice are found in appendix 2.B.1. Details of the inference problems are found in appendix 2.B. The end result is a recursive expression for the matrices governing hte law of motion in equation (2.11).

2.4.3 Structural equations, equilibrium, and solution method

In this section, I outline the structural equations of the model conditional on the proposed law of motion for the state.

Define $\bar{H}$ and H^{CB} as two matrices that select $E_t X_t^{(0:\bar{k})}$ and $E_t^{CB} A_n X_t^{(0:\bar{k})}$ from $X_t^{(0:\bar{k})}$ (See appendix 2.B.4 for details). We can rewrite the linearized model in terms of the proposed solution, known coefficient matrices, and $X_t^{(0:\bar{k})}$. To begin, write the system of structural equations:

$$
\begin{aligned}
\begin{bmatrix} \pi_t \\ y_t \end{bmatrix} &= \begin{bmatrix} (1-\theta) & (1-\theta)(1-\beta\theta) \\ 0 & 0 \end{bmatrix} \bar{E}_t \begin{bmatrix} \pi_t \\ y_t \end{bmatrix} + \begin{bmatrix} \beta\theta & 0 \\ 0 & 0 \end{bmatrix} \bar{E}_t \begin{bmatrix} \pi_{t+1} \\ y_{t+1} \end{bmatrix} \\
&\quad + \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} E_t \begin{bmatrix} \pi_{t+1} \\ y_{t+1} \end{bmatrix} + \begin{bmatrix} 0 \\ -1 \end{bmatrix} R_t + \begin{bmatrix} -e_1'(1-\theta)(1-\beta\theta) \\ (1-\rho_g)e_3' \end{bmatrix} X_t^{0:\bar{k}} \\
R_t &= \begin{bmatrix} \phi_\pi & \phi_y \end{bmatrix} E_t^{CB} \begin{bmatrix} \pi_t \\ y_t \end{bmatrix} + \begin{bmatrix} -e_1\phi_y \end{bmatrix} E_t^{CB} X_t^{0:\bar{k}} + e_2 X_t^{0:\bar{k}}
\end{aligned}
\tag{2.14}
$$

In appendix 2.C, I show how to rewrite the expectations terms solely in terms of structural matrices and the state. The end result, after matching coefficients and conditional on the law of motion for the state, is a system of three matrix equations in three unknown matrices A, B and T:

$$
\begin{aligned}
A &= G_C A\bar{H} + G_F(A+BT)M\bar{H} + RX_F(A+BT)M + G_X \\
B &= G_C B + G_R \\
T &= F_C(AH^{CB} + BT) + F_X H^{CB} + e_2
\end{aligned}
\tag{2.15}
$$

B can be solved for explicitly, and in the appendix I derive an explicit expression for T in terms of A. With all the elements in place, I define equilibrium:

Definition 1 (Equilibrium)**.** *An equilibrium is a collection of matrices A, M and N such that the system of matrix equations* (2.15) *and appendix equations* (2.45)*,* (2.46) *are satisfied.*

The definition of equilibrium is that agents' beliefs about the evolution of the state and the mapping between the state and endogenous variables is consistent with the actual mapping between the state and the exogenous variables, and also consistent with individual optimality conditions. All of these restrictions are summarized the cross-equation restrictions that define A, M and N.

Unfortunately, even with truncation, we can not solve the model explicitly in terms of the parameters; in general there is not a closed-form expression for the Kalman gain matrices embedded in M and N. However, by the arguments in Nimark (2007) and Kohlhas (2014), the mapping from $\{A, M, N\} \rightarrow \{A, M, N\}'$ is a contraction mapping so an equilibrium exists.

In practice, I find a solution to the model by initializing at the full-information solution, using that to back out an implied M and N, then iterating between updating A and M, N until convergence.

2.5 Quantitative Results

In this section, I illustrate through calibrated examples the ability of the model to generate macroeconomic and forecast dynamics that can rationalize some of the facts established in section 2.2. As a preview, the model requires some degree of central bank noise to match comovement of inflation and firm and central bank inflation forecasts. But it is essentially unable to capture the correlation of measured total factor productivity with other variables, or capturing correlations of output with inflation and inflation forecasts. and has a hard time matching the relationship of output to other variables as well. The experiments reveal that the type of and degree of central bank noise matters for the model's ability to match forecast moments.

The model calibration is shown in table 2.3. The degree of price stickiness, the variance of the shocks, and the persistence of demand and monetary shocks are close to the mean estimates in Melosi (2017). Technology shocks are calibrated to match the empirical persistence of utilization-adjusted TFP for non-investment goods (from Fernald (2012)), which is much lower than the estimated value in Melosi (2017). The weight on the output gap in the Taylor rule is set to a lower value in line

Parameter	Value
β	.99
θ	0.35
ϕ_π	1.5
ϕ_y	0.05
ρ_a	0.8796
$100\sigma_a$	1.5
ρ_g	0.9038
$100\sigma_g$	3.75
ρ_r	0.9468
$100\sigma_r$	0.85
$100\widetilde{\sigma_a}$	2.6
$100\widetilde{\sigma_g}$	8
$100\widetilde{\sigma_a}^{CB}$	Varies (see table 2.4)
$100\widetilde{\sigma_g}^{CB}$	Varies (see table 2.4)
$\bar{k}$	10

Table 2.3: Calibrated parameters of the cross-higher-order-expectations model.

with most of the literature. $\widetilde{\sigma_g}$ is set to a much lower value than in Melosi (2017) because highly uniformative firm signals about demand shocks lead to nonexistence of equilibria when the central bank is also uninformed about technology shocks.[8]

Given this calibration, I examine the role of central bank information by varying over different calibrations for the central bank's two noise parameters. Details of the calibration are in table 2.4. The first calibration essentially corresponds to perfect central bank information; the second, to a situation where the central bank has strictly more accurate exogenous signal than firms; the third explores what happens when the central bank and firms have equally noisy exogenous signals, while calibrations 4 and 5 explore asymmetric quality of information across technology and demand shocks. I evaluate the ability of the model to match the data by examining the response of the economy conditional on exogenous shocks and the ability of the model to match unconditional second moments. For the latter, I use the correlogram (out to 6 lags) of detrended data on GDP deflator inflation, real GDP (from the BEA's NIPA tables) the Federal Funds Rate (the quarterly average of the effective daily rate reported by the Federal Reserve Bank of St. Louis), utilization-adjusted TFP for non-investment goods, and 1-4 quarter ahead forecasts of inflation from the SPF

[8] I comment on some of the technical issues at the end of this section.

Calibration	$100\widetilde{\sigma_a}^{CB}$	$100\widetilde{\sigma_g}^{CB}$
1	2×10^{-14}	2×10^{-14}
2	2	2
3	2.6	8
4	1.3	12
5	3.9	4

Table 2.4: Calibrated central bank noise parameters.

(using the median forecast) and the Greenbook. The data run from 1960Q1-2010Q4, with moments calculated pairwise to account for jagged start dates in the forecast data. Data are detrended using the Hamilton (2017) filter.

Plots of impulse responses (to a 100× standard deviation shock) and correllograms are shown in appendix 2.D.[9].

Full central bank information. Calibration 1 essentially implies central bankers have full information while firm information is dispersed. Impulse responses and correlations are displayed in 2.D.1. The high quality of central bank information is apparent from the impulse responses, where central bank expectations are arbitrarily close to the realization of variables. As the evidence in section 2.2 shows, this is counterfactual. Under this calibration, firms underestimate technology shocks slightly, and also misattribute part of the effect of a technology shock to negative demand or interest rate shocks. There is a slight price puzzle in response to interest rate shocks, although inflation expectations rise in response to this shock, which is consistent with evidence from a sign-restricted VAR estimated by Melosi (2017). Such shocks are contractionary, and firms attribute part of what they observe to negative technology shocks and a small, positive demand shock. Demand shocks, however, do not have their “typical” effect in New Keynesian models. Inflation essentially does not change, while output expands. This is likely because firms’ information about demand is quite poor and prices are relatively flexible. This means that losses of mispricing (due to not raising prices

[9] I also solve a version of the model using Melosi (2017)’s mean posterior results for all the parameters. At his calibration, more than an arbitrarily small amount of noise in the observation of the level of technology results in the absence of a solution, although there can be somewhat more noise in demand. Results for that calibration and some additional details are in appendix 2.E.

when household demands are higher) are relatively low (since they will likely be adjusted in the near future), and firms underestimate those losses further because they underestimate the positive demand shock. Hence, firms do not believe that others will raise prices much (inflation expectations do not increase), nor do they believe they themselves should raise prices for fundamental reasons. In general, the model has difficulty generating expansions in inflation as a result of demand shocks, regardless of the calibration; it is most stark here, when there are not complicated confounding effects from central bank information.

The "supply shock" behavior of inflation, interest rates, and output is the emphasis of Melosi (2017)'s estimated results because it allows the model to match the data in the 1970s. However, in his estimated results, the impact response of demand shocks is of a similar magnitude and inflation expectations fall essentially immediately. However, as the impulse responses in appendix 2.E shows, this results is fragile to the extent that introducing noise in the central bank's observation of demand shocks restores their typical behavior when other parameters are left at the posterior mean of hist estimation.

Turning to unconditional second moments, the model with near-perfect central bank information captures qualitatively the autocorrelation of inflation and (to an extent) its co-movement with interest rates. It generates essentially no feedback from activity in the current or previous quarters to inflation, whereas in the data there is a weak positive relationship between lagged output and inflation. As mentioned at the start of the section, the model (regardless of calibration) misses a great deal regarding the relationship between detrended technology and other variables. Finally, the model has difficulty capturing the relationship between firm expectations and interest rates, and features of far-ahead central bank inflation expectations. In short, the model captures some features of (firm) beliefs but makes counterfactual predictions about central bank beliefs; to capture the fact that central bankers make forecast errors, we have to introduce noise into their beliefs.

Noisy but superior central bank information. Calibration two assumes central bank observations are noisy, but strictly better than those of firms. The slight addition of noise, however, changes the conditional and unconditional predictions of the model as shown in 2.D.2. Firms' forecasts are immediately worse conditional on each kind of shock, in the sense that their absolute forecast errors are larger.[10] This represents the loss of reliable information coming from the interest rate signal. The central bank also makes mistakes; in fact their mistakes are somewhat larger than those of firms after technology shocks, because they under-estimate the size of the shock to a greater extent.[11] Technology shocks cause a greater fall in inflation than in the no-noise case, despite *current* inflation expectations reacting about the same amount; inspecting the New Keynesian Phillips Curve (2.7), this implies that agents believe that future inflation (or higher-order average firm beliefs about inflation) will be lower for longer than in the full-central-bank-information case. Essentially, firms correctly assume that the central bank will underreact to the shock and, indeed, the gain in output is smaller than in the full-information case because nominal do not fall as much. In response to a monetary policy shock, the overall path of inflation, output, and rates is reasonably similar to the previous case. Owing to its (relative more) superior information, the central banks' beliefs are more accurate than firms' in response to a demand shock. The last two sets of impulse responses show that, owing to the slow learning process by central banks and firms, *i.i.d* noise in central bank beliefs has persistent effects on inflation and especially output; Technology noise shocks are qualitatively similar to an expansionary monetary policy shock in a full-information New Keyesnian model. By contrast, demand noise shocks are quite small and have barely any effect. The difference is that a one-standard deviation demand noise shock produces relatively little movement in interest rates and is quickly reversed, so firms' beliefs do not change in a particularly persistent way; households, of course, do not react to the aggregate noise shock except to the extent interest rates or prices change.

[10] The "jagged" movements in impulse responses are related to the accuracy of the solution, as I discuss at the end of the section.

[11] Essentially, firms are able to construct a more accurate signal of technology by combining their idiosyncratic information with the interest rate signal, at the expense of mis-attributing some of the interest rate movement endogenously to other shocks. The central bank must rely on a single signal, but only under-estimates fundamental changes because it attributes some of what it observes to its own noise.

In terms of unconditional moments, the addition of noise qualitatively improves the cross-correlation structure between central bank and firm beliefs, and (to an extent) the relationship between interest rates and lags of central bank forecasts. The relationship between inflation and contemporaneous and lagged output is somewhat improved as well, although the model continues to predict too great of a correlation between contemporaneous inflation and lagged inflation expectations relative to the data.

Overall, introducing noise into the signals of the central bank qualitatively matches a few features of section 2.2. The central bank makes forecasting errors, just as firms do; the extent of those errors is conditional on the type of shock (hence, the advantage might show up more in time periods where fluctuations are demand- or policy-driven). Consistent with the empirical evidence cited in the introduction, monetary policy actions are still informative to firms.

Symmetry in noise Calibration 3 illustrates an economy where the central bank and firms have equally noisy signals about technology and demand. Recall that this does not mean they will have equal forecast errors; the Fed has a strict advantage over firms in that it observes η_t perfectly, but firms have the advantage of combining their signals with the interest rate when forming beliefs. It turns out that calibrations 3 and 4 produce reasonably similar impulse responses (except following technology shocks) and unconditional correlations, so appendix 2.D.3 contains the results for only calibration 3 with only one additional impulse response for calibration 4. When the central bank has relatively better information about technology, technology shocks become contractionary. As an accounting matter, aggregate hours must be strongly falling in response to the technology shock in order for technology improvements to generate a decrease in output. This can come from price rigidity induced by both the Calvo friction and the underestimate of the shock - firms need fewer workers to produce input matching the demand they face from households at the prevailing price. Indeed inflation is (conditionally) slower to react here than in the first calibration. At the same time, real rates are rising (since inflation is falling more than nominal rates) so that consumers are satisfied with a reduction in consumption.

Monetary shocks are qualitatively similar to calibration 2, although firms do not attribute as much of the change in interest rates to demand shocks. Demand shocks do not generate much inflation because firms systematically underestimate the size of the shock and do not get much help from the central bank's policy signal. Output increases by more, however, because households are able to borrow cheaply to finance the additional consumption they desire, and prices are not adjusting to counteract this effect. Unconditionally, the model somewhat more closely captures the relationship of federal reserve inflation forecasts to lags of other observables, although it does worse at matching the unconditional relationship between TFP and output.

Adjusting the central bank's information so that they have an advantage in technology shocks but a relative disadvantage in forecasting demand shocks is qualitatively quite similar to the symmetric noise case. The largest difference is in impulse responses to technology shocks. Inflation falls less and interest rates fall slightly more; this is enough that real rates do not rise as much and technology shocks are again expansionary. The reaction of endogenous variables and beliefs to monetary shocks is more or less similar, as is the qualitative/quantitative effect of demand shocks. Accordingly, the main difference between the 3rd and 4th calibrations in terms of unconditional second moments is that the higher-noise-on-technology-shocks quantitatively matches the low, positive, TFP and output correlation structure while calibration 3 basically implies they are uncorrelated. Firms' expectations are still estimated to have a too-high correlation with interest rates (driven by too high of a correlation with near-term central bank inflation expectations), but both calibrations 3 and 4 are able to capture correlations of inflation forecasts reasonably well.

Asymmetric noise. Finally, appendix 2.D.4 illustrates an economy where the central bank has relatively noisier technology signals but relatively superior demand signals. Unsurprisingly, under this calibration firms' forecasts for output (and the output gap) are more accurate than the central banks' following technology shocks (their inflation forecasts appear to be roughly equally accurate), but the central bank has more accurate forecasts following demand shocks. Shocks also have effects at odds with the typical results in New Keynesian models; technology shocks produce business cycle comovements similar to negative demand shocks (excepting the change in TFP itself) owing to the rise in real rates, and demand shocks produce mild disinflation.[12] models are important for business cycle accounting; Unconditionally, the model generates negative correlation between output and TFP, but does similarly in terms of matching the correlogram of inflation expectations.

Summary of the quantitative exercise. Regardless of the calibration, the model has some difficulty matching features of the data; particularly, the relatively mild relationship between TFP, inflation, and output, and comovements between output and other variables. However, adding noise to the central bank's observations of exogenous demand and technology processes allows the data to qualitatively match the fact that interest rate movements are informative for firms, and neither firms nor the central bank have an absolute advantage in forecasting.

Since the dynamics of the model depend nonlinearly on all the parameters of the model, any lessons about what features models with dispersed information among firms and the central bank should have are necessarily tentative. Adjustments to the processes of exogenous processes and signal noise appear to have complicated effects on both conditional and unconditional predictions of the data. Nevertheless, a few lessons emerge. First, when the central bank has overly-noisy observations of technology, its failure to accommodate technology shocks by lowering the nominal rate can lead to contractionary expansions in technology and a negative correlation between technology and output. Since the correlation in the data used here is weakly positive, this suggests that (at least for the United States), the central bank's observation error for technology is not too large (at least here, not larger than firms'). Too noisy technology observations by the central bank can also lead to the nonexistence of fixed points, which may be a purely numerical issue (discussed below), but also may

[12]The unusual impulse responses illustrate that both the form of the monetary policy rule and assumptions about what information agents have are important for accounting for business cycle fluctuations.

be related to satisfying the Taylor principle.[13] However, it may be that firms have weakly better signals about aggregate demand shocks, even if those signals are quite noisy; this is consistent with both firms and the central bank having different expertise in forecasting that leads to differences in relative accuracy at different times. Second, when the central bank faces a signal extraction problem, its measurement error shocks can resemble monetary policy shocks; this depends both on the size of its noise (which determines how quickly the central bank learns about its error) but also how firms and the private sector endogenously incorporate that error into their own beliefs. Third, relative to a model which features exogenous measures of forecast error (such as Melosi (2017)), the model has a more difficult time matching business cycle movements. This partially seems related to the assumption that households have perfect information, unlike firms and the central bank. Essentially, the model lacks a mechanism to limit household reactions to shocks; firms do not raise prices very quickly and the central bank does not adjust rates as much as it would otherwise. This is a distinction between the results here and those of Melosi (2017). In Melosi's model, mismeasurement of inflation is just another shock, and one could always engineer observationally equivalent sequences of interest rate and inflation forecast error shocks. But when central bank expectations are endogenous, the central bank *systematically* under-reacts to exogenous shocks. Hence, the assumption of a representative household with full information, made in Melosi (2017) and made here, is not innocuous.

Computational issues. Several of the impulse responses, especially those related to monetary policy shocks, feature unusual saw-tooth dynamics. This appears to be related to the accuracy of the truncated solution; it shows up more frequently when low orders of truncation are used, when shocks are more persistent, and when signal noise is larger. This is partially why I have emphasized qualitative features of the impulse responses rather than the magnitudes. Kohlhas (2015) finds that he is able to find accurate solutions for $\bar{k} = 8$ but I find even for a higher level of truncation, accuracy

[13] A similar issue is discussed in Boehm and House (2014).

becomes an issue. However, the model with explicit cross-higher-order-beliefs encounters the curse of dimensionality rather quickly as the order of truncation is increased.[14]

2.6 Conclusion

Inspired by evidence that monetary policy has an information and coordination role, and that central bankers' forecasts are not systematically better than the forecasts of firms, I examine the ability of a monetary policy featuring both dispersed firm information and endogenous central bank forecasts to match both macroeconomic data and agents' beliefs. The model can successfully match features of both firm and inflation forecasts, and under certain calibrations captures the feedback between central bank beliefs, interest rates, and firm beliefs. However, the model has difficulty matching output dynamics despite these successes. This is likely due to a simplifying assumption that households have full information, which is less innocuous here than in the estimated model of Melosi (2017). The feedback effects between central bank information and household consumption are difficult to study given the complicated nature of the model, but they appear to be critical in getting the model to match data on beliefs and outcomes simultaneously.

Empirical evidence - including Romer and Romer (2000), Tang (2015), and Nakamura and Steinsson (2013) - suggests that in the United States, the private sector learns about the macroeconomy in part by observing what the Federal Reserve does. The reduced-form evidence in 2.2 suggests that the forecasts which inform monetary policy rates are imperfect; the results in the previous section suggest treating central bank forecast error as exogenous delivers quite different positive implications than when it is endogenous. My results suggest that understanding this in a unified way is important for our understanding of an informational monetary transmission mechanism, but also for how

[14]Calculating Kalman gains for both firms and the central bank becomes computationally burdensome and raises problems with double-precision arithmetic when the state covariance matrix becomes large; see Simon (2006), for a discussion of some of these issues from the engineering literature. In contrast to Kohlhas (2015), the three persistent shocks in the model apparently require a higher truncation to achieve an accurate solution; $\bar{k} = 10$ implies that X_t has 1,125 elements, which is sufficiently computationally costly to preclude even generalized method of moments estimation, much less full-information methods like those used in Melosi (2017).

monetary policymakers can adjust their policy strategies (either their policy rules or communication strategies) to optimally stabilize the economy. Unfortunately, besides Kohlhas (2015), many of the modeling techniques developed for dealing with dispersed information are not well-equipped for handling "informationally large" agents like a central banker, or dispersed information that differs in quality across types of agents. Frequency-domain techniques, such as Huo and Takayama (2014), Tan and Walker (2015), and Miao et al. (2017) show promise in delivering closed-form solutions of dispersed information models, but are less well-equipped for dealing with endogenous information (or differences in information across agent types) except in simple cases. An alternative for exploring these issues is non-parametric approaches which sidestep the forecasting-the-forecasts of others problem and the accompanying curse of dimensionality, such as Chahrour and Ulbricht (2017). While this type of approach can generate information-based business cycles and match aggregate data well, it is less obvious how they can be brought to bear on understanding how agents learn from particular information sources. An important step for future work is developing techniques that can accurately and quickly solve models with endogenous information and which can accommodate "noise shocks" which do not disappear with aggregation. More accurate solutions will clarify the complicated informational feedback mechanisms; more computationally tractable techniques will also enable full estimation rather than limited comparisons across calibrations.

Appendix

2.A Derivation of the linearized Phillips curve

The derivation of the NKPC is as in Melosi (2017) and Nimark (2008), but is reproduced below.

The starting point is the first order condition for the firm's price setting problem:

$$E_{j,t}\left\{\sum_{s=0}^{\infty}(\beta\theta)^s\Xi_{t|t+s}\left[(1-v)\pi_*^s+v\frac{MC_{j,t+s}}{P_{j,t}^*}\right]Y_{j,t+s}\right\}=0 \tag{2.16}$$

We want to transform this so it is in terms of stationary variables. This involves (mostly) multiplying and dividing by γ^t and aggregate price levels:

$$E_{j,t}\Bigg\{\Xi_{t|t}\left[(1-v)+v\frac{MC_{j,t}}{P_{j,t}^*}\frac{P_t}{P_t}\right]\gamma^t\frac{Y_{j,t}}{\gamma^t} \\ +\sum_{s=1}^{\infty}(\beta\theta)^s\left[(1-v)\pi_*^s+v\frac{MC_{j,t+s}}{P_{j,t}^*}\frac{P_{t+s}}{P_t}\frac{P_t}{P_{t+s}}\right]Y_{j,t+s}\gamma^{t+s}\Xi_{t|t+s}\frac{1}{\gamma^{t+s}}\Bigg\}=0$$

We can write P_{t+s}/P_t as $\Pi_{\tau=1}^s\pi_{t+\tau}$. Using the definitions of the stationary variables:

$$E_{j,t}\left\{\left[(1-v)+v\frac{mc_{j,t}}{p_{j,t}^*}\right]y_{j,t}\zeta_{t|t}+\sum_{s=1}^{\infty}(\beta\theta)^s\zeta_{t|t+s}\left[(1-v)\pi_*^s+v\frac{mc_{j,t+s}}{p_{j,t}^*}\Pi_{\tau=1}^s\pi_{t+\tau}\right]y_{j,t+s}\right\}=0$$

At the nonstochastic steady state, given CES preferences, optimal prices are a constant markup over marginal cost, with the markup term being $v/(v-1)$. This immediately implies that in the above expression the bracketed terms will be zero at the nonstochastic steady state. Furthermore, the necessary condition holds if and only if the bracketed terms are zero, so for the linearizion about the nonstochastic steady state we can focus on the bracketed terms.

Denote nonstochastic steady state variable with a $*$ subscript, e.g. $p_{j,*}$ is the steady state price relative to the price level of firm j. Rewrite the inner terms:

$$E_{j,t}\left\{\left[(1-v)+v\frac{mc_{j,t}}{p_{j,t}^*}\right]+\sum_{s=1}^{\infty}(\beta\theta)^s\left[(1-v)\pi_*^s+v\frac{mc_{j,t+s}}{p_{j,t}^*}\Pi_{\tau=1}^s\pi_{t+\tau}\right]\right\}=0 \tag{2.17}$$

Since we are taking a first order approximation in logs, write:

$$E_{j,t}\left\{\left[(1-v)+v\frac{mc_{j,*}}{p^*_{j,*}}(1+\widehat{mc}_{j,t}-\widehat{p}^*_{j,t})\right]+\sum_{s=1}^{\infty}(\beta\theta)^s\left[(1-v)\pi_*^s+v\frac{mc_{j,*}}{p^*_{j,*}}(1+\widehat{mc}_{j,t+s}-\widehat{p}^*_{j,t}+\sum_{\tau=1}^{s}\widehat{\pi}_{t+\tau})\right]\right\}=0$$

In steady state:

$$1-v+v\frac{mc_*}{p^*_{j,*}}+\sum_s(\beta\theta)^s+(1-v)\pi_*+v\frac{mc_*}{p^*_{j,*}}=0$$

so collecting those terms and eliminating, we focus on the log deviations:

$$E_{j,t}\left[\widehat{mc}_{j,t}-\widehat{p}^*_{j,t}+\sum_{s=1}^{\infty}(\beta\theta)^s\left(\widehat{mc}_{j,t+s}-\widehat{p}^*_{j,t}+\sum_{\tau=1}^{s}\widehat{\pi}_{t+\tau}\right)\right]=0$$

Collecting all the $\widehat{p}^*_{j,t}$ terms and applying the formula for an infinite geometric sum we write

$$E_{j,t}\left[\widehat{mc}_{j,t}-\frac{1}{1-\beta\theta}\widehat{p}^*_{j,t}+\sum_{s=1}^{\infty}(\beta\theta)^s\left(\widehat{mc}_{j,t+s}+\sum_{\tau=1}^{s}\widehat{\pi}_{t+\tau}\right)\right]=0$$

We defined $\widehat{p}^*_{j,t}$ as $\ln P^*_{j,t}-\ln P_t$ Note that aggregate prices aren't in the agent's information set so we can't pull them out of the sum. But own optimal prices are, hence we can substitute in the identity and rearrange to obtain

$$\ln P^*_{j,t}=(1-\beta\theta)E_{j,t}\left[\widehat{mc}_{j,t}+\frac{1}{1-\beta\theta}\ln P_t+\sum_{s=1}^{\infty}(\beta\theta)^s\left(\widehat{mc}_{j,t+s}+\sum_{\tau=1}^{s}\widehat{\pi}_{t+\tau}\right)\right] \tag{2.18}$$

Roll this equation forward one period and apply the law of iterated expectations to obtain

$$\ln P^*_{j,t+1}=(1-\beta\theta)E_{j,t+1}\left[\widehat{mc}_{j,t+1}+\frac{1}{1-\beta\theta}\ln P_{t+1}+\sum_{s=1}^{\infty}(\beta\theta)^s\left(\widehat{mc}_{j,t+s+1}+\sum_{\tau=1}^{s}\widehat{\pi}_{t+\tau+1}\right)\right]$$

Take $\widehat{mc}_{j,t+1}$ inside the sum operator and solve for the sum of marginal costs to obtain:

$$\sum_{s=1}^{\infty}(\beta\theta)^sE_{j,t}\widehat{mc}_{j,t+s}=\frac{\beta\theta}{1-\beta\theta}\left[E_{j,t}\ln P^*_{j,t+1}-E_{j,t}\ln P_{t+1})\right]-\beta\theta\sum_{s=1}^{\infty}(\beta\theta)^s\sum_{\tau=1}^{s}E_{j,t}\widehat{\pi}_{t+\tau+1} \tag{2.19}$$

Substitute (2.19) into (2.18) to obtain

$$
\begin{aligned}
\ln P^*_{j,t} =& (1-\beta\theta)\left[E_{j,t}\widehat{mc}_{j,t} + \frac{1}{1-\beta\theta}E_{j,t}\ln P_t\right] \\
&+ \beta\theta\left[E_{j,t}\ln P^*_{j,t+1} - E_{j,t}P_{t+1}\right] - (1-\beta\theta)\sum_{s=1}^{\infty}(\beta\theta)^s\sum_{\tau=1}^{s}E_{j,t}\widehat{\pi}_{t+\tau+1} \\
&+ \sum_{s=1}^{\infty}(\beta\theta)^s\sum_{\tau=1}^{s}E_{j,t}(\widehat{\pi}_{t+\tau})
\end{aligned}
\tag{2.20}
$$

Some algebra allows us to write this last term as

$$
(1-\beta\theta)\sum_{s=1}^{\infty}(\beta\theta)^s\sum_{\tau=1}^{s}E_{j,t}(\widehat{\pi}_{t+\tau}) = \quad \beta\theta E_{j,t}\widehat{\pi}_{t+1} + (1-\beta\theta)\sum_{s=1}^{\infty}(\beta\theta)^{s+1}\sum_{\tau=1}^{s}E_{j,t}\widehat{\pi}_{t+\tau+1}
$$

substituting this into (2.20) and simplifying yields

$$
\ln P^*_{j,t} = (1-\beta\theta)E_{j,t}\widehat{mc}_{j,t} + E_{j,t}\ln P_t + \beta\theta\left[E_{j,t}\ln P_{j,t+1} + E_{j,t}\widehat{\pi}_{t+1} - E_{j,t}\ln P_{t+1}\right]
$$

By definition, $\widehat{\pi}_{t+1} = \ln P_{t+1} - \ln P_t - \ln \pi_*$. Substitute this in:

$$
\ln P^*_{j,t} = (1-\beta\theta)E_{j,t}\widehat{mc}_{j,t} + (1-\beta\theta)E_{j,t}\ln P_t + \beta\theta E_{j,t}\ln P^*_{j,t+1} - \beta\theta\ln\pi_* \tag{2.21}
$$

Denote the average reset price as $\ln P^*_t = \int \ln P^*_{j,t}dj$. The evolution of the price level can be written

$$
\ln P_t = \theta(\ln P_{t-1} + \ln\pi_*) + (1-\theta)\ln P^*_t \tag{2.22}
$$

Integrate (2.21) across firms and solve for the average reset price

$$
\ln P^*_t = (1-\beta\theta)\widehat{mc}^{(1)}_{t|t} + (1-\beta\theta)\ln P^{(1)}_{t|t} + \beta\theta\ln P^*_{t+1|t}(1) - \beta\theta\ln\pi_* \tag{2.23}
$$

Plug (2.23) into (2.22) to show

$$
\ln P_t = \theta\ln P_{t-1} + (\theta-(1-\theta)\beta\theta)\ln\pi_* + (1-\theta)\left[(1-\beta\theta)\widehat{mc}^{(1)}_{t|t} + (1-\beta\theta)\ln P^{(1)}_{t|t} + \beta\theta\ln P^*_{t+1|t}(1)\right] \tag{2.24}
$$

Given the price index (2.22) and the definition of $\widehat{\pi}_t$, we can show two facts:

$$\ln P^*_{t+1} = \frac{\widehat{\pi}_{t+1}}{1-\theta} + \ln P_t + \ln \pi_*$$
$$\ln P_{t+1} = \widehat{\pi}_{t+1} + \ln P_t + \ln \pi_*$$

We can plug these into (2.24) in order to obtain

$$\begin{aligned}\widehat{\pi}_t + \ln P_{t-1} + \ln \pi_* =& \theta \ln P_{t-1} + (\theta - (1-\theta)\beta\theta) \ln \pi_* \\ &+ (1-\theta)\Bigg[(1-\beta\theta)\widehat{mc}^{(1)}_{t|t} + (1-\beta\theta)\ln P^{(1)}_{t|t} \\ &+ \beta\theta \left(\frac{\widehat{\pi}^{(1)}_{t+1|t}}{1-\theta} + \ln P^{(1)}_{t|t} + \ln \pi_*\right)\Bigg]\end{aligned}$$

Collecting terms:

$$\begin{aligned}\widehat{\pi}_t =& (\theta-1)\ln P_{t-1} + (\theta - (1-\theta)\beta\theta - 1)\ln \pi_* \\ &+ (1-\theta)\left[(1-\beta\theta)\widehat{mc}^{(1)}_{t|t} + (1-\beta\theta)\ln P^{(1)}_{t|t}\right] \\ &+ \beta\theta\widehat{\pi}^{(1)}_{t+1|t} + (1-\theta)\beta\theta \ln P^{(1)}_{t|t} + (1-\theta)\beta\theta \ln \pi_*\end{aligned}$$

Simplifying this expression:[15]

$$\begin{aligned}\widehat{\pi}_t =& (\theta-1)\ln P_{t-1} + (\theta-1)\pi_* + (1-\theta)\ln P^{(1)}_{t|t} \\ &(1-\theta)(1-\beta\theta)\widehat{mc}^{(1)}_{t|t} + \beta\theta\widehat{\pi}^{(1)}_{t+1|t}\end{aligned}$$

Rewriting this yields the expression in the paper (using the fact that $\widehat{mc}^{(1)}_t = \int_j E^j_t(y_t - a^j_t)dj) = y^{(1)}_t - a_t$

$$\pi_t = (1-\theta)\bar{E}_t\pi_t + (1-\theta)(1-\beta\theta)\bar{E}_t y_t - (1-\theta)(1-\beta)a_t + \beta\bar{E}_t\pi_{t+1}$$

[15]Like Melosi (2017), I assume that past inflation is revealed but not used in expectations formation.

2.B Details of the inference problem

2.B.1 Preliminaries

As Kohlhas (2014) notes, the orders of expectation in (2.12) increase in a predictable way, according to Binet's formula

$$\sum_{k=0}^{\bar{k}} \left(\frac{z_1^{k+2} - z_2^{k+2}}{z_1 - z_2} \right)$$

where $z_1 = (1 + \sqrt{5})/2$ and $z_2 = (1 - \sqrt{5})/2$.

At a given order $k > 0$ there will be $\frac{z_1^{k+2} - z_2^{k+2}}{z_1 - z_2}$ orders of expectation, of which $\frac{z_1^{k+1} - z_2^{k+1}}{z_1 - z_2}$ are average expectations of the firms and $\frac{z_1^{k} - z_2^{k}}{z_1 - z_2}$ are central bank expectations.

We can also define appropriate sub-matrices $\bar{H}$ and H^{CB} to grab only the higher-order expectations of firms or households. Figure 2.4 shows the basic idea for $\bar{k} = 5$, where H is 159×159; there are three "fundamental" states, 96 states which correspond with higher order firm expectations, and 60 which correspond with higher order central bank expectations.

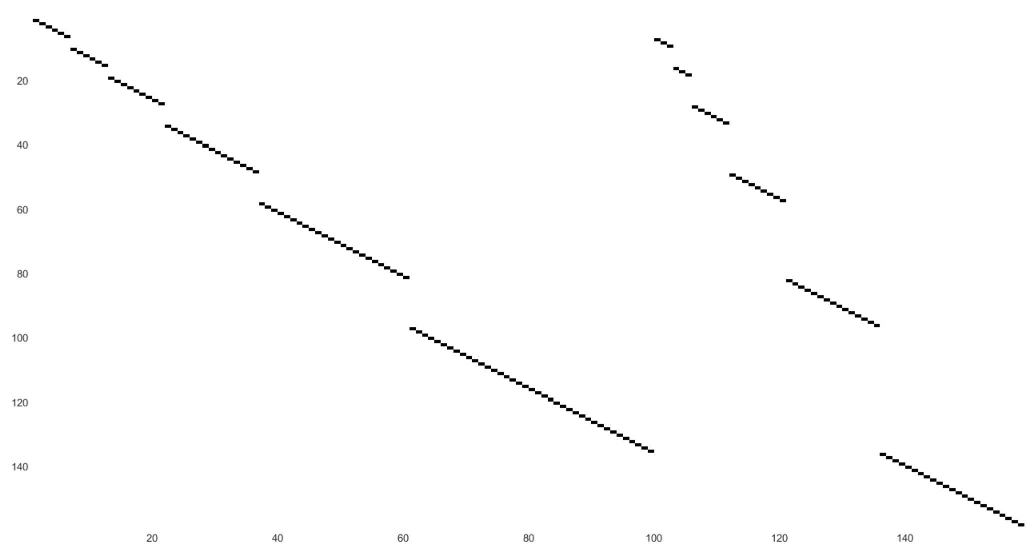


Figure 2.4: Example of H when $\bar{k} = 5$. Black squares indicate positions where there is an element equal to 1, white squares indicate elements equal to 0.

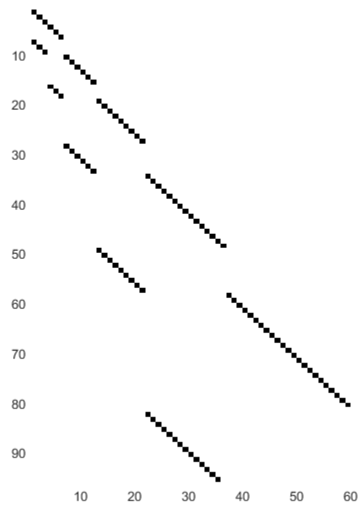


Figure 2.5: Example of R, the reordering matrix that adds in redundant so that $E_t^{CB} X_t^{0:\bar{k}} = R E_t^{CB} A_n X_t^{0:\bar{k}}$. Black boxes denote 1 entries, while white boxes are 0 entries.

R is a matrix that "adds back in" redundant orders of expectation. For example, when $\bar{k} = 5$, R is 96×60. The grid for nonzero values appears in figure 2.5.

2.B.2 The inference problem

Given Gaussian signals with i.i.d. measurement error, and given our assumption that agents have an infinite history of signals, it is optimal for them to use the steady-state Kalman filter for inference. I first characterize the inference problem for firms (assuming a solution exists) before characterizing the inference problem for the central bank. Using the results, I verify that the law of motion for the state takes the proposed form (2.11) in appendix 2.B.3. For ease of exposition, I deal with the truncated vector of higher order expectations, rather than the infinite order one. The derivation is exactly the same, replacing $X_t^{(0:\bar{k})}$ with $X_t^{(0:\infty)}$ throughout.

It will be useful to define a particular re-ordering matrix $\tilde{H}$ implicitly:

$$X_t^{1:\bar{k}+1} = \widetilde{H} \begin{bmatrix} \bar{E}_t X_t^{0:\bar{k}} \\ E_t^{CB} A_n X_t^{0:\bar{k}} \end{bmatrix} \tag{2.25}$$

And a matrix R that "adds back in" the redundant orders of expectation.

The firm inference problem

Recall the idiosyncratic firm j's information set is

$$\Omega_t^j = \left\{a_\tau^j, g_\tau^j, R_\tau; \tau \leq t\right\} \tag{2.26}$$

The signals observed by the jth firm has the truncated representation:

$$z_{jt} = \underbrace{\begin{bmatrix} e_1 \\ T \\ e_3 \end{bmatrix}}_{\bar{D}} X_t^{(0:\bar{k})} + \underbrace{\begin{bmatrix} \tilde{\sigma}_a & 0; \\ 0 & 0 \\ 0 & \tilde{\sigma}_g \end{bmatrix}}_{\bar{Q}} \begin{bmatrix} e_t^{a,j} \\ e_t^{g,j} \end{bmatrix} \tag{2.27}$$

where e_j is a selection vector with a one in the jth entry and zeros elsewhere. (2.27), in combination with (2.11) forms a state-space representation which enables use of the Kalman filter.

The standard Kalman filter derivation (see, for example Hamilton (1994b)) yields the following set of recursive expressions for The firm's current-period state estimate error covariance matrix ($P_{t|t}^j$), its one-step ahead state forecasting error covariance matrix $P_{t+1|t}^j$ and its one step ahead forecasting error matrix for the signals $\Omega_{t|t-1}^j$:

$$\Omega_{t|t-1}^j = \quad \bar{D}P_{t|t-1}\bar{D}' + \bar{Q}\bar{Q}' \tag{2.28}$$

$$P_{t|t-1}^j = \quad MP_{t-1|t-1}M' + NN' \tag{2.29}$$

$$P_{t|t}^j = \quad P_{t|t-1} - P_{t|t-1}\bar{D}'[\Omega_{t|t-1}^j]^{-1}DP_{t|t-1}^j \tag{2.30}$$

Iterating on this system until it converges yields the steady state Kalman gain $\bar{K}$ = which is the same across firms

$$\bar{K} = PD'\Omega^{-1} \tag{2.31}$$

The Kalman filter implies that firm j's expectation of the state is

$$E_t^j X_t^{(0:\bar{k})} = (M - \bar{K}\bar{D}M)E_{t-1}^j X_{t-1}^{(0:\bar{k})} + \bar{K}\bar{D}MX_{t-1}^{(0:\bar{k})} + \bar{K}\bar{D}(N\varepsilon_t + \bar{Q}e_t^j) \tag{2.32}$$

Integrating over firms to find the average expectation yields

$$\bar{E}_t X_t^{0:\bar{k}} \equiv \int_0^1 E_t^j X_t^{(0:\bar{k})} dj = (M - \bar{K}\bar{D}M)\bar{E}^j X_{t-1}^{(0:\bar{k})} + \bar{K}\bar{D}MX_{t-1}^{(0:\bar{k})} + \bar{K}\bar{D}N\varepsilon_t \tag{2.33}$$

since the idiosyncratic shocks are mean zero.

Note our definition of $\widetilde{H}$ implies that

$$\widetilde{H}^{-1} X_t^{1:\bar{k}+1} = \begin{bmatrix} \bar{E}_t X_t^{0:\bar{k}} \\ E_t^{CB} A_n X_t^{0:\bar{k}} \end{bmatrix} \tag{2.34}$$

Then define a matrix $\bar{S}$ that selects only the firm expectations from the right hand side of the previous expression, so we have

$$\underbrace{\begin{bmatrix} I & 0 \end{bmatrix} \tilde{H}^{-1}}_{\bar{S}} X_t^{1:\bar{k}+1} = \bar{E}_t X_t^{0:\bar{k}} \tag{2.35}$$

and we can substitute this into (2.33) to obtain

$$\bar{E}_t X_t^{0:\bar{k}} = (M - \bar{K}\bar{D}M)\bar{S}X_{t-1}^{0:\bar{k}} + \bar{K}\bar{D}MX_t^{0:\bar{k}} + \bar{K}\bar{D}N\varepsilon_t \tag{2.36}$$

Inference problem for the central bank

Recall the central bank's information set is assumed to be

$$\Omega_t^j = \left\{ a_\tau^{CB}, \eta_\tau^{CB}, g_\tau^{CB}; \tau \leq t \right\} \tag{2.37}$$

The central bank's signals are

$$z_t^{CB} = \underbrace{\begin{bmatrix} e_1' \\ e_2' \\ e_3' \end{bmatrix}}_{D^{CB}} X_t^{0:\bar{k}} + \underbrace{\begin{bmatrix} \tilde{\sigma}_a^{CB} & 0; \\ 0 & 0 \\ 0 & \tilde{\sigma}_g^{CB} \end{bmatrix}}_{Q^{CB}} \begin{bmatrix} e_t^{a,CB} \\ e_t^{g,CB} \end{bmatrix} \tag{2.38}$$

Note that the noisy part of the central bank's signals will not average out, which differentiates their error from firms'.

As before (with some obvious changes in notation), the Kalman recursion will yield the following expression for the central bank's beliefs about the state:

$$E_t^{CB} X_t^{0:\bar{k}} = (M - K^{CB} D^{CB} M) E_{t-1}^{CB} X_{t-1}^{0:\bar{k}} + K^{CB} D^{CB} M X_{t-1}^{0:\bar{k}} + K^{CB} (D^{CB} N + Q^{CB}) \varepsilon_t \tag{2.39}$$

Then recalling

$$\tilde{H}^{-1} X_t^{1:\bar{k}+1} = \begin{bmatrix} \bar{E}_t X_t^{0:\bar{k}} \\ E_t^{CB} A_n X_t^{0:\bar{k}} \end{bmatrix} \tag{2.40}$$

we can write

$$\begin{aligned} R \begin{bmatrix} 0 & I \end{bmatrix} \tilde{H}^{-1} X_t^{1:\bar{k}+1} &= R E_t^{CB} A_n X_t^{0:\bar{k}} \\ \Rightarrow S^{CB} X_t^{1:\bar{k}+1} &= R E_t^{CB} A_n X_t \end{aligned} \tag{2.41}$$

Substituting this expression into (2.39) yields the following expression:

$$E_t^{CB} X_t^{0:\bar{k}} = (M - K^{CB} D^{CB} M) S^{CB} X_t^{1:\bar{k}+1} + K^{CB} D^{CB} M X_{t-1}^{0:\bar{k}} + K^{CB} (D^{CB} N + Q^{CB}) \varepsilon_t \tag{2.42}$$

However, our law of motion is written in terms of $E_t A_n X_t^{(0:\bar{k})}$, whereas the above expression contains redundant central bank expectations about its own expectations. To get this in the form of $E_t^{CB} A_n X_t^{0:\bar{k}}$ take the left inverse of R (note that R is not square but the left inverse generally exists)

$$\begin{aligned} E_t^{CB} A_n X_t^{0:\bar{k}} =& R_{\text{left}}^{-1} (M - K^{CB} D^{CB} M) S^{CB} X_t^{1:\bar{k}+1} \\ &+ R_{\text{left}}^{-1} K^{CB} D^{CB} M X_{t-1}^{0:\bar{k}} + R_{\text{left}}^{-1} K^{CB} (D^{CB} N + Q^{CB}) \varepsilon_t \end{aligned} \tag{2.43}$$

The law of motion for the state

We had proposed that the vector of higher-order expectations $X_t^{0:\bar{k}}$ followed a VAR(1) in (2.11). We can combine the results from the previous section to show that

$$
\begin{aligned}
X_{t+1}^{1:\bar{k}+1} =& \tilde{H}\begin{bmatrix} \bar{E}_t X_t^{0:\bar{k}} \\ E_t^{CB} A_n X_t^{0:\bar{k}} \end{bmatrix} \\
&+ \tilde{H}\begin{bmatrix} (M - \bar{K}\bar{D}M)\bar{S} \\ R_{\text{left}}^{-1}(M - K^{CB}D^{CB}M)S^{CB} \end{bmatrix} X_t^{1:\bar{k}+1} \\
&+ \tilde{H}\begin{bmatrix} \bar{K}\bar{D}M \\ R_{\text{left}}^{-1}K^{CB}D^{CB}M \end{bmatrix} X_{t-1}^{0:\bar{k}} \\
&\tilde{H}\begin{bmatrix} \bar{K}\bar{D}N \\ R_{\text{left}}^{-1}K^{CB}(D^{CB}N + Q^{CB}) \end{bmatrix} \varepsilon_t
\end{aligned} \tag{2.44}
$$

and thus (implicitly) derive expressions for M and N:

$$
M = \begin{bmatrix} \Phi & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} + \begin{bmatrix} & \mathbf{0} \\ \mathbf{0} & \tilde{H}\begin{pmatrix} \bar{M}_1 \\ M_1^{CB} \end{pmatrix} \end{bmatrix} \quad + \begin{bmatrix} \mathbf{0} & \\ \tilde{H}\begin{pmatrix} \bar{M}_2 \\ M_2^{CB} \end{pmatrix} & \mathbf{0} \end{bmatrix} \tag{2.45}
$$

$$
N = \begin{bmatrix} \Sigma \\ \tilde{H}\begin{pmatrix} \bar{N} \\ N^{CB} \end{pmatrix} \end{bmatrix} \tag{2.46}
$$

where $\bar{M}_1, M_1^{CB}, \bar{M}_2, M_2^{CB}, \bar{N}, N^{CB}$ are implicitly defined by (2.44).

2.B.3 Showing the proposed law of motion holds

We show that (2.11) holds by combining (2.36) with (2.43) and matching coefficients. Stacking the the latter two equations yields

$$
\begin{aligned}
X_{t+1}^{1:\bar{k}+1} =& \tilde{H}\begin{bmatrix} \bar{E}_t X_t^{0:\bar{k}} \\ E_t^{CB} A_n X_t^{0:\bar{k}} \end{bmatrix} \\
&+ \tilde{H}\begin{bmatrix} (M - \bar{K}\bar{D}M)\bar{S} \\ R_{\text{left}}^{-1}(M - K^{CB}D^{CB}M)S^{CB} \end{bmatrix} X_t^{1:\bar{k}+1} \\
&+ \tilde{H}\begin{bmatrix} \bar{K}\bar{D}M \\ R_{\text{left}}^{-1}K^{CB}D^{CB}M \end{bmatrix} X_{t-1}^{0:\bar{k}} \\
&\tilde{H}\begin{bmatrix} \bar{K}\bar{D}N \\ R_{\text{left}}^{-1}K^{CB}(D^{CB}N + Q^{CB}) \end{bmatrix} \varepsilon_t
\end{aligned} \tag{2.47}
$$

Given the assumed structure for exogenous variables and shocks, we can write the law of motion for x_t as:

$$x_t = \begin{bmatrix} a_t \\ \eta_t \\ g_t \end{bmatrix} = \begin{bmatrix} \rho_a & 0 & 0 \\ 0 & \rho_r & 0 \\ 0 & 0 & \rho_g \end{bmatrix} x_{t-1} + \begin{bmatrix} \sigma_a & 0 & 0 & 0 & 0 \\ 0 & \sigma_r & 0 & 0 & 0 \\ 0 & 0 & \sigma_g & 0 & 0 \end{bmatrix} \begin{bmatrix} \varepsilon_t^a \\ \varepsilon_t^r \\ \varepsilon_t^g \\ e_t^{CB,a} \\ e_t^{CB,g} \end{bmatrix} \tag{2.48}$$

$$x_t = \Phi x_{t-1} + \Sigma \varepsilon_t \tag{2.49}$$

$$\begin{aligned} X_t^{0:\bar{k}+1} = \begin{bmatrix} x_t \\ X_t^{1:\bar{k}+1} \end{bmatrix} &= \begin{bmatrix} \Phi & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} x_{t-1} \\ X_{t-1}^{1:\bar{k}+1} \end{bmatrix} \\ &+ \begin{bmatrix} & \mathbf{0} \\ \mathbf{0} & \tilde{H} \begin{pmatrix} \bar{M}_1 \\ M_1^{CB} \end{pmatrix} \end{bmatrix} \begin{bmatrix} x_{t-1} \\ X_{t-1}^{1:\bar{k}+1} \end{bmatrix} \\ &+ \begin{bmatrix} \mathbf{0} & \\ \tilde{H} \begin{pmatrix} \bar{M}_2 \\ M_2^{CB} \end{pmatrix} & \mathbf{0} \end{bmatrix} \begin{bmatrix} x_{t-1} \\ X_{t-1}^{1:\bar{k}+1} \end{bmatrix} \\ &+ \begin{bmatrix} \Sigma \\ \tilde{H} \begin{pmatrix} \bar{N} \\ N^{CB} \end{pmatrix} \end{bmatrix} \varepsilon_t \end{aligned}$$

Where $\bar{M}_1$, etc are implicitly defined.

We then truncate rows/columns to ensure we respect the highest order of expectation $\bar{k}$.

We then have the following mapping

$$M = \begin{bmatrix} \Phi & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} + \begin{bmatrix} & \mathbf{0} \\ \mathbf{0} & \tilde{H} \begin{pmatrix} \bar{M}_1 \\ M_1^{CB} \end{pmatrix} \end{bmatrix} + \begin{bmatrix} \mathbf{0} & \\ \tilde{H} \begin{pmatrix} \bar{M}_2 \\ M_2^{CB} \end{pmatrix} & \mathbf{0} \end{bmatrix}$$

$$N = \begin{bmatrix} \Sigma \\ \tilde{H} \begin{pmatrix} \bar{N} \\ N^{CB} \end{pmatrix} \end{bmatrix}$$

where $\bar{M}_1, M_1^{CB}, \bar{M}_2, M_2^{CB}, \bar{N}, N^{CB}$ are implicitly defined by (2.44).

2.B.4 $\bar{H}$ and H^{CB}

To find $\bar{H}$ and H^{CB} (the submatrices that select only average firm expectations and central bank expectations, respectively) note that

$$X_t^{0:\bar{k}} = H^{trunc} \begin{bmatrix} x_t \\ \bar{E}_t X_t^{0:\bar{k}} \\ E_t^{CB} A_n X_t^{0:\bar{k}} \end{bmatrix} \tag{2.50}$$

where H^{trunc} enforces the truncation that orders of expectation greater than $\bar{k} = 0$. Then we can write

$$\text{pinv}(H^{trunc}) X_t^{0:\bar{k}} = \begin{bmatrix} x_t \\ \bar{E}_t X_t^{0:\bar{k}} \\ E_t^{CB} A_n X_t^{0:\bar{k}} \end{bmatrix} \tag{2.51}$$

and then premultiplying by appropriate selection matrices will give us H^{CB} and $\bar{H}$.

2.C Details of the structural equations

We begin with:

$$\begin{aligned} \begin{bmatrix} \pi_t \\ y_t \end{bmatrix} = G_C \bar{E}_t \begin{bmatrix} \pi_t \\ y_t \end{bmatrix} + G_F \bar{E}_t \begin{bmatrix} \pi_{t+1} \\ y_{t+1} \end{bmatrix} \\ + RX_F E_t \begin{bmatrix} \pi_{t+1} \\ y_{t+1} \end{bmatrix} + G_R R_t + G_X X_t^{0:\bar{k}} \end{aligned} \tag{2.52}$$

$$R_t = F_C E_t^{CB} \begin{bmatrix} \pi_t \\ y_t \end{bmatrix} + F_X E_t^{CB} X_t^{0:\bar{k}} + e_2 X_t^{0:\bar{k}} \tag{2.53}$$

Next we need to substitute out the expectations terms, using our proposed form of the solution. Particularly, assuming interest rates are observed by everyone:

$$\bar{E}_t \begin{bmatrix} \pi_t \\ y_t \end{bmatrix} = \bar{E}_t (A X_t^{0:\bar{k}} + B R_t) = A \bar{H} X_t^{0:\bar{k}} + B R_t \tag{2.54}$$

$$\bar{E}_t \begin{bmatrix} \pi_{t+1} \\ y_{t+1} \end{bmatrix} = \bar{E}_t (A X_{t+1}^{0:\bar{k}} + B R_{t+1}) = A M \bar{H} X_t^{0:\bar{k}} + B (T M \bar{H}) X_t^{0:\bar{k}} \tag{2.55}$$

$$E_t \begin{bmatrix} \pi_{t+1} \\ y_{t+1} \end{bmatrix} = E_t (A X_{t+1}^{0:\bar{k}} + B R_{t+1}) = (A + BT) M X_t^{0:\bar{k}} \tag{2.56}$$

$$E_t^{CB} \begin{bmatrix} \pi_t \\ y_t \end{bmatrix} = E_t^{CB} (A X_t + B R_t) = A H^{CB} X_t^{0:\bar{k}} + B T X_t^{0:\bar{k}} \tag{2.57}$$

Rewriting the original system by making these substitutions:

$$
\begin{aligned}
\begin{bmatrix}\pi_t \\ y_t\end{bmatrix} &= G_C\left(A\bar{H}X_t^{0:\bar{k}} + BR_t\right) + G_F\left(AM\bar{H}X_t^{0:\bar{k}} + B(TM\bar{H})X_t^{0:\bar{k}}\right) \\
&\quad + RX_F(A+BT)MX_t^{0:\bar{k}} + G_R R_t + G_X X_t^{0:\bar{k}}
\end{aligned} \tag{2.58}
$$

$$
R_t = F_C\left(AH^{CB}X_t^{0:\bar{k}} + BTX_t^{0:\bar{k}}\right) + F_X H^{CB} X_t^{0:\bar{k}} + e_2 X_t^{0:\bar{k}} \tag{2.59}
$$

We can solve for B explicitly:

$$
\begin{bmatrix}B_1 \\ B_2\end{bmatrix} = \begin{bmatrix}(1-\theta) & (1-\theta)(1-\beta\theta) \\ 0 & 0\end{bmatrix}\begin{bmatrix}B_1 \\ B_2\end{bmatrix} + \begin{bmatrix}0 \\ -1\end{bmatrix} \tag{2.60}
$$

which implies

$$
B = \begin{bmatrix}-\frac{1}{\theta}(1-\theta)(1-\beta\theta) \\ -1\end{bmatrix} \tag{2.61}
$$

Second, we can rewrite T solely as a function of A:

$$
T = \begin{bmatrix}\phi_\pi & \phi_y\end{bmatrix}\left(AH^{CB} + \begin{bmatrix}-\frac{1}{\theta}(1-\theta)(1-\beta\theta) \\ -1\end{bmatrix}T\right) + F_X H^{CB} + e_2 \tag{2.62}
$$

which implies

$$
T = \left(-\phi_\pi\frac{1}{\theta}(1-\theta)(1-\beta\theta) - \phi_y\right)T + \left(F_C AH^{CB} + F_X H^{CB} + e_2\right) \tag{2.63}
$$

$$
T = \frac{1}{1-\left(-\phi_\pi\frac{1}{\theta}(1-\theta)(1-\beta\theta) - \phi_y\right)}\left(F_C AH^{CB} + F_X H^{CB} + e_2\right) \tag{2.64}
$$

For completeness, the A expression becomes:

$$
\begin{aligned}
A &= G_C A\bar{H} \\
&\quad + G_F\left(A + B\left(\frac{1}{1-\left(-\phi_\pi\frac{1}{\theta}(1-\theta)(1-\beta\theta) - \phi_y\right)}\left(F_C AH^{CB} + F_X H^{CB} + e_2\right)\right)\right)M\bar{H} \\
&\quad + RX_F\left(A + B\left(\frac{1}{1-\left(-\phi_\pi\frac{1}{\theta}(1-\theta)(1-\beta\theta) - \phi_y\right)}\left(F_C AH^{CB} + F_X H^{CB} + e_2\right)\right)\right)M \\
&\quad + G_X
\end{aligned} \tag{2.65}
$$

Conditional on A, we can find a new expression for T when computing a fixed point.

2.D Figures from calibrated model

2.D.1 Calibration 1

Impulse responses

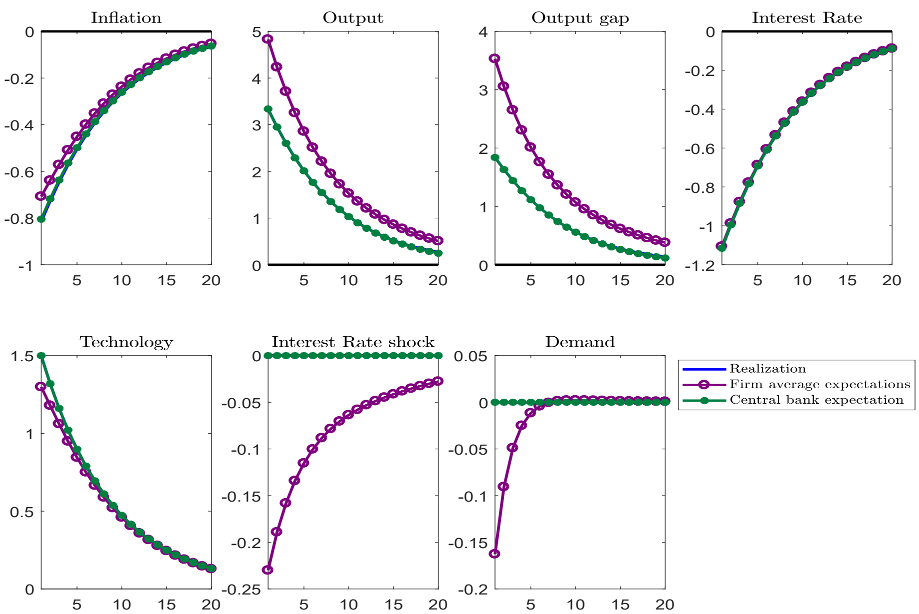


Figure 2.6: Impulse response to TFP shock, calibration 1.

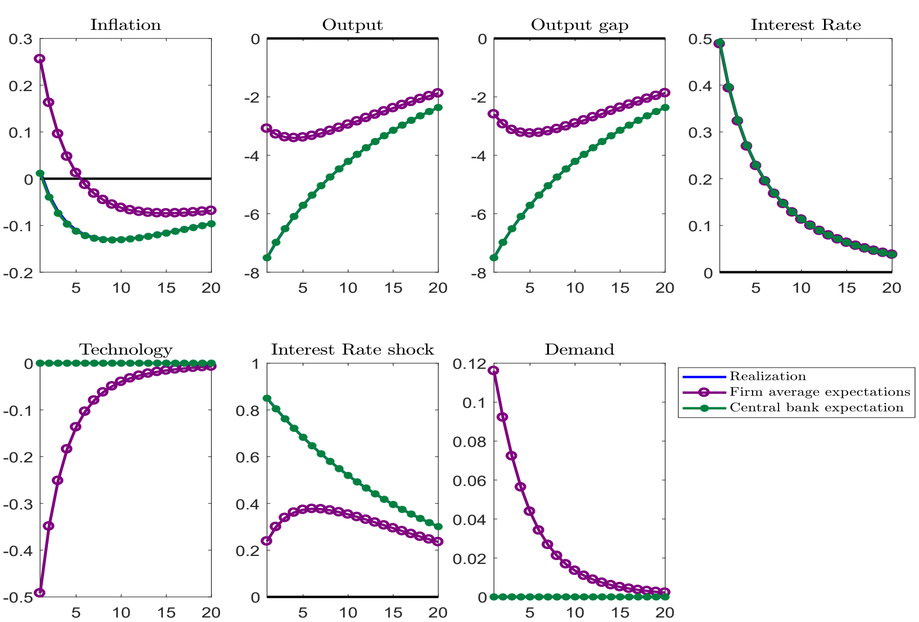


Figure 2.7: Impulse response to monetary policy shock, calibration 1.

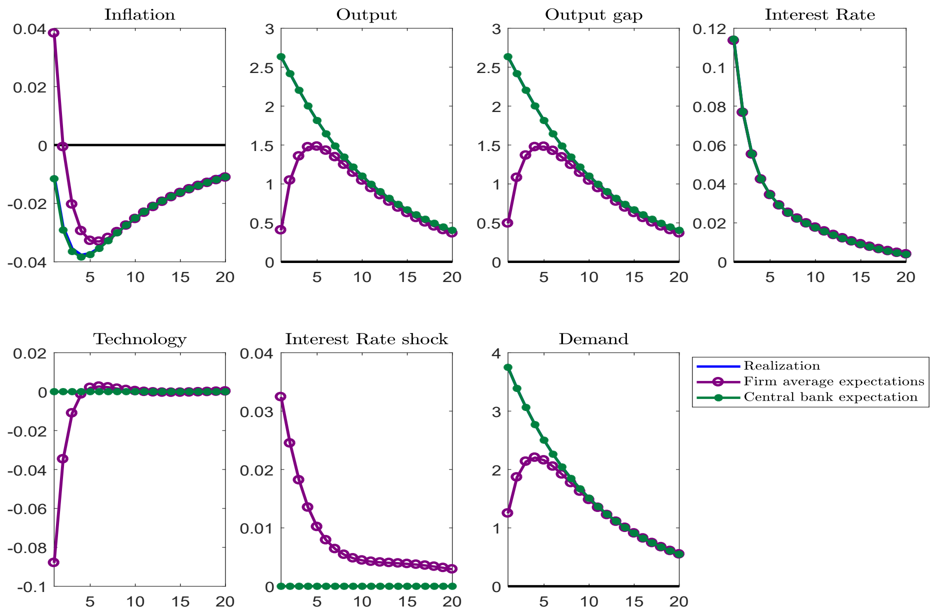


Figure 2.8: Impulse response to demand shock, calibration 1.

Correlations

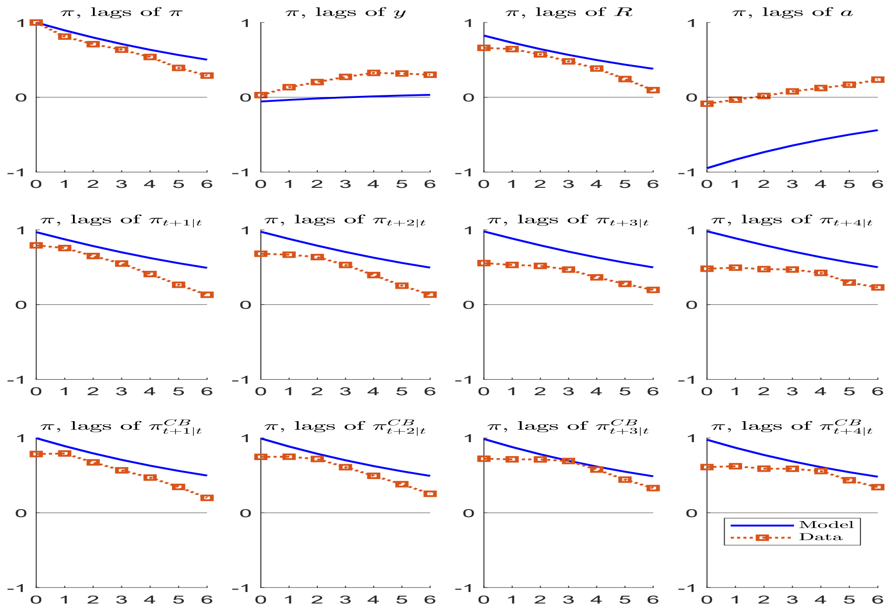


Figure 2.9: Correlation of inflation with other observables, calibration 1.

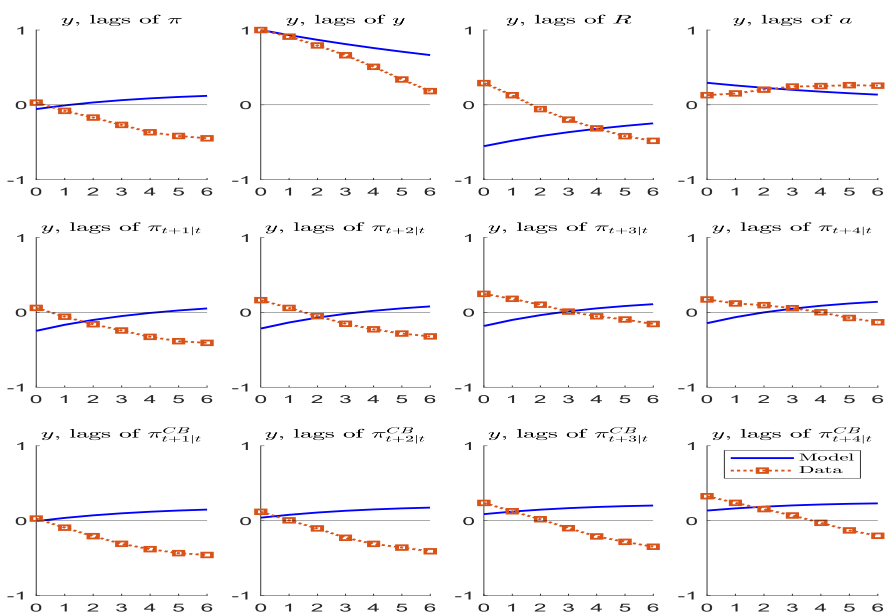


Figure 2.10: Correlation of output with other observables, calibration 1.

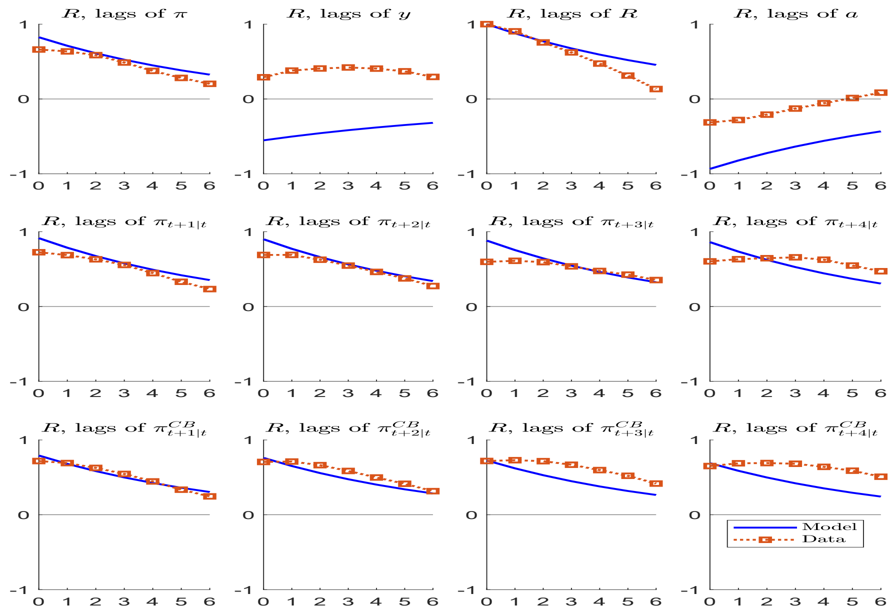


Figure 2.11: Correlations between interest rate and other variables, calibration 1.

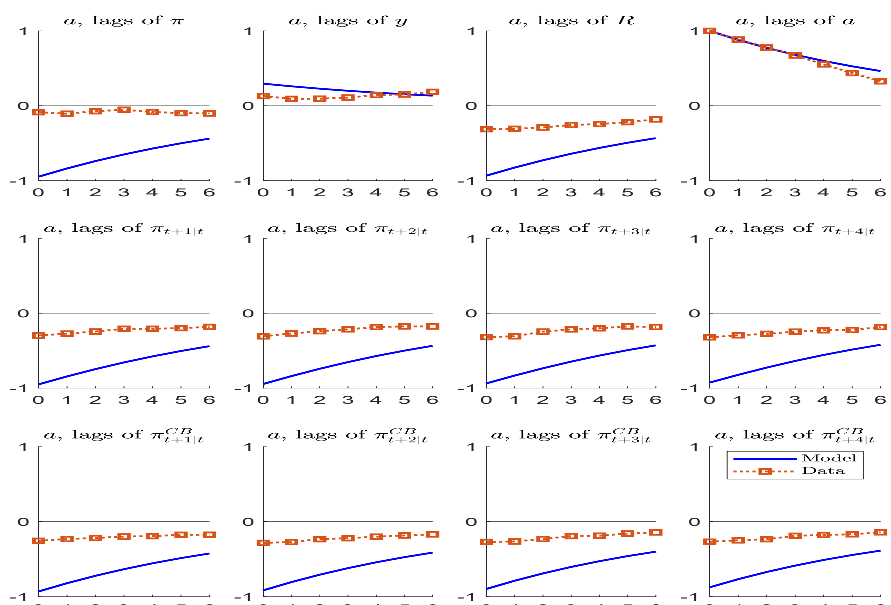


Figure 2.12: Correlations between TFP and other variables, calibration 1.

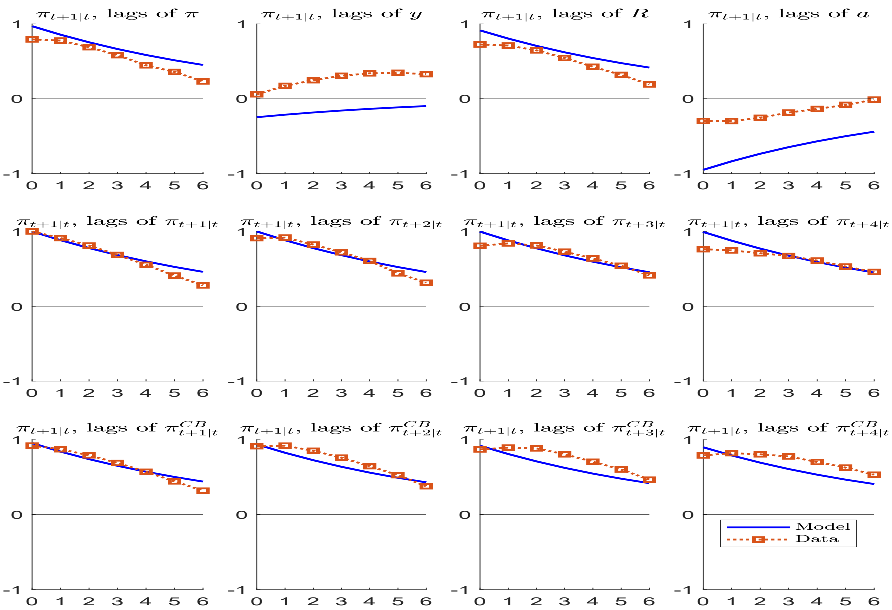


Figure 2.13: Correlations between one-step ahead private inflation expectations and other variables, calibration 1.

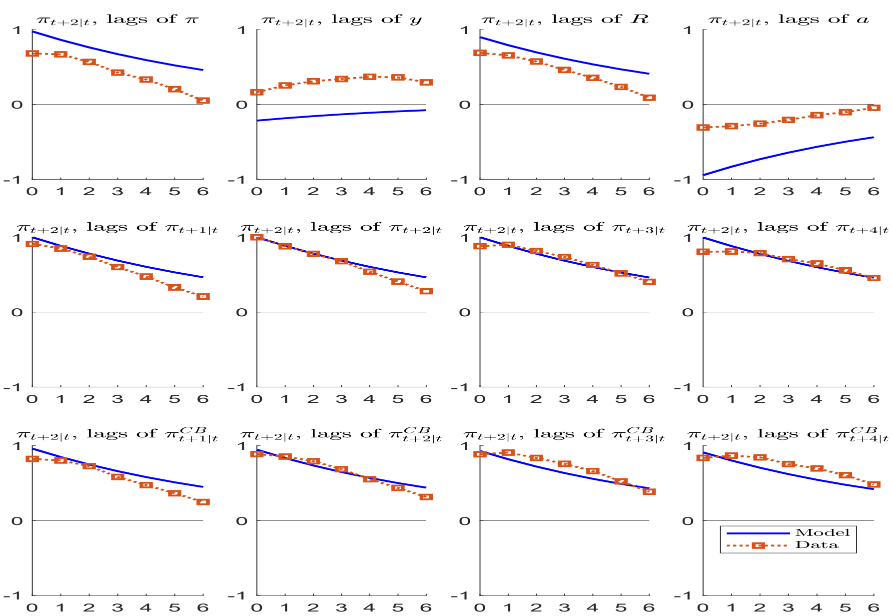


Figure 2.14: Correlations between two step ahead private inflation expectations and other variables, calibration 1.

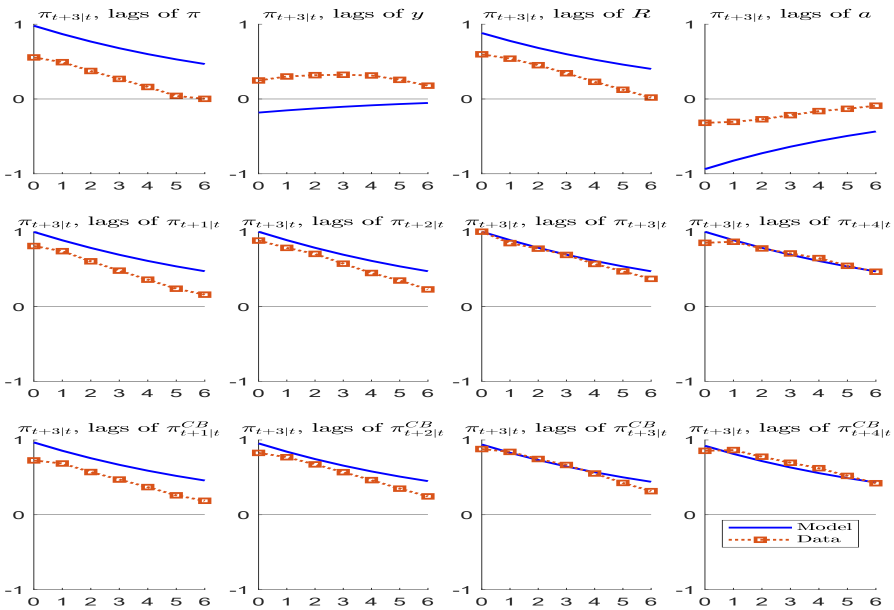


Figure 2.15: Correlations between three step ahead private inflation expectations and other variables, calibration 1.

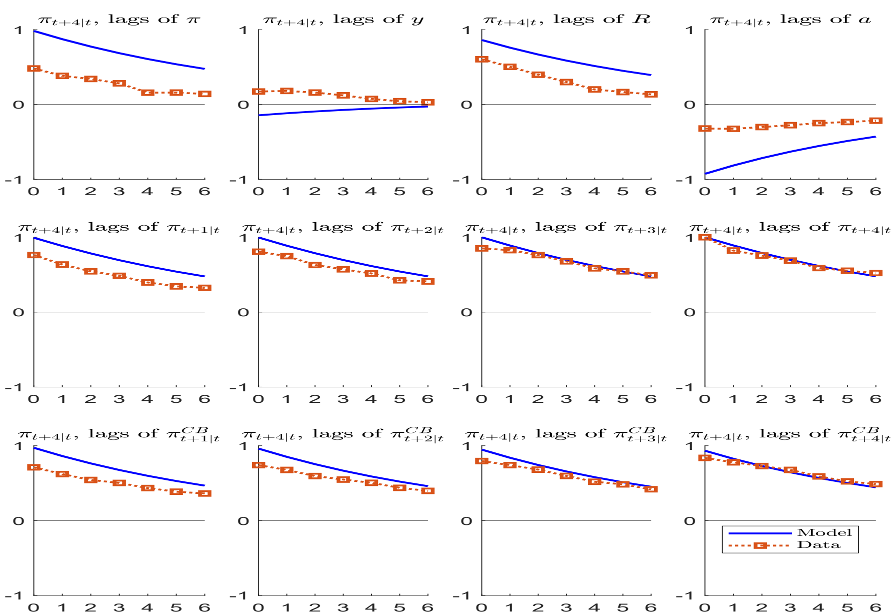


Figure 2.16: Correlations between four step ahead private inflation expectations and other variables, calibration 1.

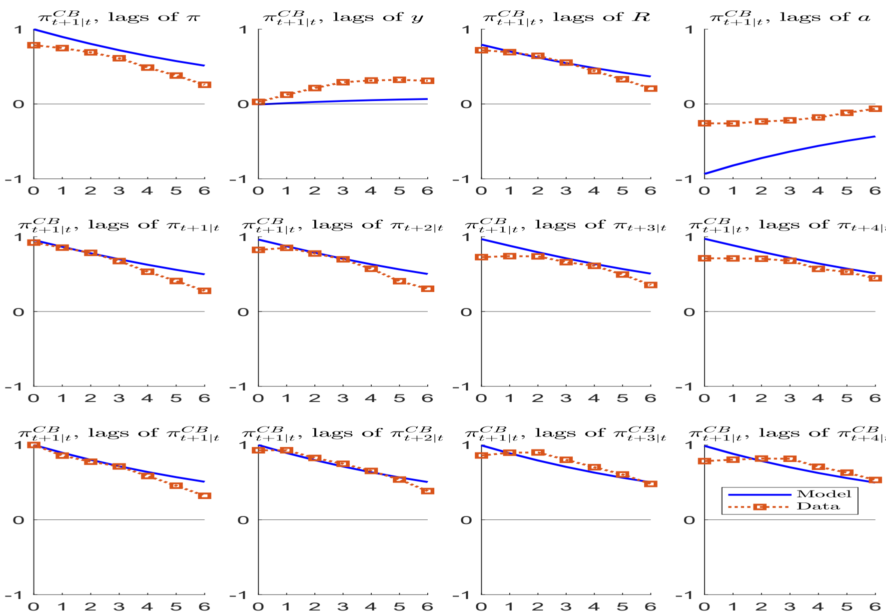


Figure 2.17: Correlations between one step ahead central bank inflation expectations and other variables, calibration 1.

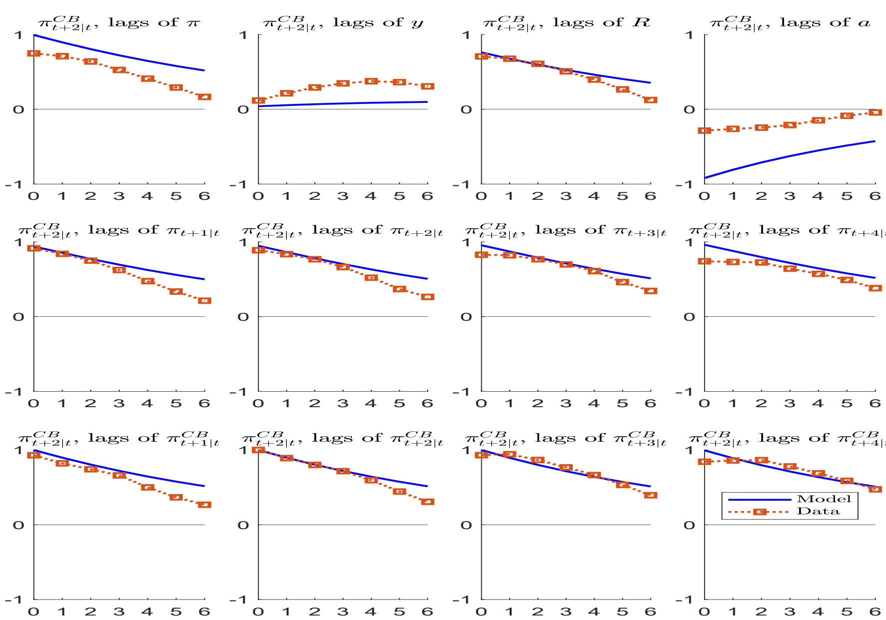


Figure 2.18: Correlations between two step ahead central bank inflation expectations and other variables, calibration 1.

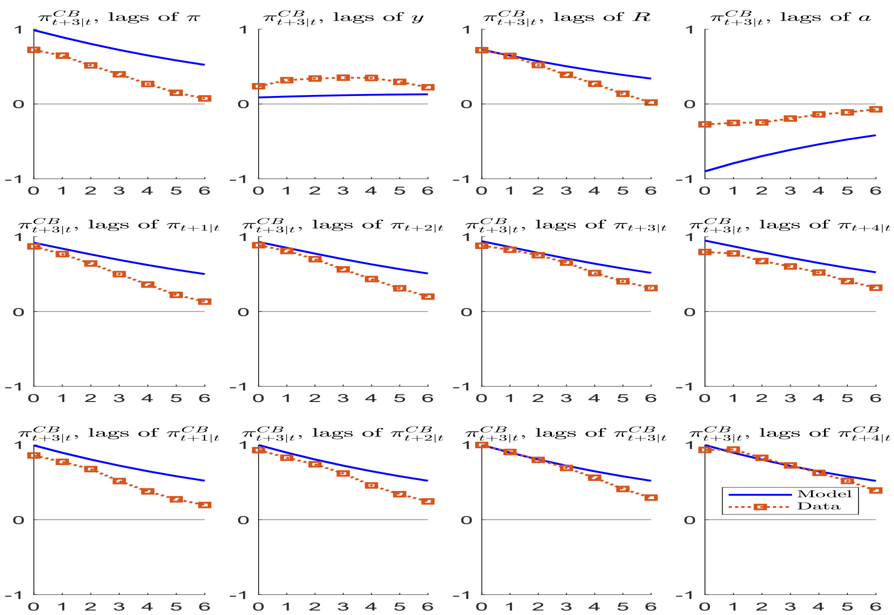


Figure 2.19: Correlations between three step ahead central bank inflation expectations and other variables, calibration 1.

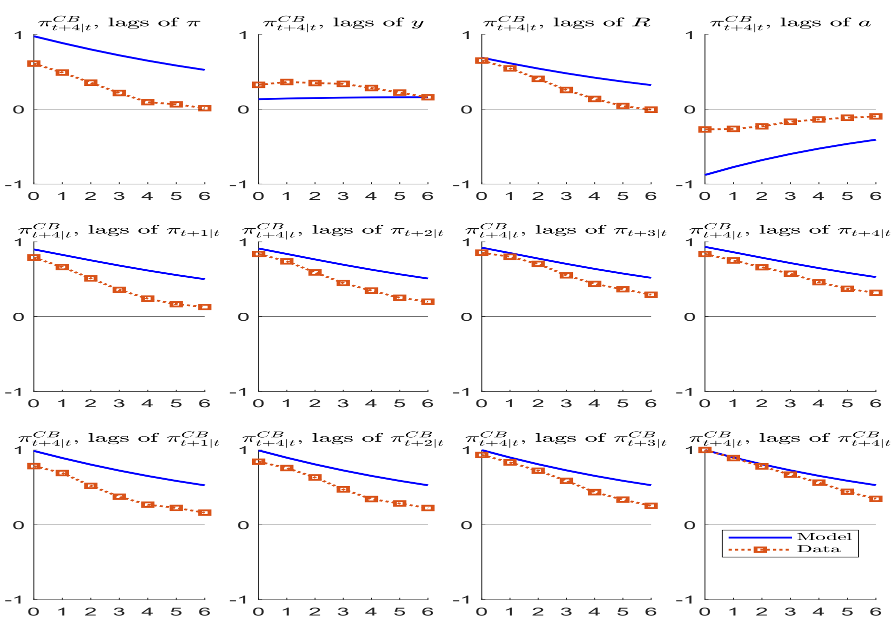


Figure 2.20: Correlations between four step ahead central bank inflation expectations and other variables, calibration 1.

2.D.2 Calibration 2

Impulse responses

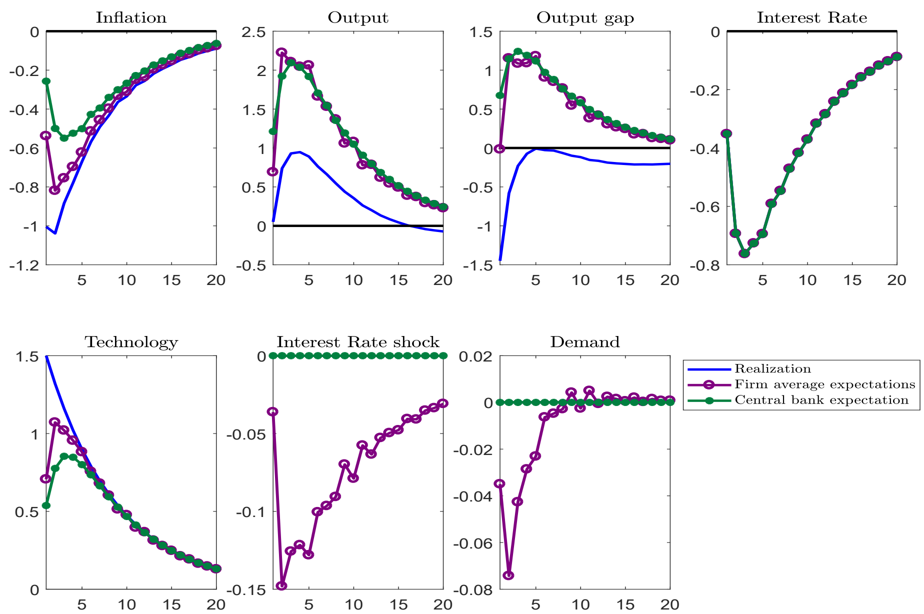


Figure 2.21: Impulse response to TFP shock, calibration 2.

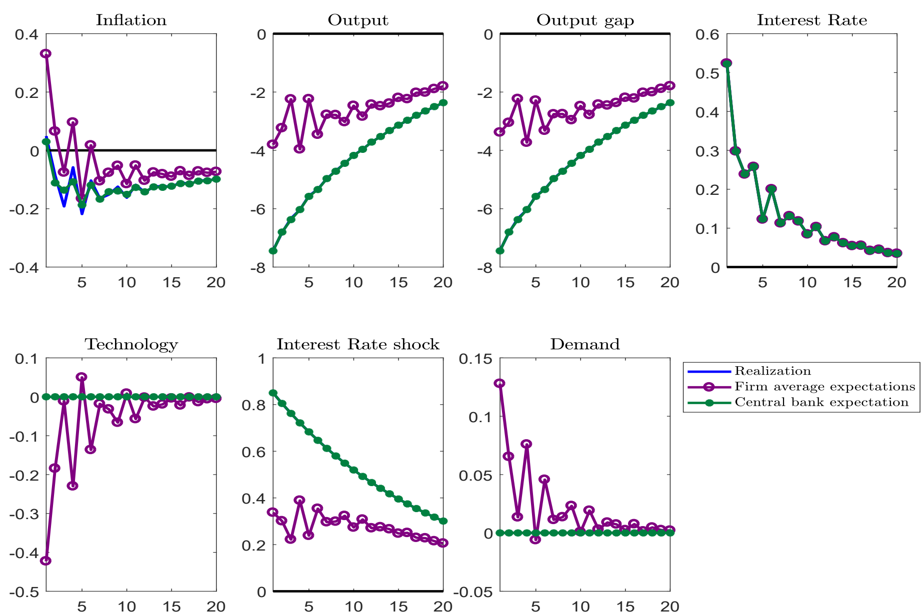


Figure 2.22: Impulse response to monetary policy shock, calibration 2.

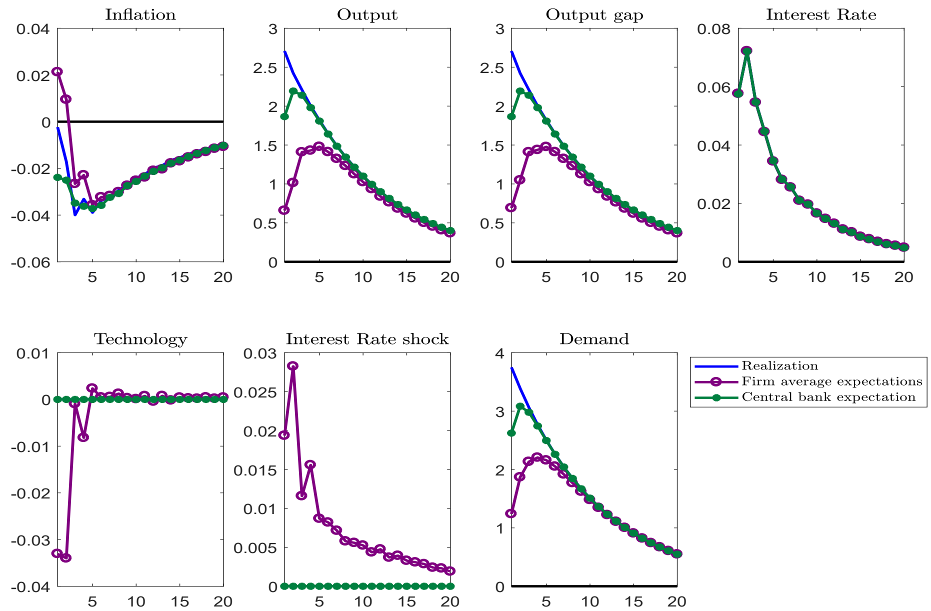


Figure 2.23: Impulse response to demand shock, calibration 2.

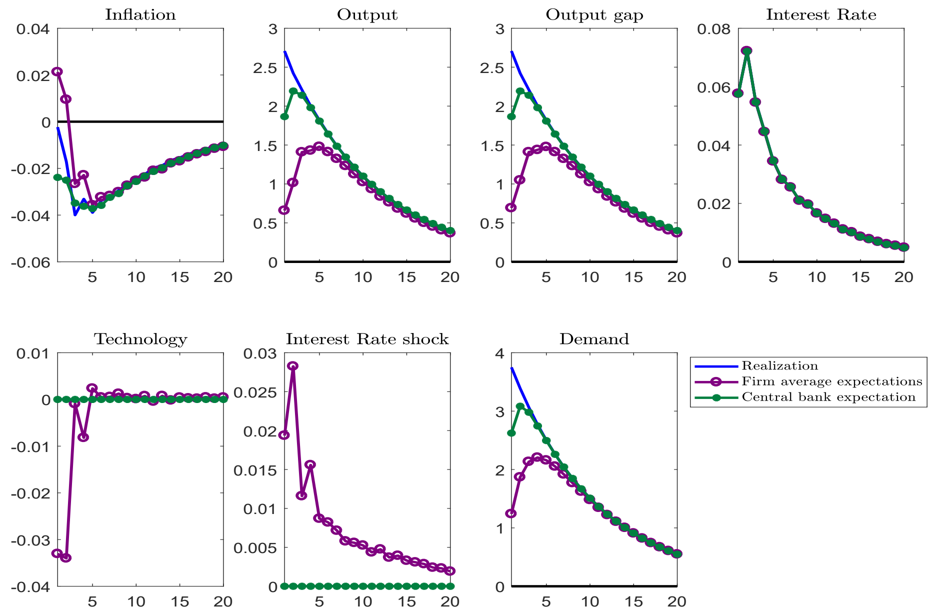


Figure 2.24: Impulse response to demand shock, calibration 2.

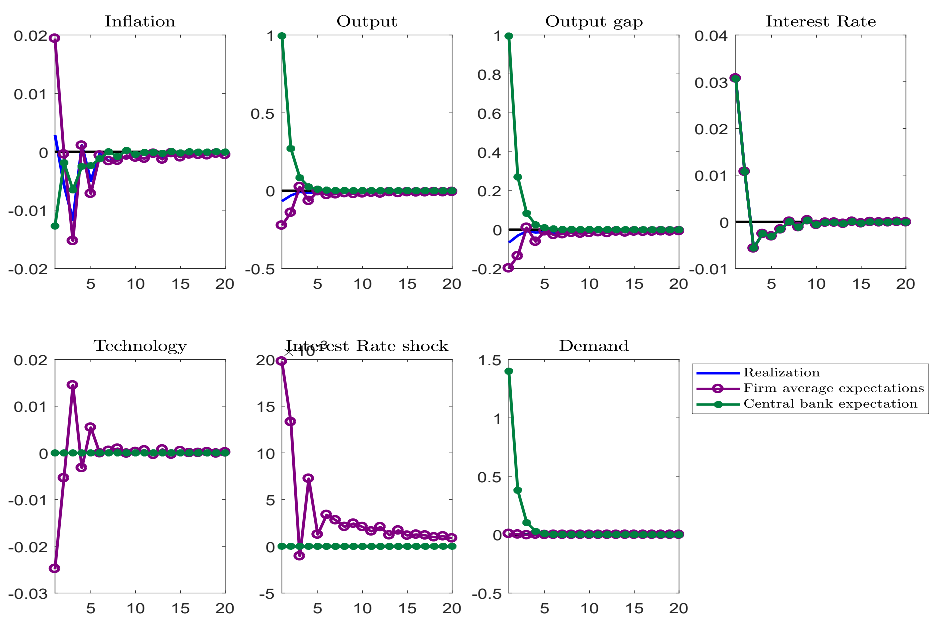


Figure 2.25: Impulse response to central bank demand measurement error shock shock, calibration 2.

Correlations

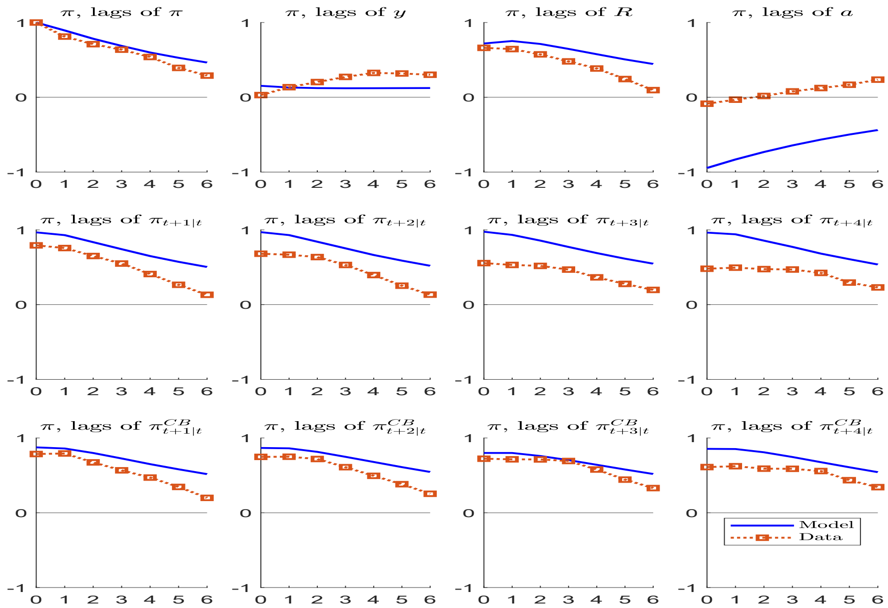


Figure 2.26: Correlation of inflation with other observables, calibration 2.

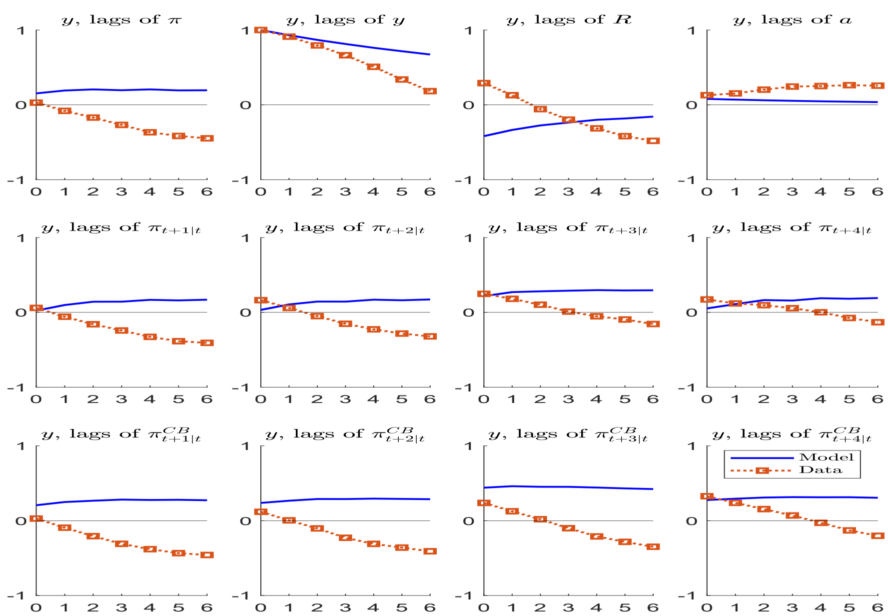


Figure 2.27: Correlation of output with other observables, calibration 2.

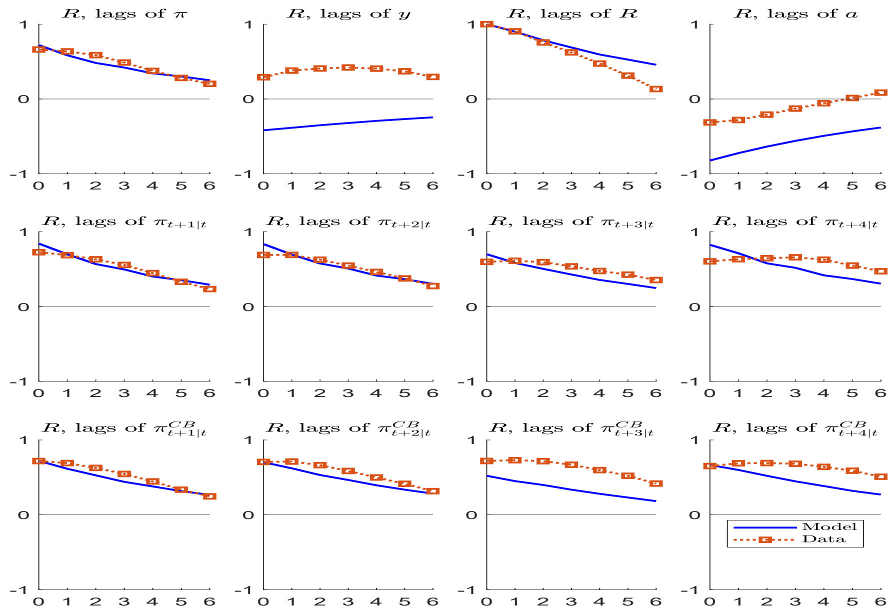


Figure 2.28: Correlations between interest rate and other variables, calibration 2.

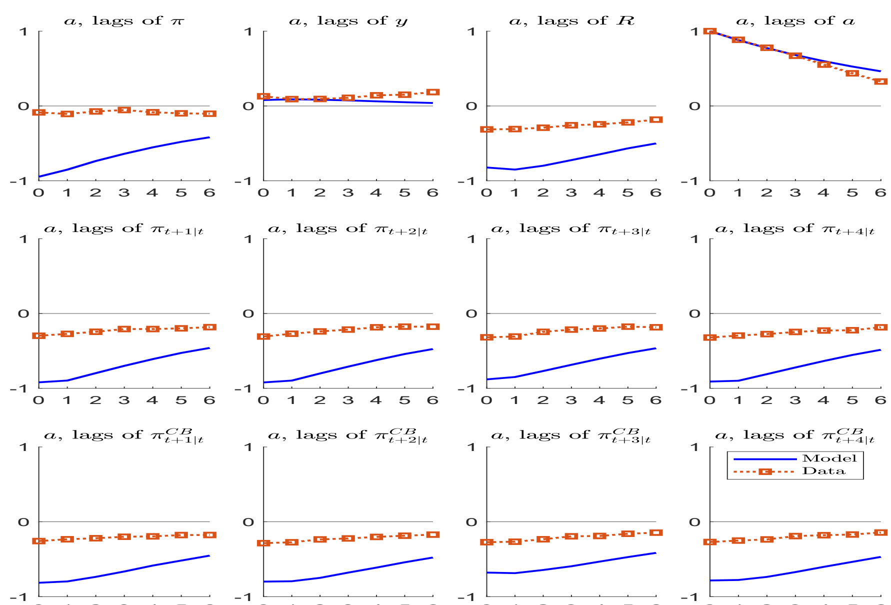


Figure 2.29: Correlations between TFP and other variables, calibration 2.

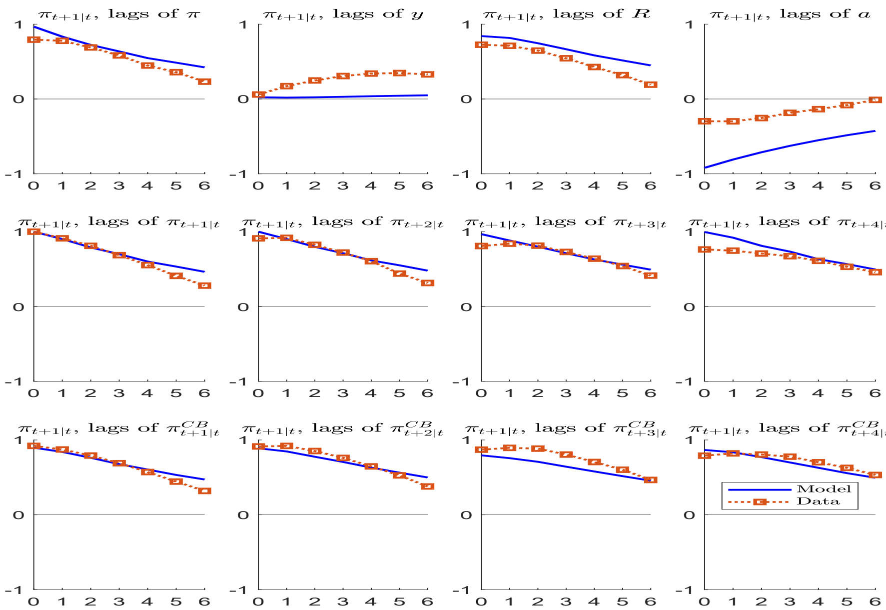


Figure 2.30: Correlations between one-step ahead private inflation expectations and other variables, calibration 2.

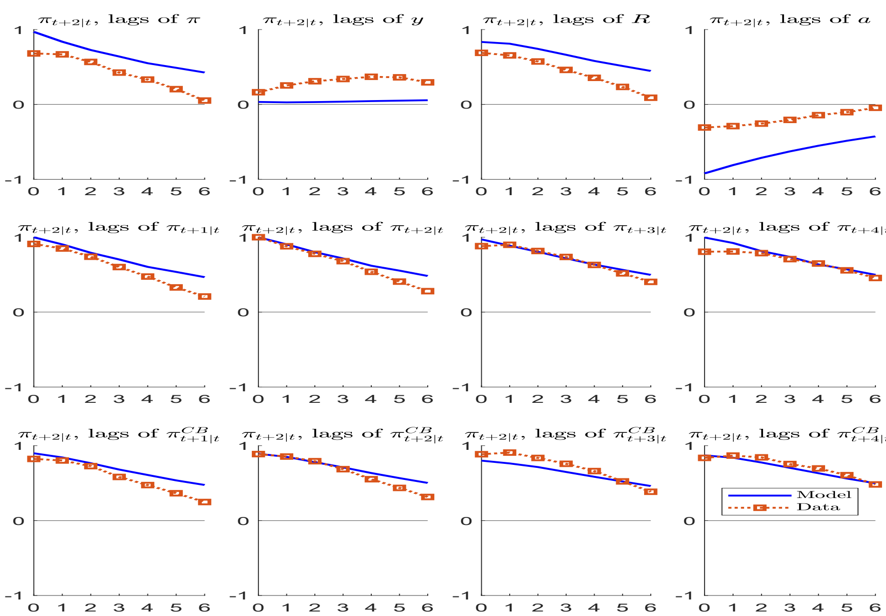


Figure 2.31: Correlations between two step ahead private inflation expectations and other variables, calibration 2.

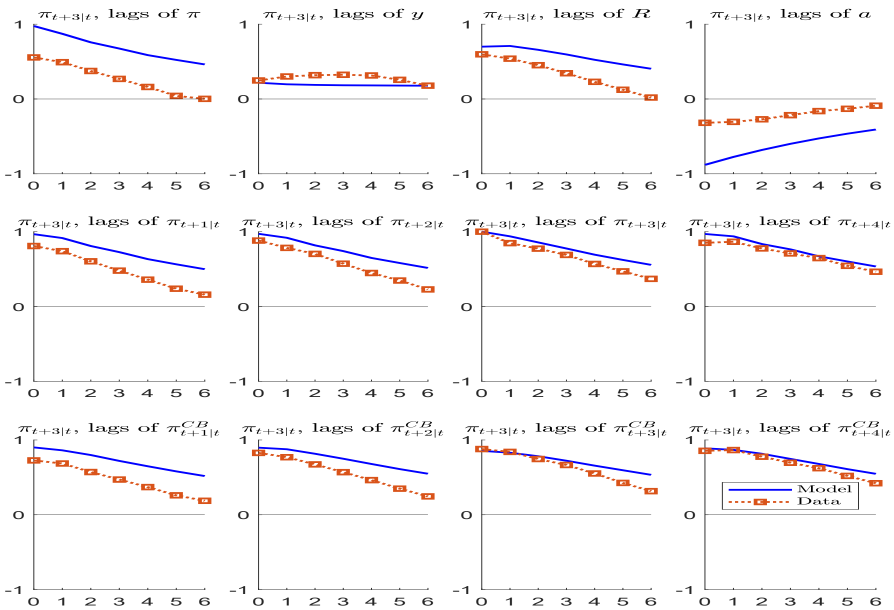


Figure 2.32: Correlations between three step ahead private inflation expectations and other variables, calibration 2.

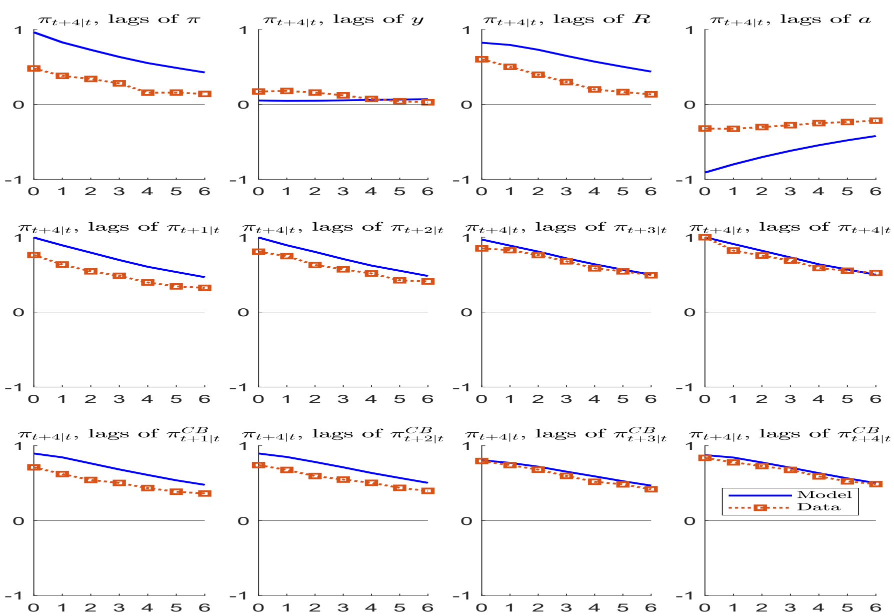


Figure 2.33: Correlations between four step ahead private inflation expectations and other variables, calibration 2.

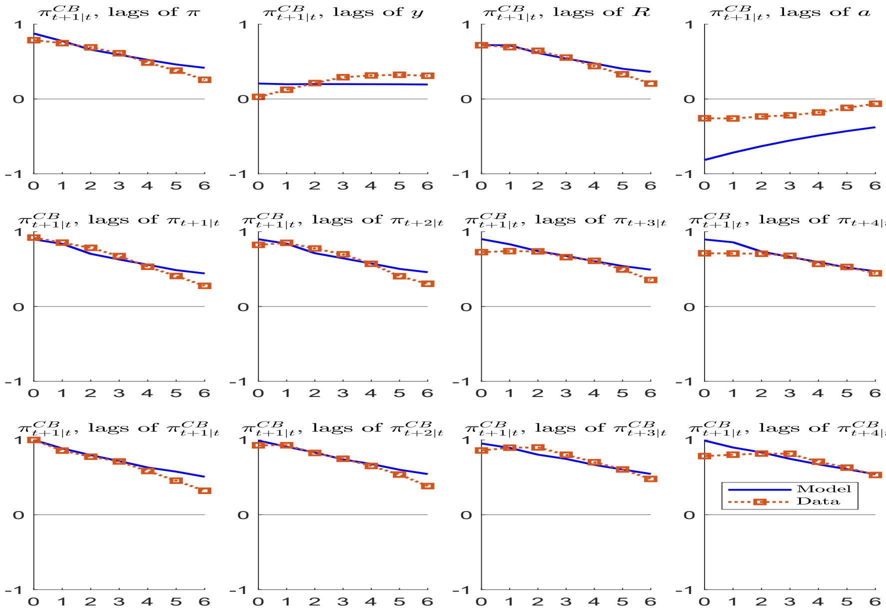


Figure 2.34: Correlations between one step ahead central bank inflation expectations and other variables, calibration 2.

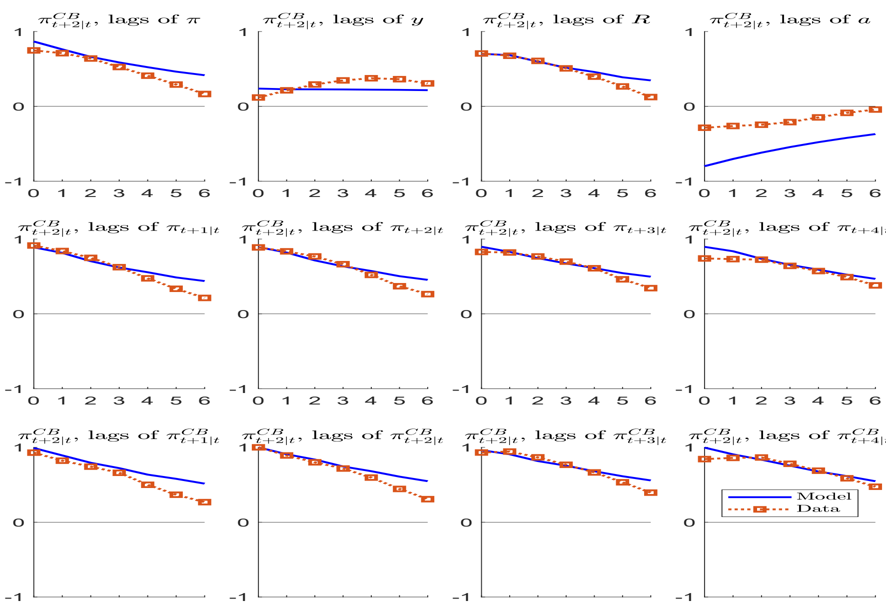


Figure 2.35: Correlations between two step ahead central bank inflation expectations and other variables, calibration 2.

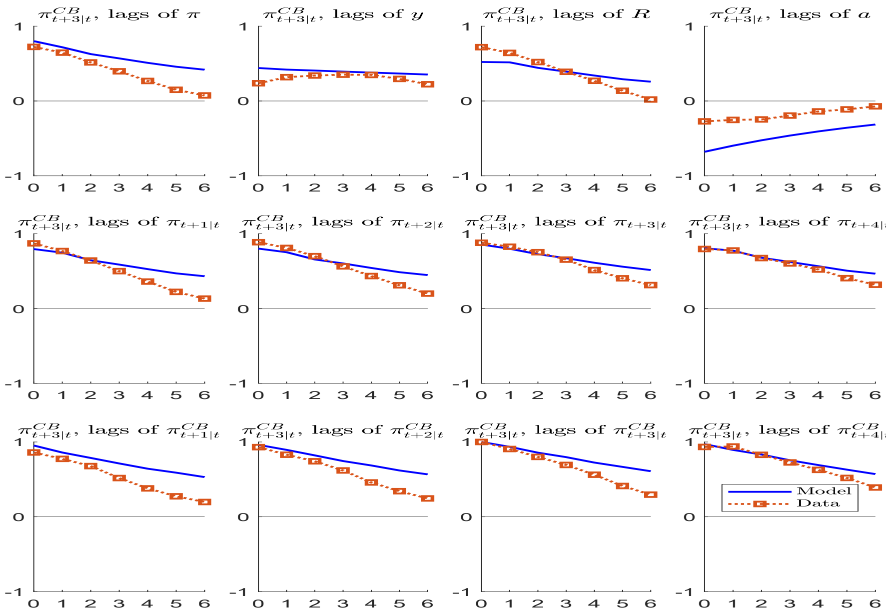


Figure 2.36: Correlations between three step ahead central bank inflation expectations and other variables, calibration 2.

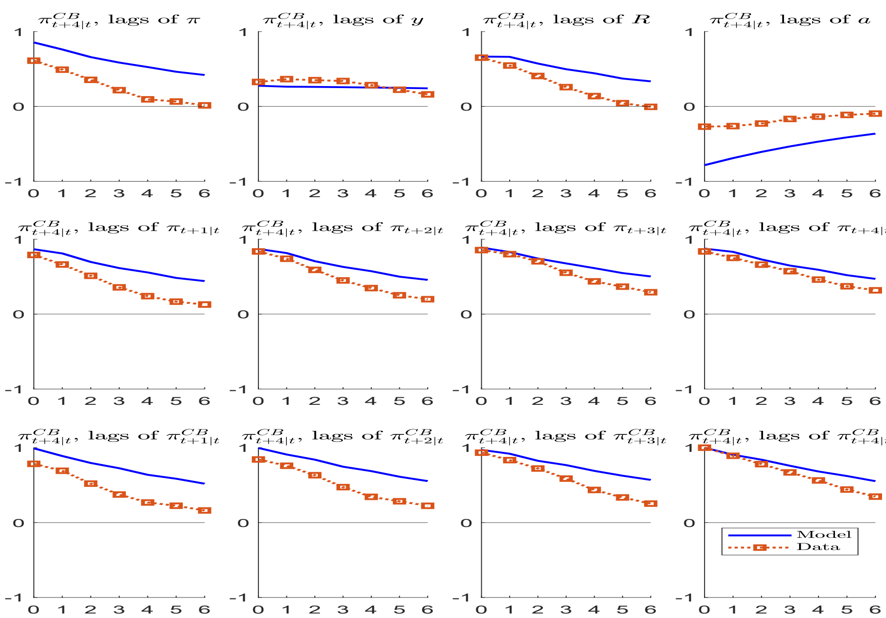


Figure 2.37: Correlations between four step ahead central bank inflation expectations and other variables, calibration 2.

2.D.3 Calibration 3 and 4

Impulse responses

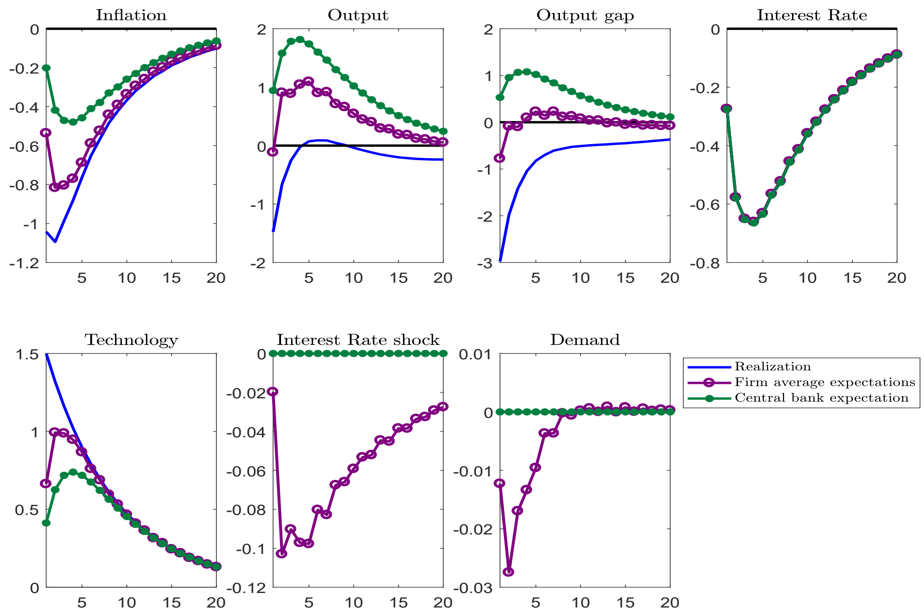


Figure 2.38: Impulse response to TFP shock, calibration 3.

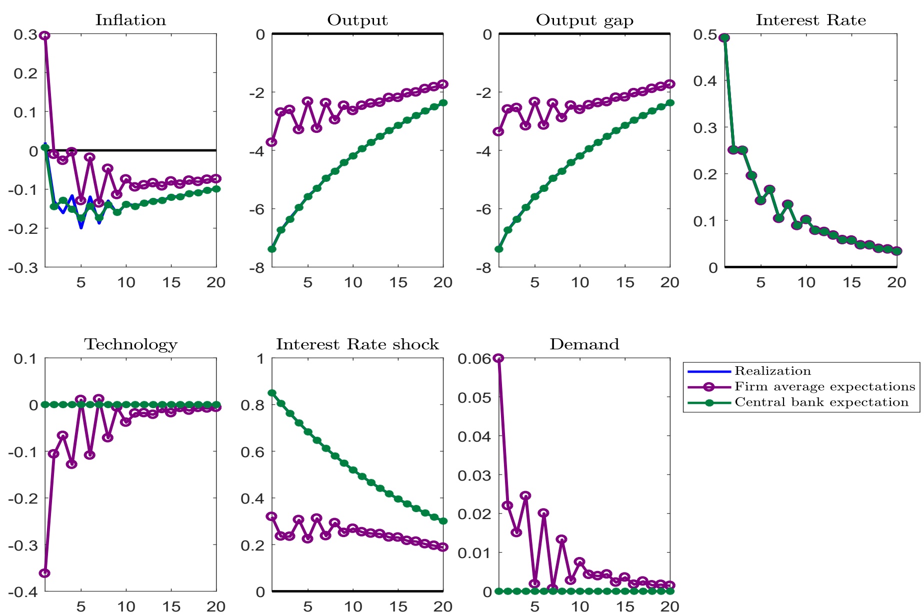


Figure 2.39: Impulse response to monetary policy shock, calibration 3.

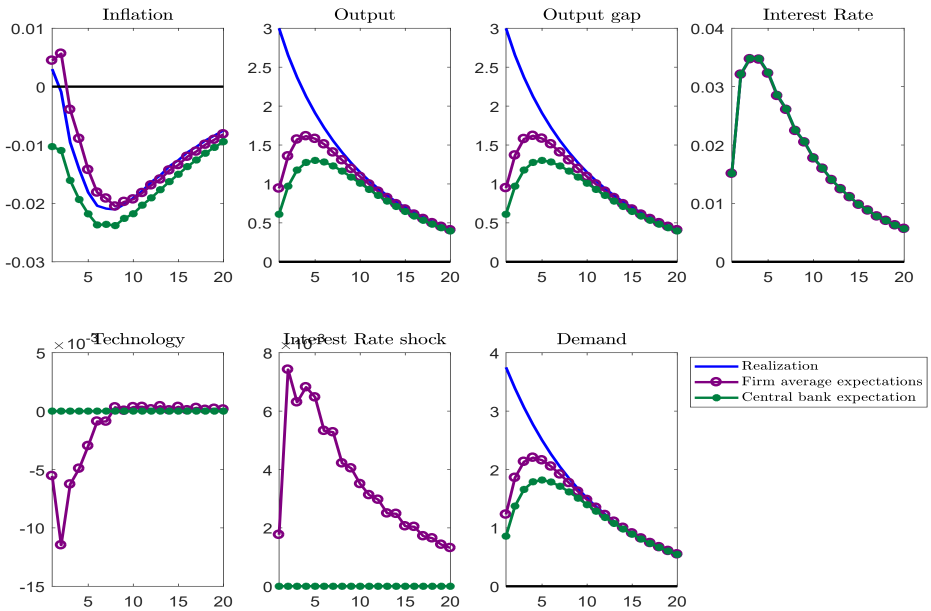


Figure 2.40: Impulse response to demand shock, calibration 3.

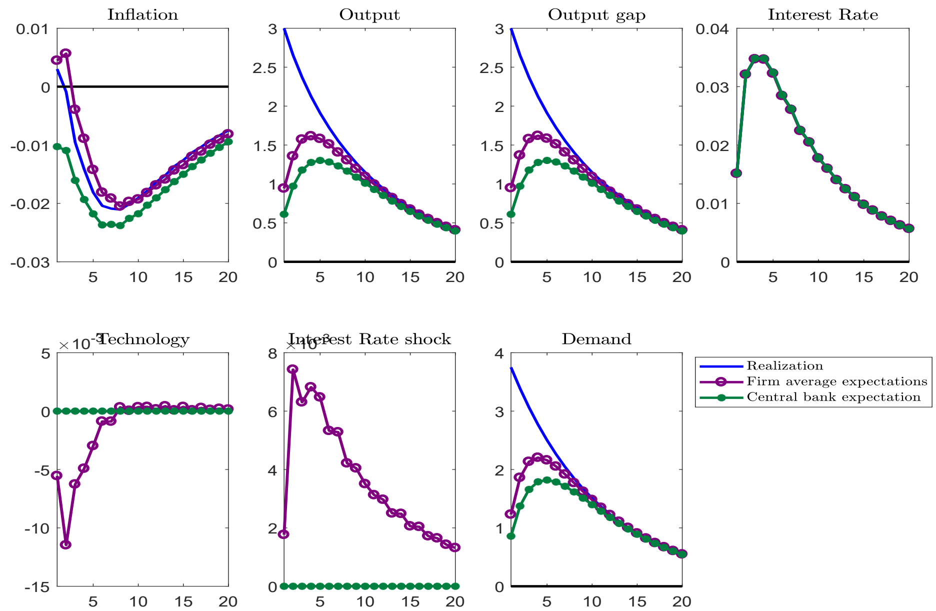


Figure 2.41: Impulse response to demand shock, calibration 3.

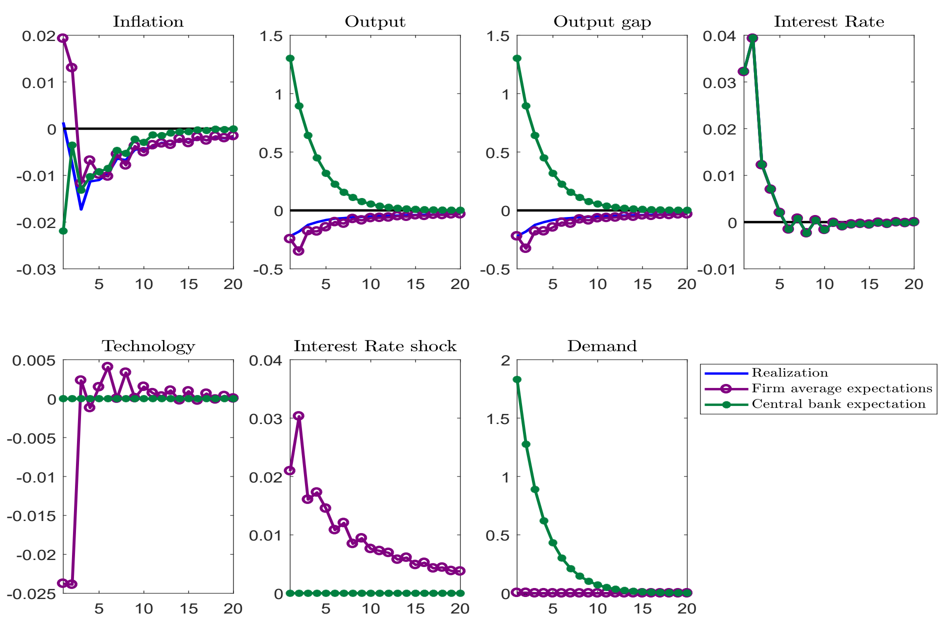


Figure 2.42: Impulse response to central bank demand measurement error shock shock, calibration 3.

Correlations

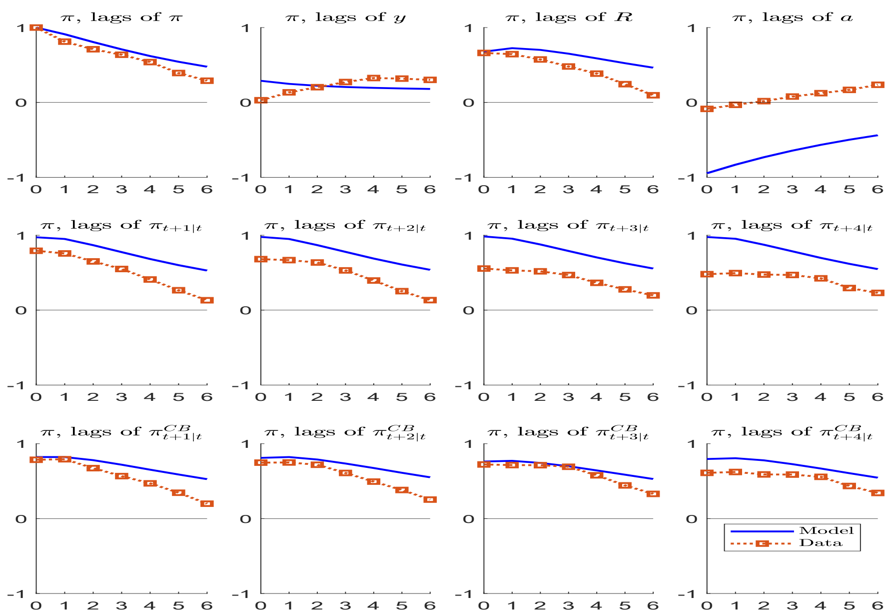


Figure 2.43: Correlation of inflation with other observables, calibration 3.

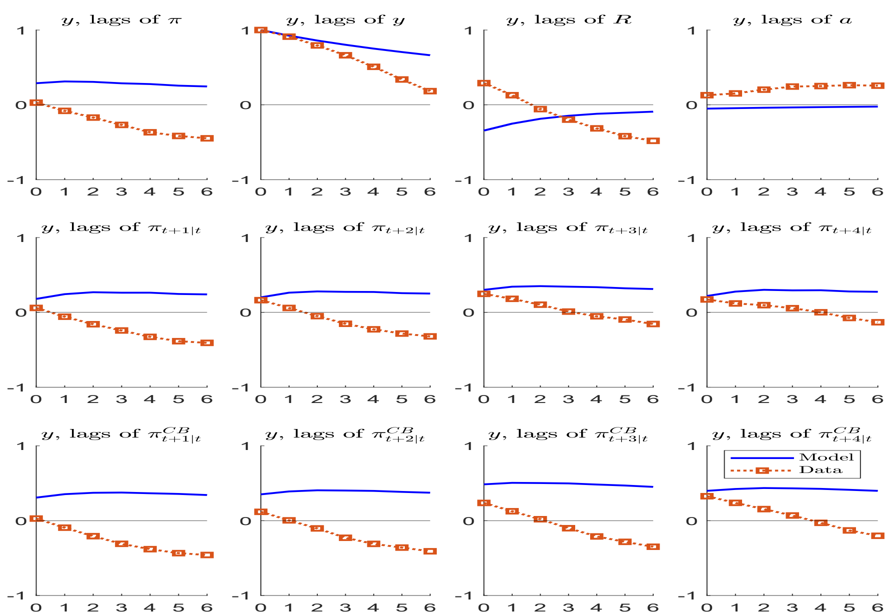


Figure 2.44: Correlation of output with other observables, calibration 3.

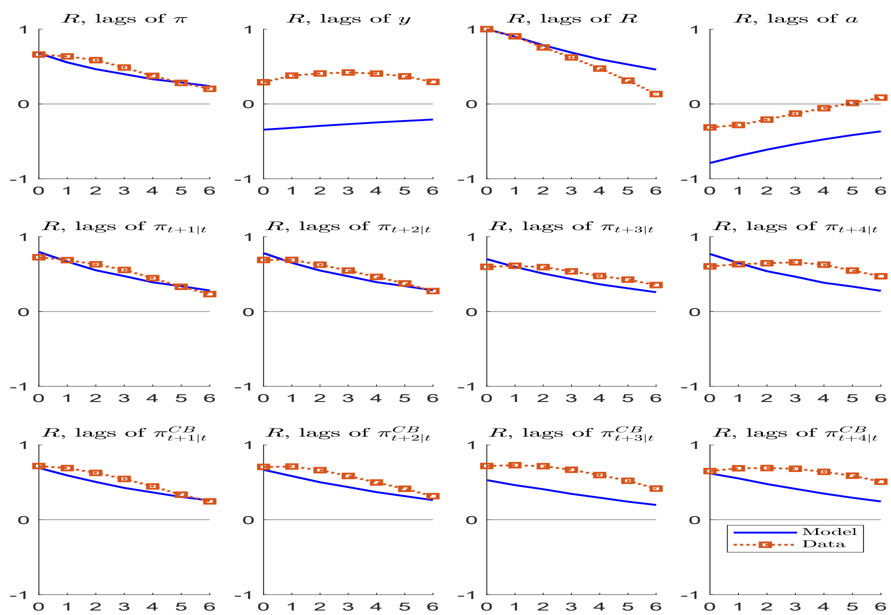


Figure 2.45: Correlations between interest rate and other variables, calibration 3.

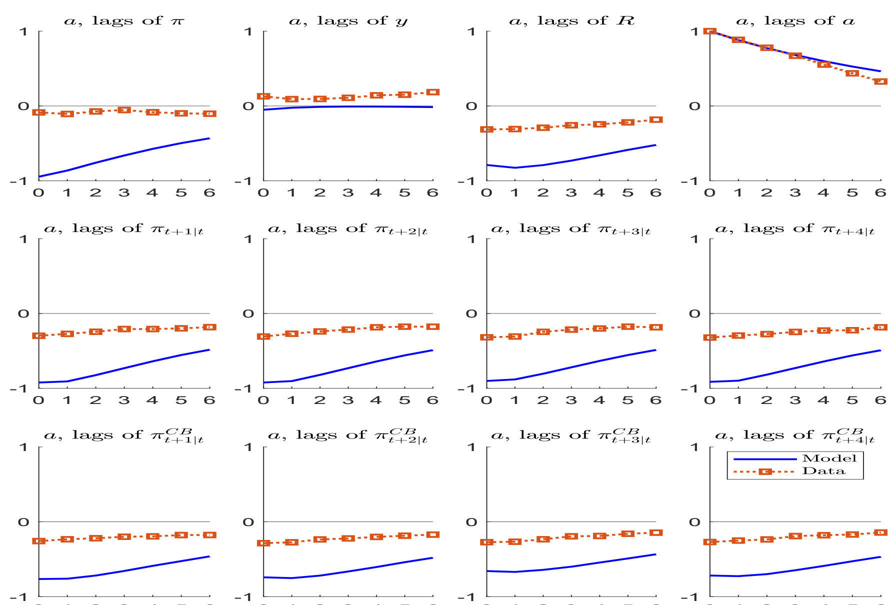


Figure 2.46: Correlations between TFP and other variables, calibration 3.

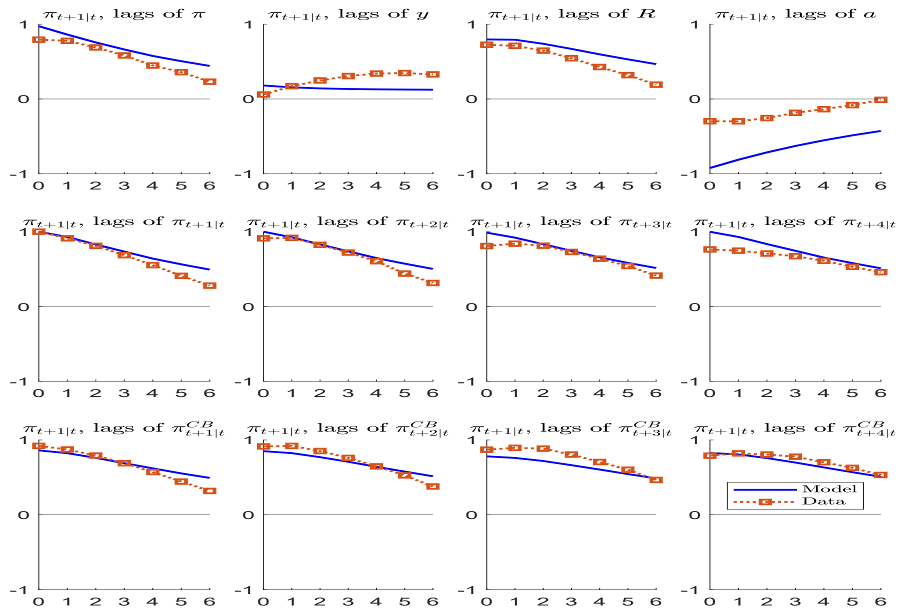


Figure 2.47: Correlations between one-step ahead private inflation expectations and other variables, calibration 3.

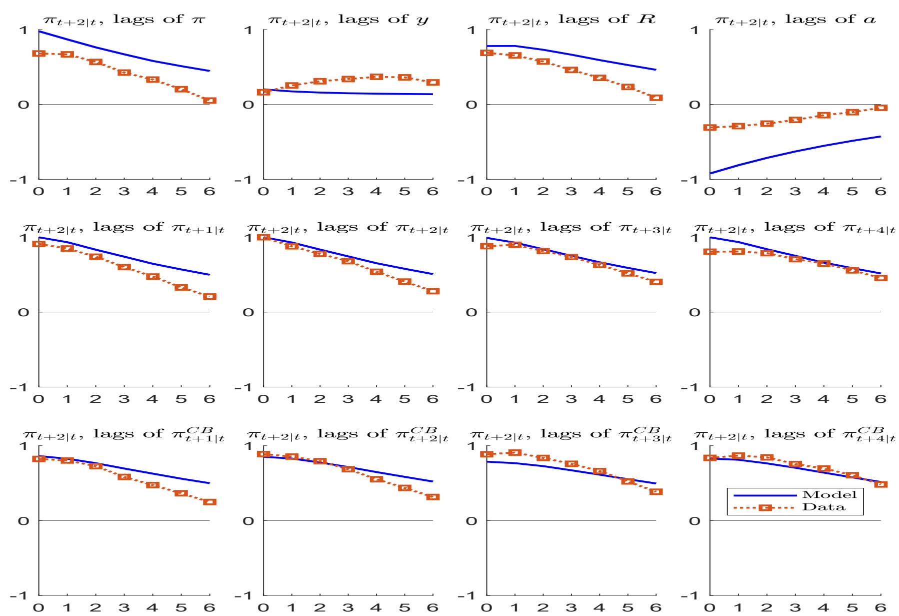


Figure 2.48: Correlations between two step ahead private inflation expectations and other variables, calibration 3.

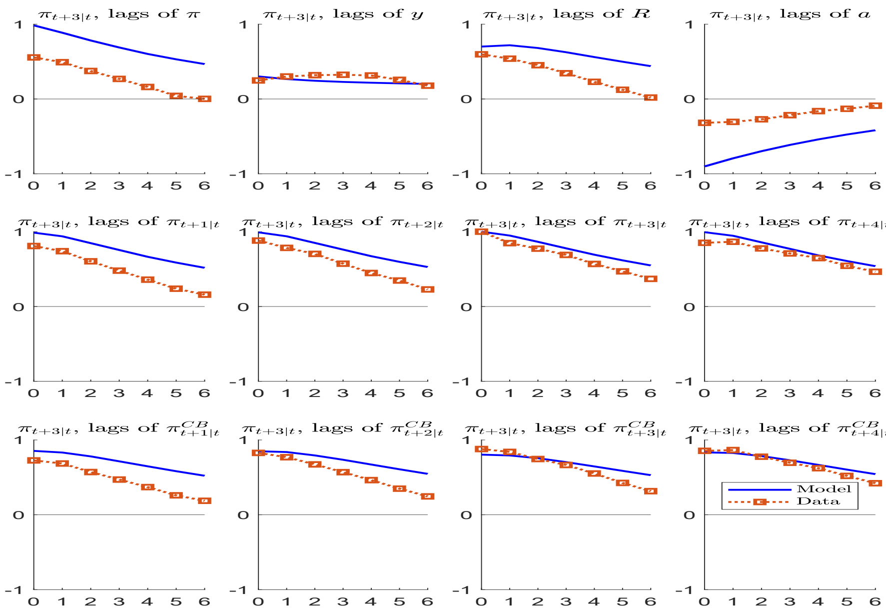


Figure 2.49: Correlations between three step ahead private inflation expectations and other variables, calibration 3.

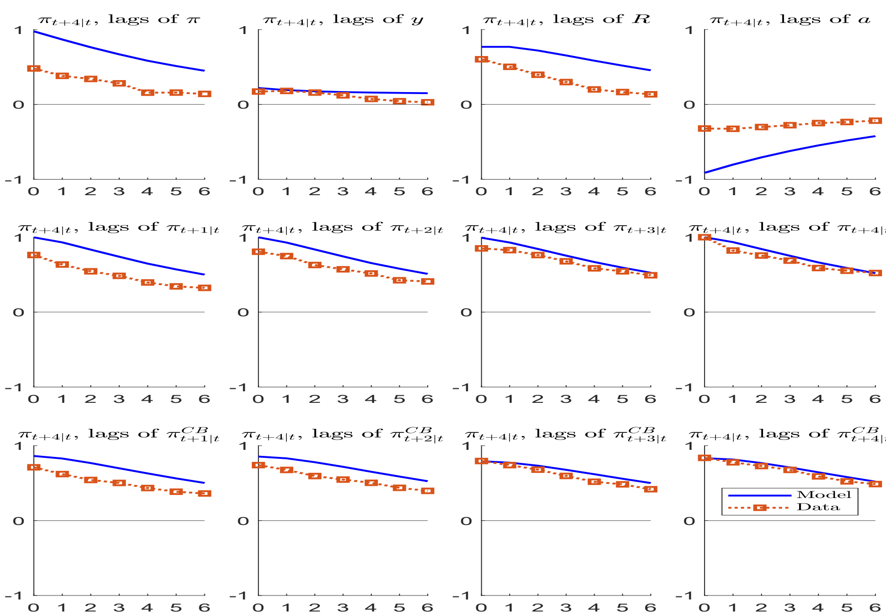


Figure 2.50: Correlations between four step ahead private inflation expectations and other variables, calibration 3.

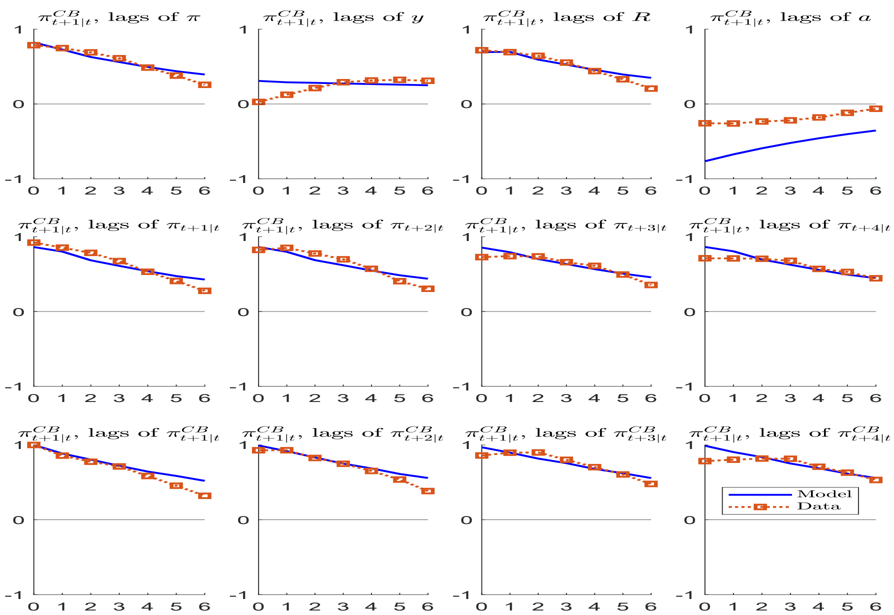


Figure 2.51: Correlations between one step ahead central bank inflation expectations and other variables, calibration 3.

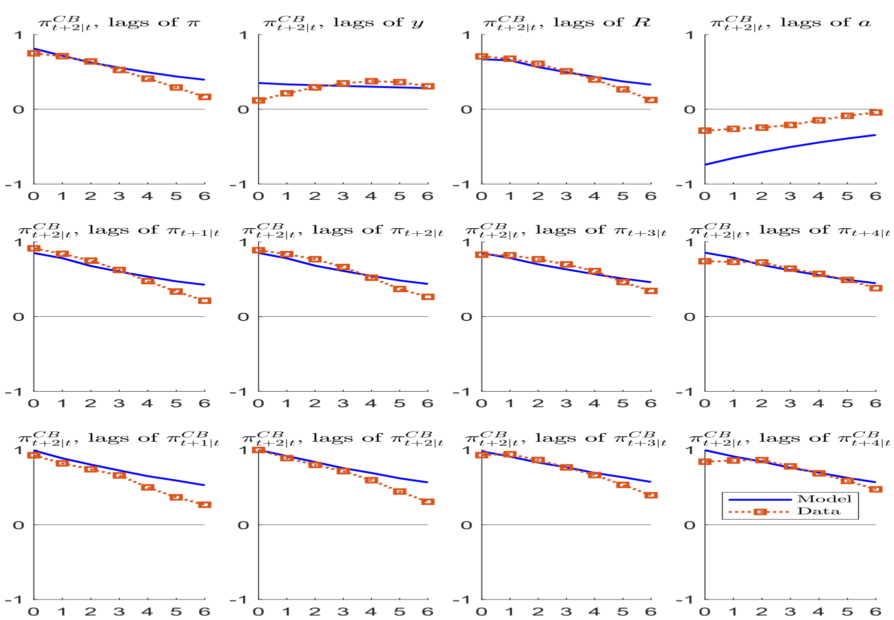


Figure 2.52: Correlations between two step ahead central bank inflation expectations and other variables, calibration 3.

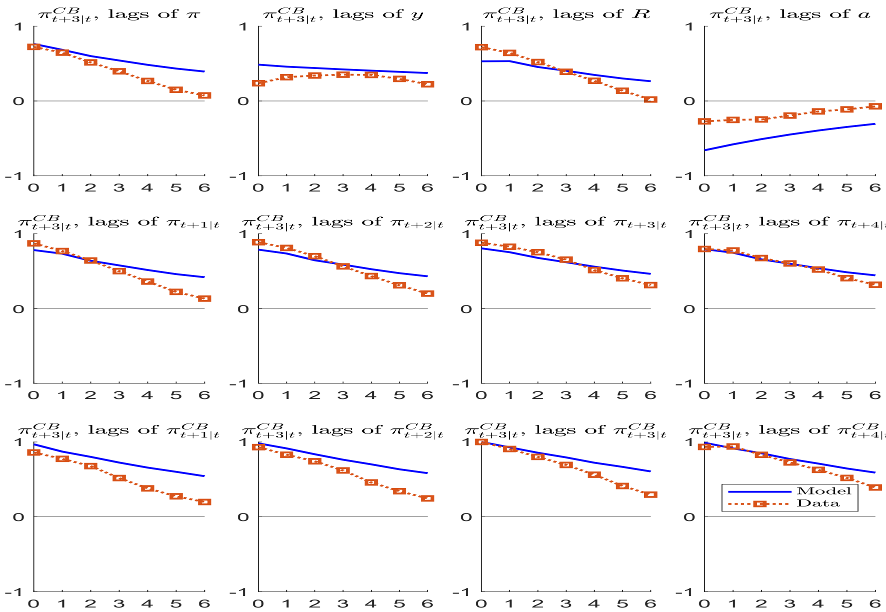


Figure 2.53: Correlations between three step ahead central bank inflation expectations and other variables, calibration 3.

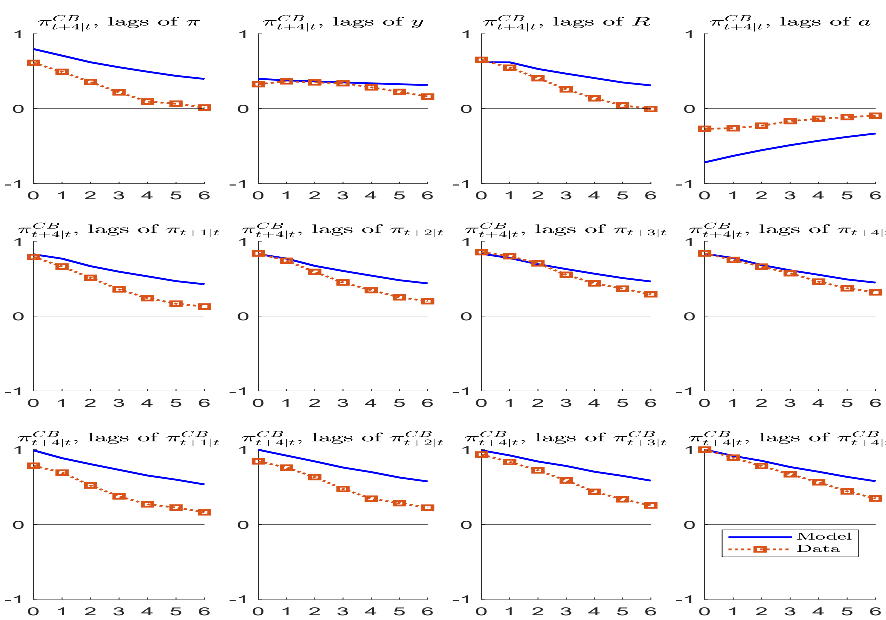


Figure 2.54: Correlations between four step ahead central bank inflation expectations and other variables, calibration 3.

Calibration 4: TFP shock

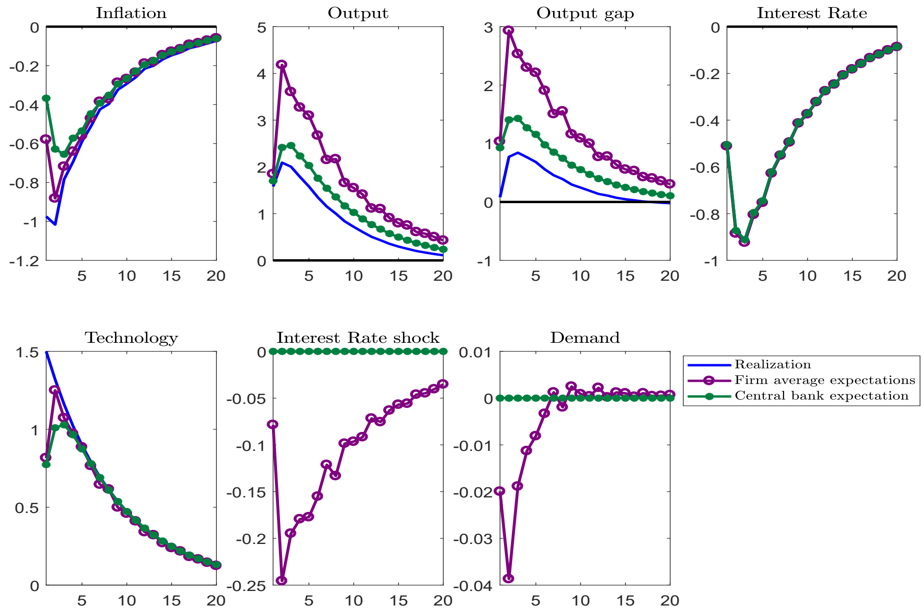


Figure 2.55: Impulse response to TFP shock, calibration 4.

2.D.4 Calibration 5

Impulse responses

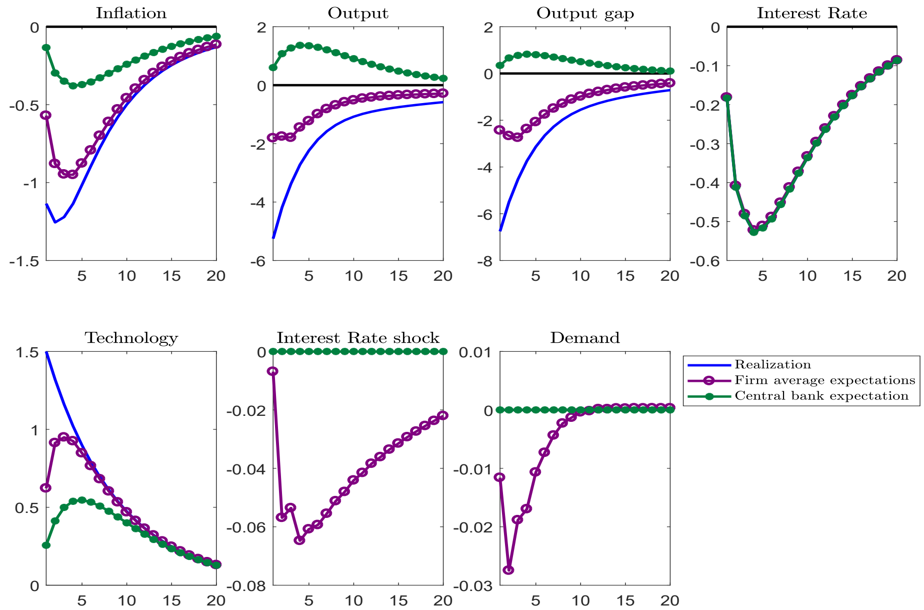


Figure 2.56: Impulse response to TFP shock, calibration 5.

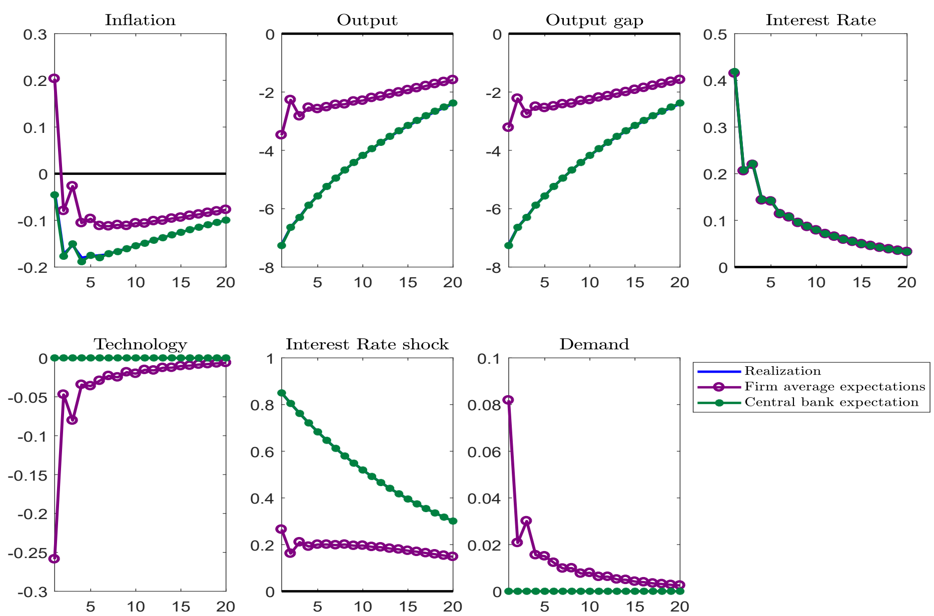


Figure 2.57: Impulse response to monetary policy shock, calibration 5.

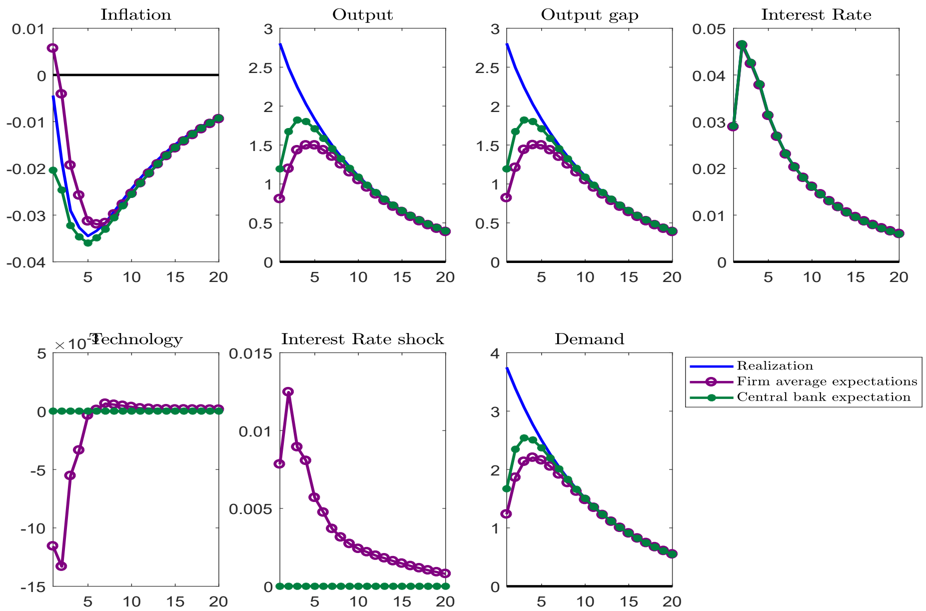


Figure 2.58: Impulse response to demand shock, calibration 5.

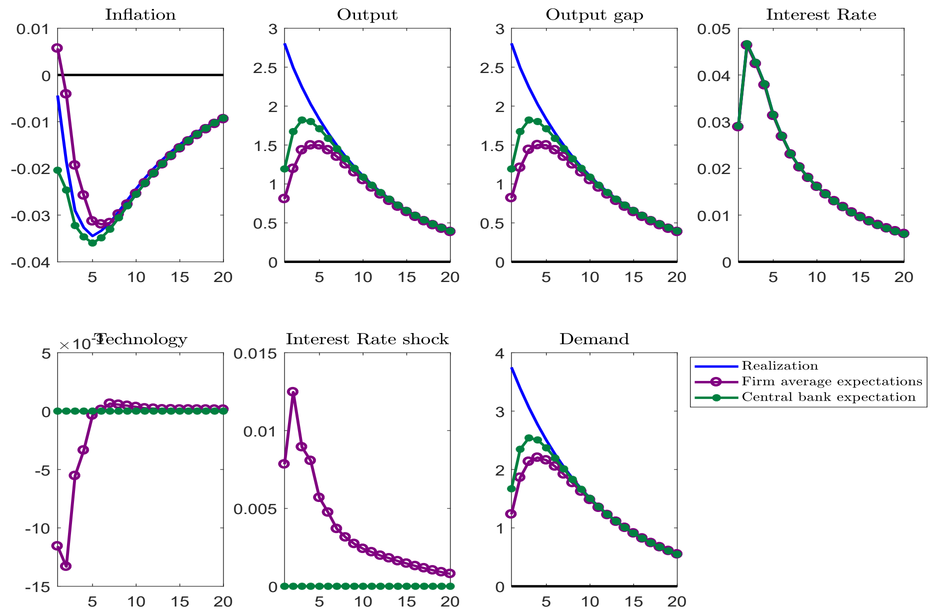


Figure 2.59: Impulse response to demand shock, calibration 5.

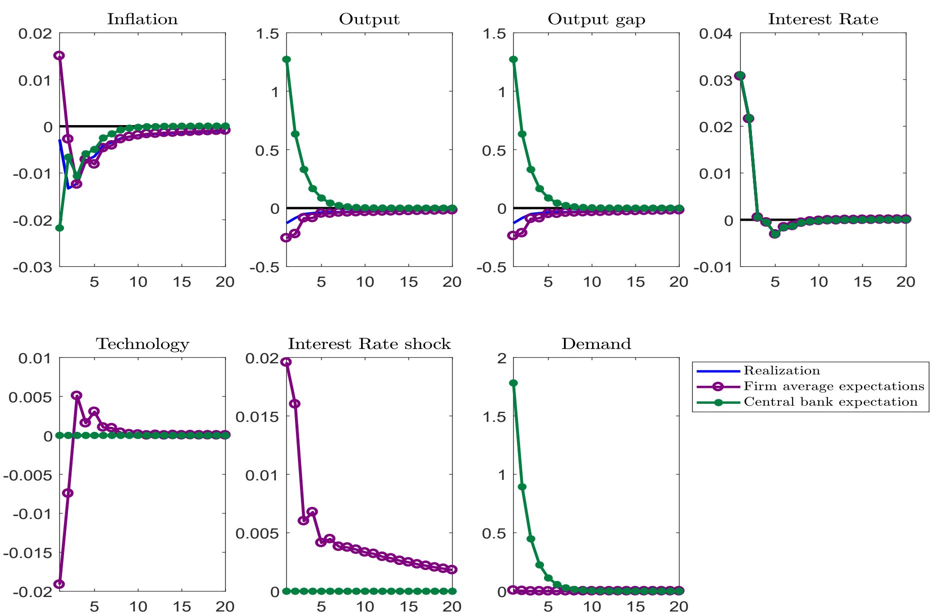


Figure 2.60: Impulse response to central bank demand measurement error shock shock, calibration 5.

Correlations

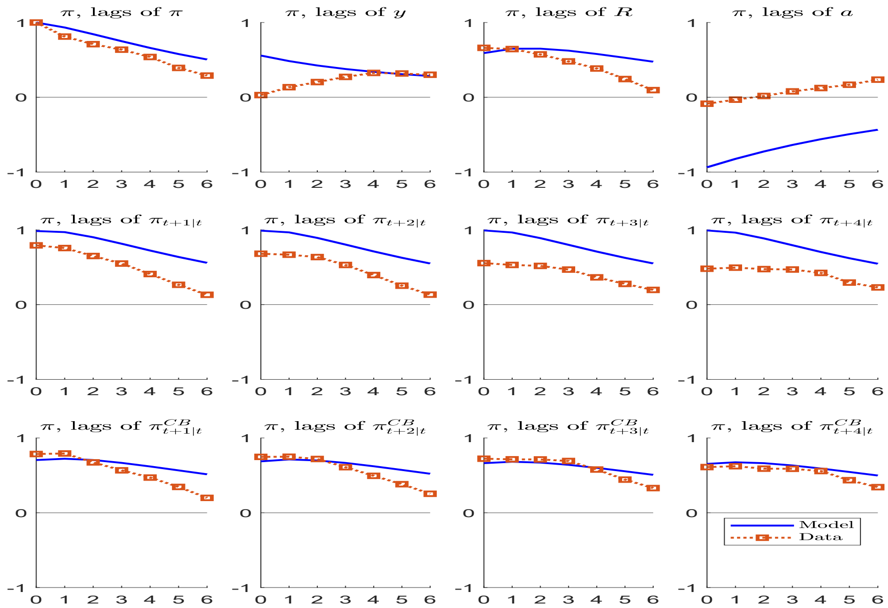


Figure 2.61: Correlation of inflation with other observables, calibration 5.

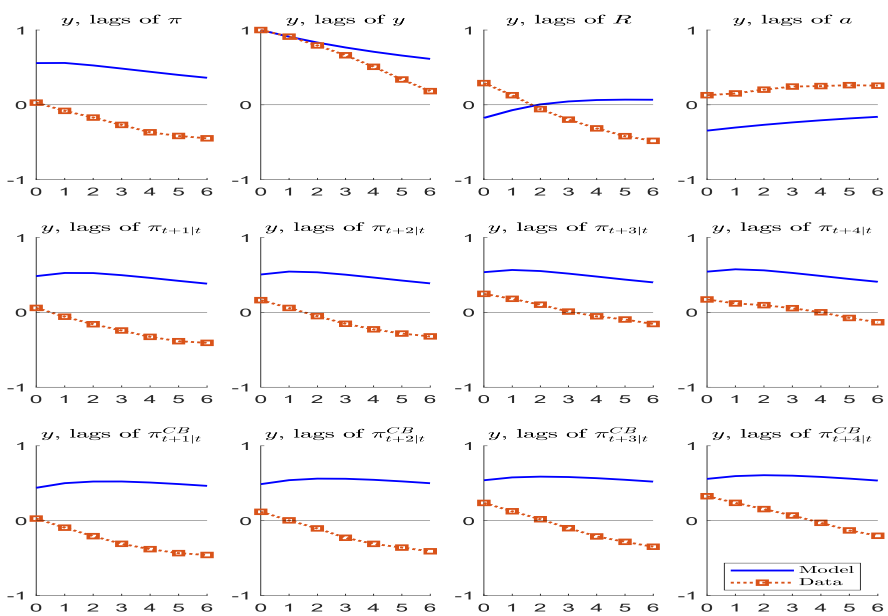


Figure 2.62: Correlation of output with other observables, calibration 5.

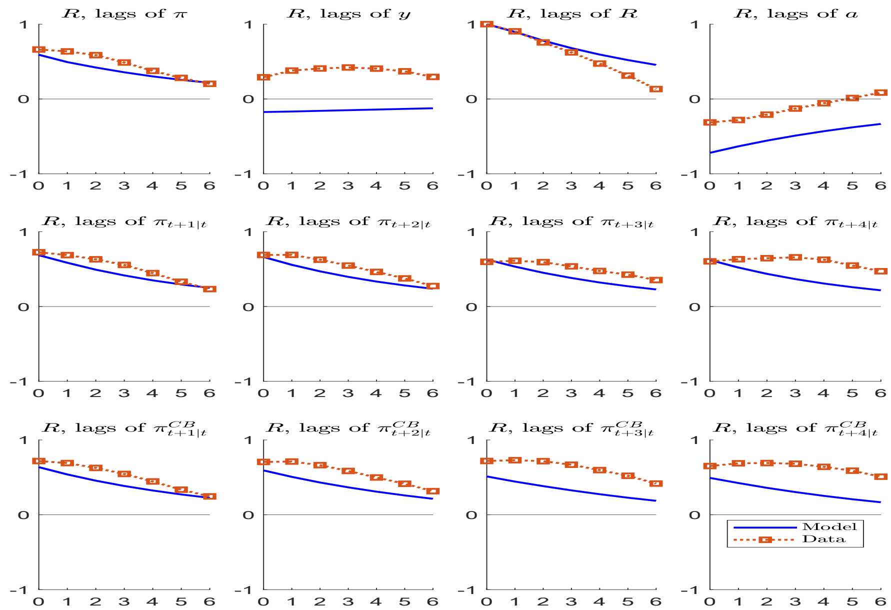


Figure 2.63: Correlations between interest rate and other variables, calibration 5.

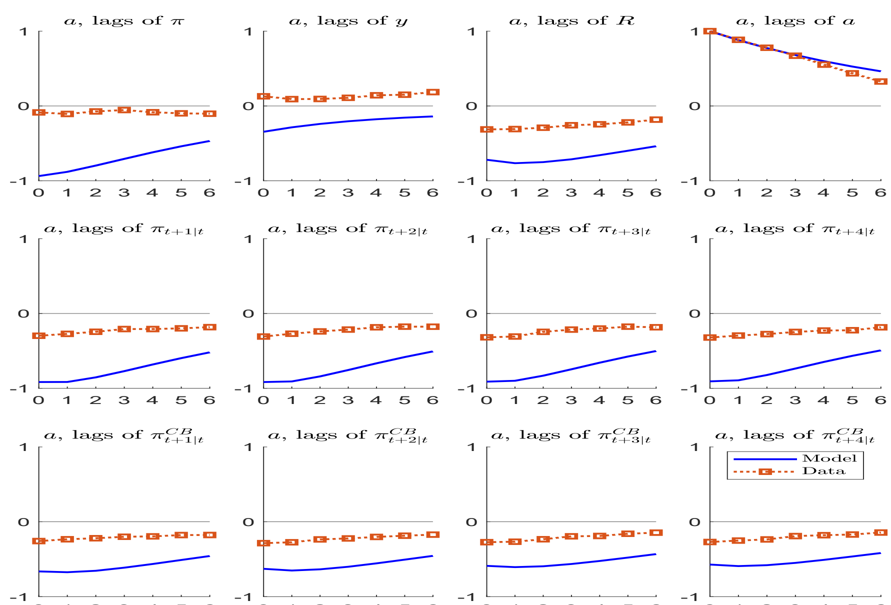


Figure 2.64: Correlations between TFP and other variables, calibration 5.

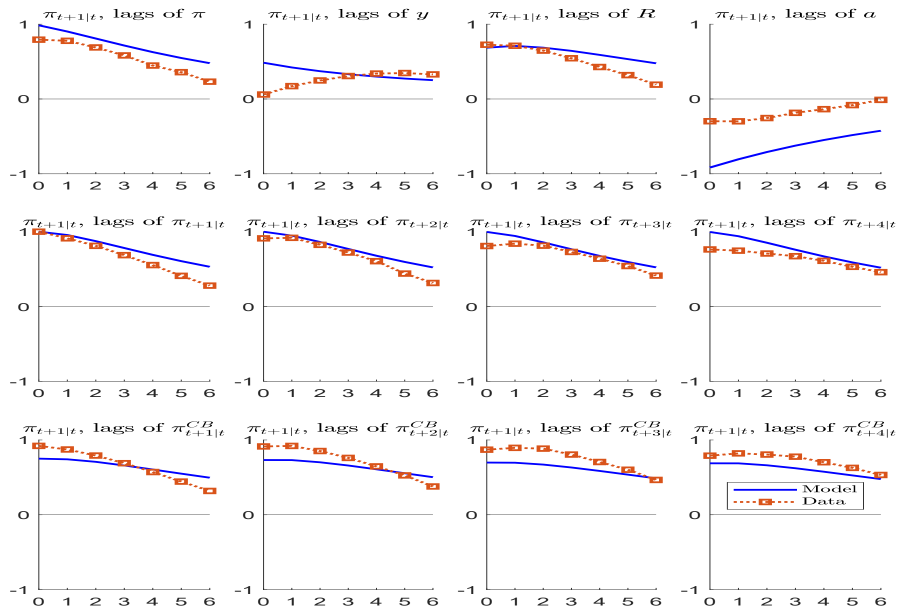


Figure 2.65: Correlations between one-step ahead private inflation expectations and other variables, calibration 5.

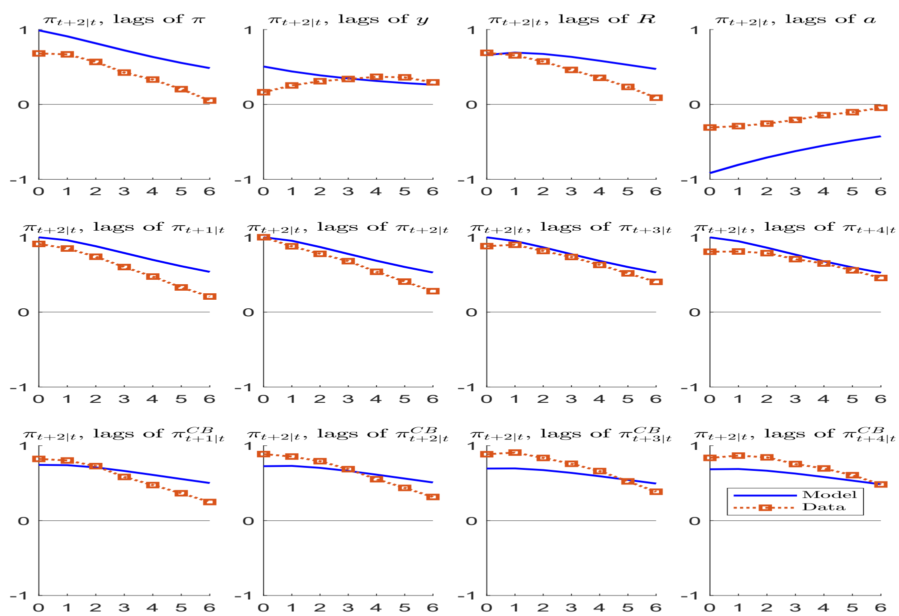


Figure 2.66: Correlations between two step ahead private inflation expectations and other variables, calibration 5.

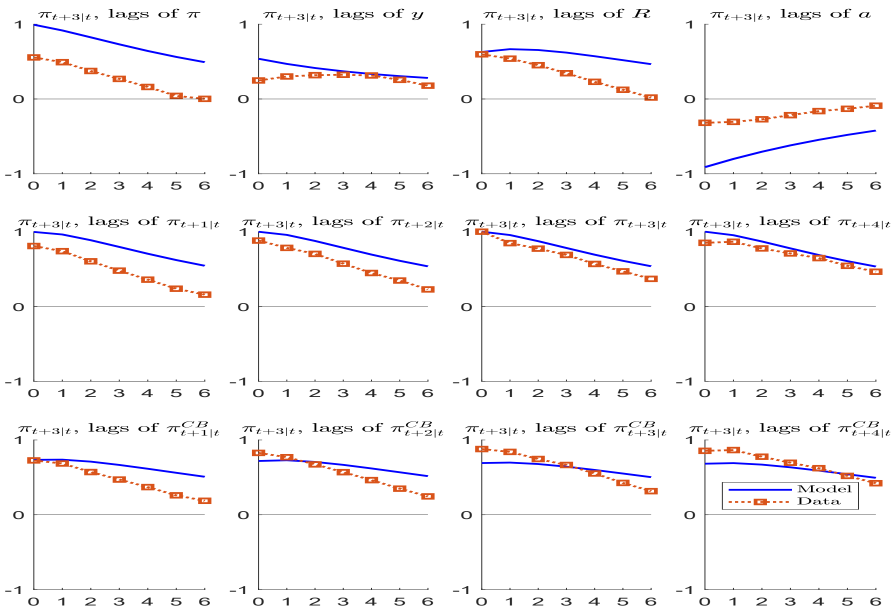


Figure 2.67: Correlations between three step ahead private inflation expectations and other variables, calibration 5.

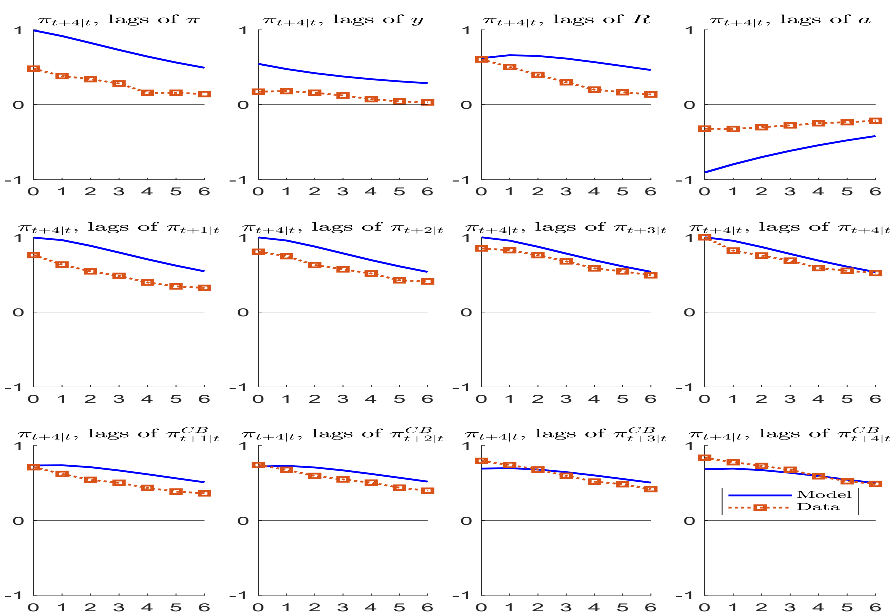


Figure 2.68: Correlations between four step ahead private inflation expectations and other variables, calibration 5.

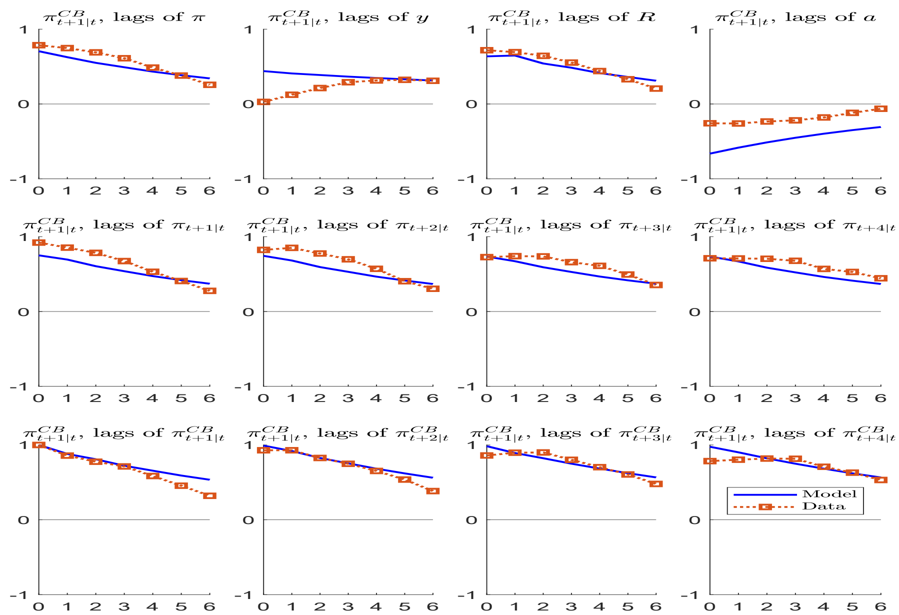


Figure 2.69: Correlations between one step ahead central bank inflation expectations and other variables, calibration 5.

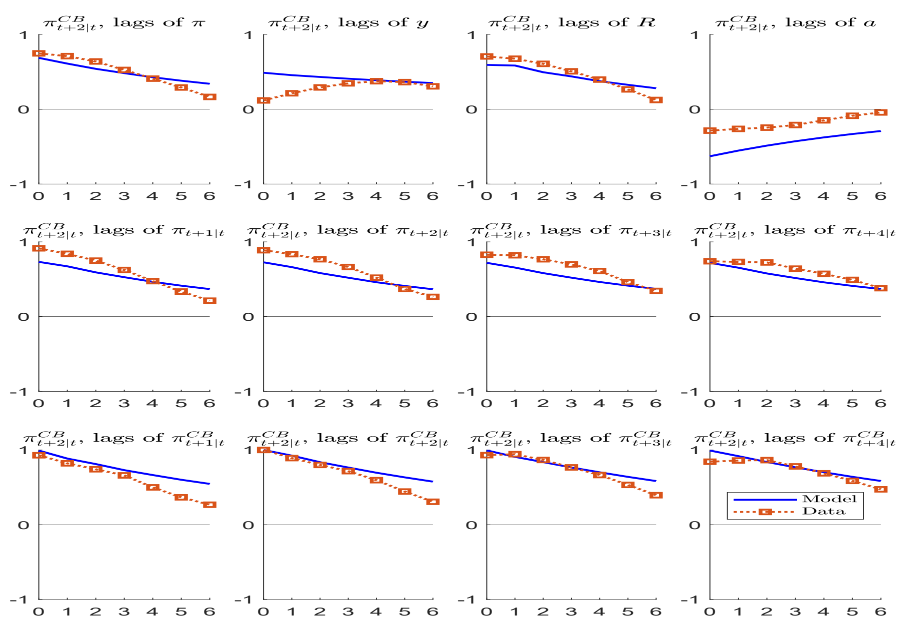


Figure 2.70: Correlations between two step ahead central bank inflation expectations and other variables, calibration 5.

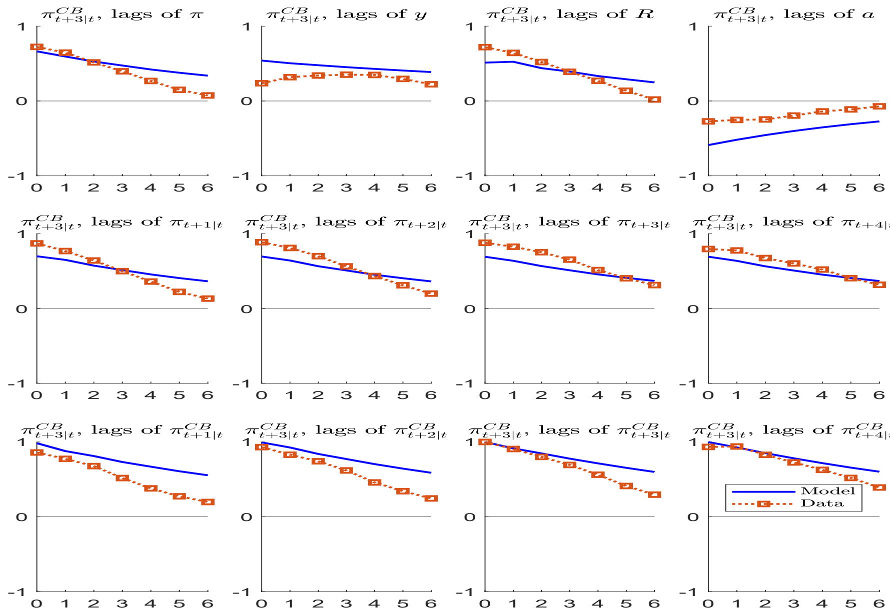


Figure 2.71: Correlations between three step ahead central bank inflation expectations and other variables, calibration 5.

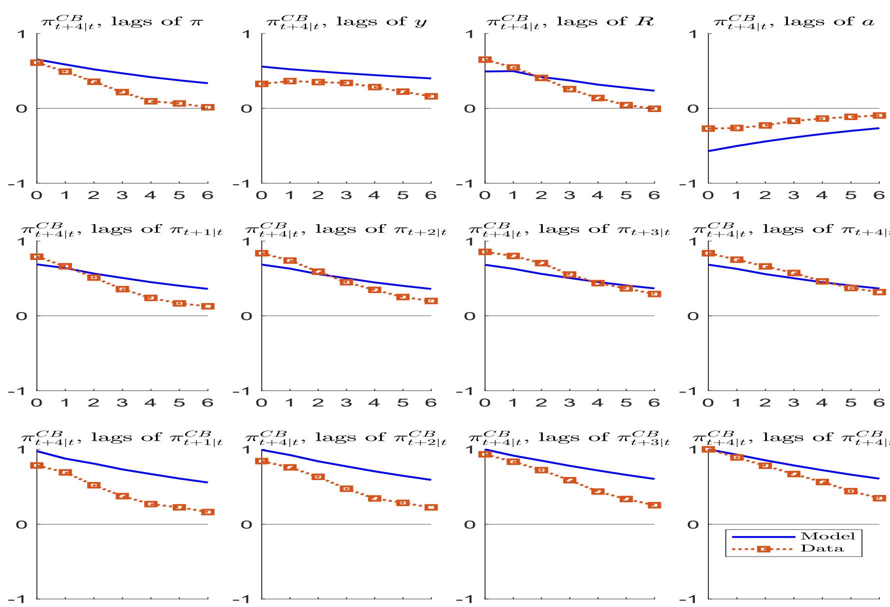


Figure 2.72: Correlations between four step ahead central bank inflation expectations and other variables, calibration 5.

2.E Melosi calibration figures

This appendix contains calibrated results for the model with a calibration similar to the estimated model in Melosi (2017), with very little noise in the central bank's observations of technology shocks ($\widetilde{\sigma_a}^{CB} \approx 2 \times 10^{-16}$) and a small amount of noise in their observations of demand shocks ($\widetilde{\sigma_g}^{CB} = 0.0033$). Higher levels of noise in technology shocks results in there being no solution to the fixed point problem. Here I present two sets of results: A baseline, and one where the persistence of technology shocks is calibrated to the same as the data used in the rest of the paper, and a third where in addition to the lower persistence of technology shocks the central bank strictly targes inflation in its policy rule.

2.E.1 Baseline Melosi calibration

Impulse responses

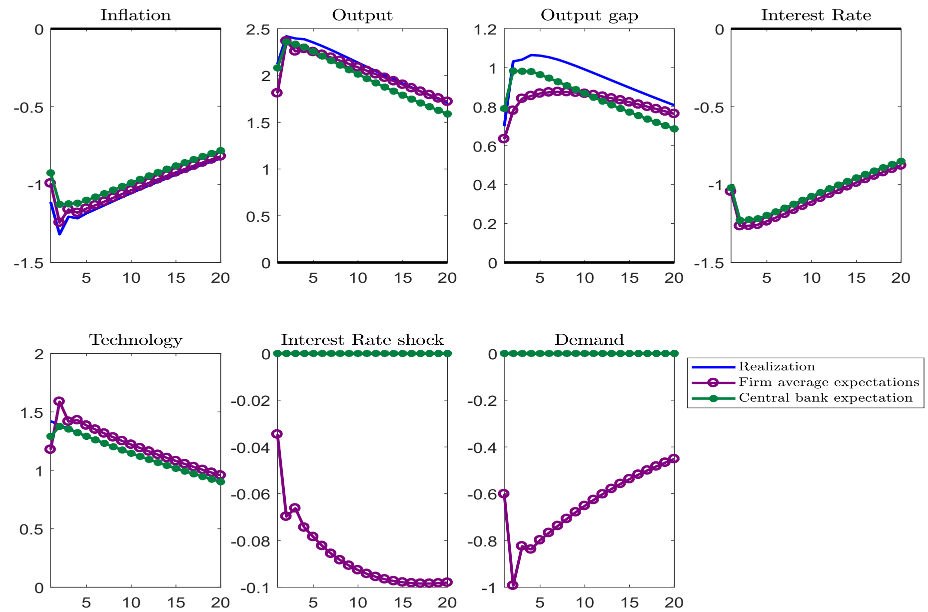


Figure 2.73: Impulse response to TFP shock, Melosi calibration.

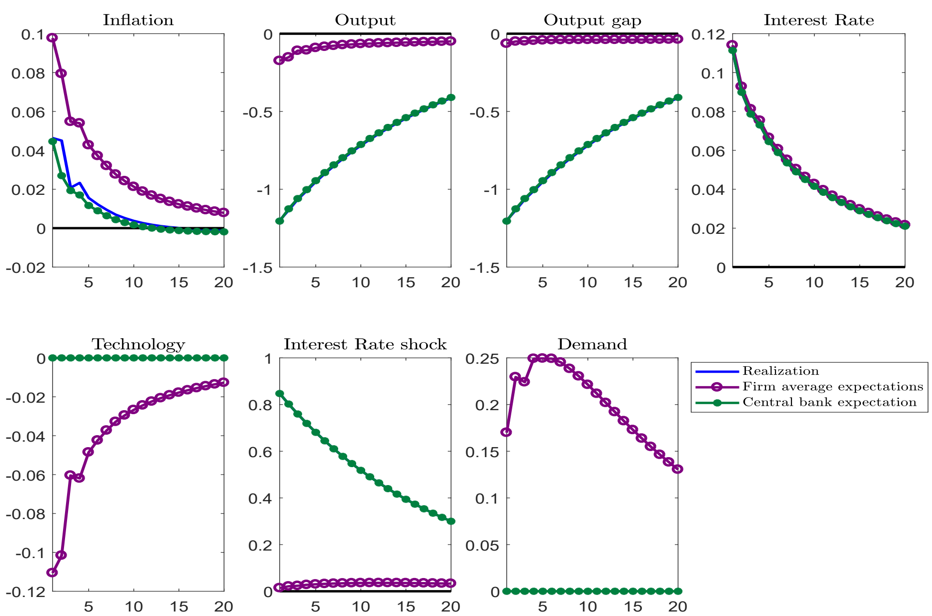


Figure 2.74: Impulse response to monetary policy shock, Melosi calibration.

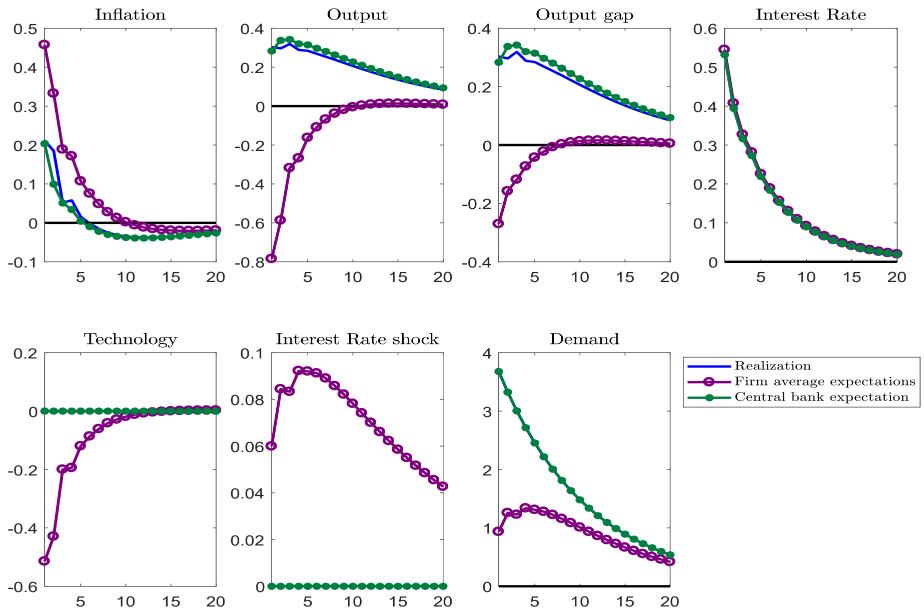


Figure 2.75: Impulse response to demand shock, Melosi calibration.

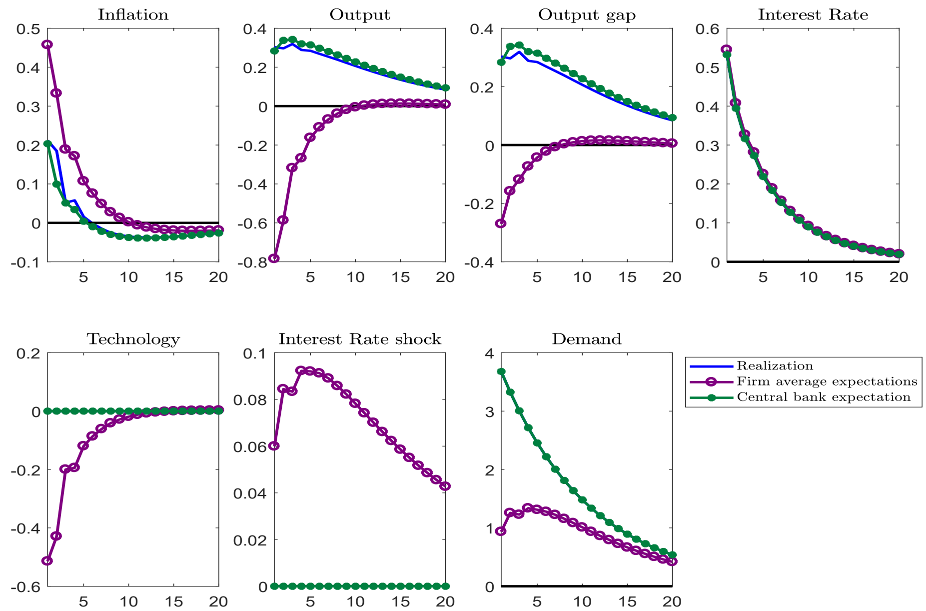


Figure 2.76: Impulse response to demand shock, Melosi calibration.

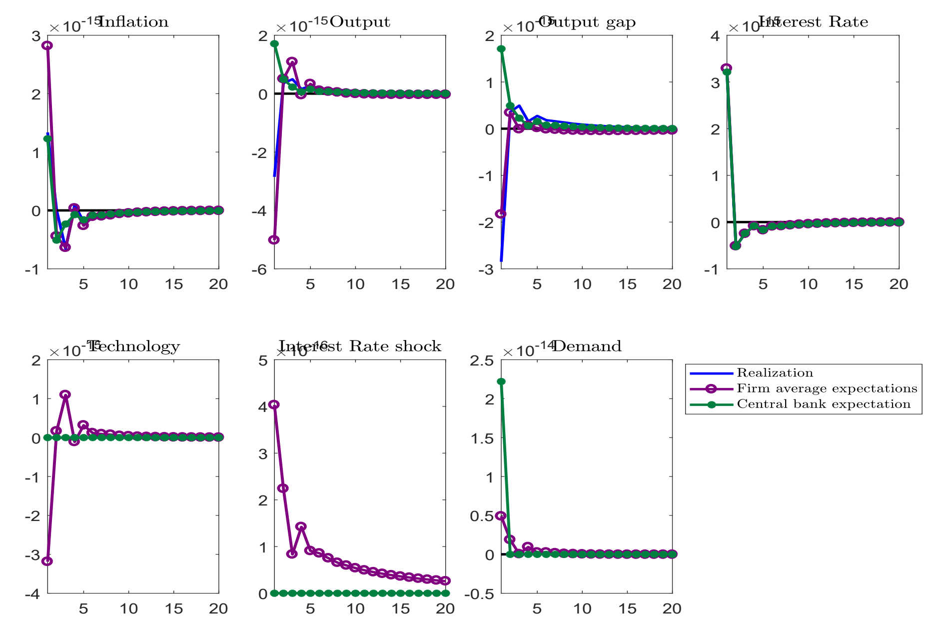


Figure 2.77: Impulse response to central bank demand measurement error shock shock, Melosi calibration.

Correlations

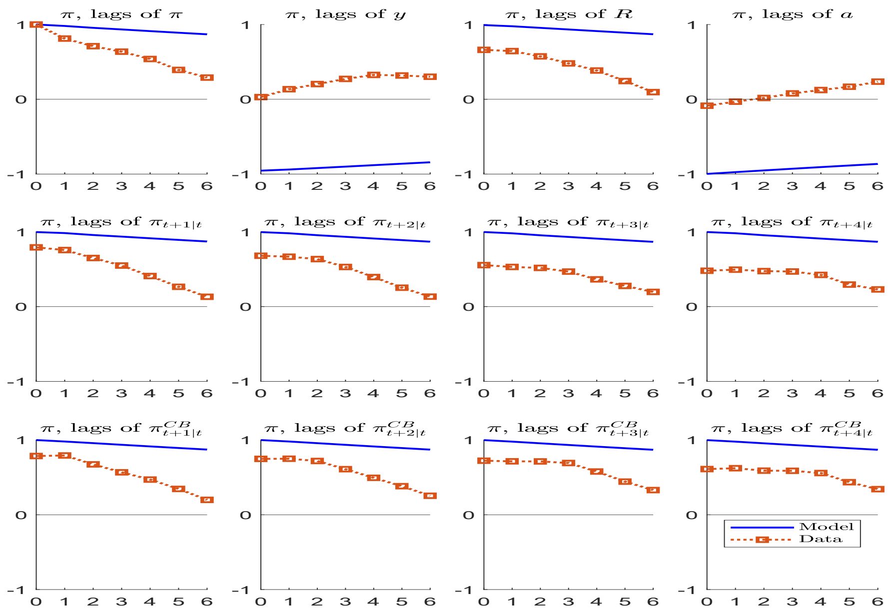


Figure 2.78: Correlation of inflation with other observables, Melosi calibration.

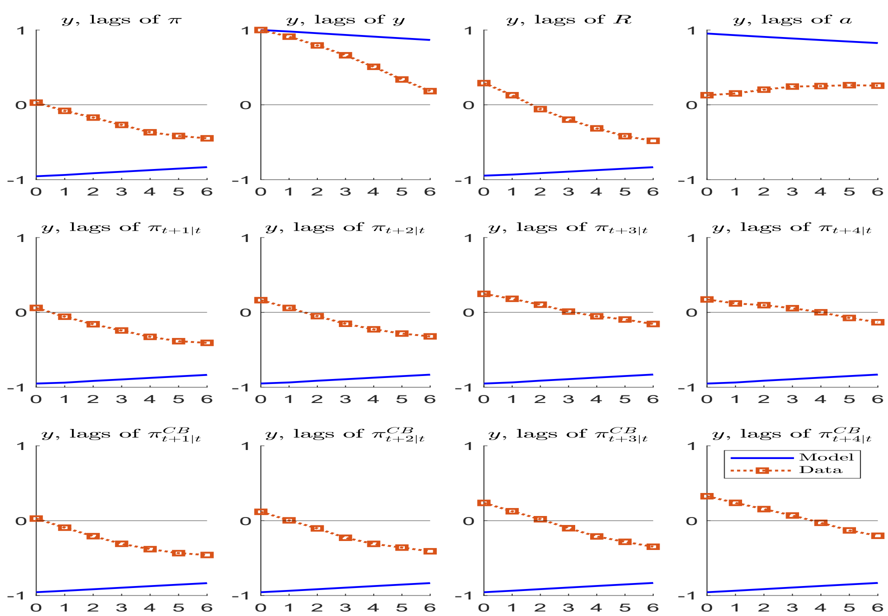


Figure 2.79: Correlation of output with other observables, Melosi calibration.

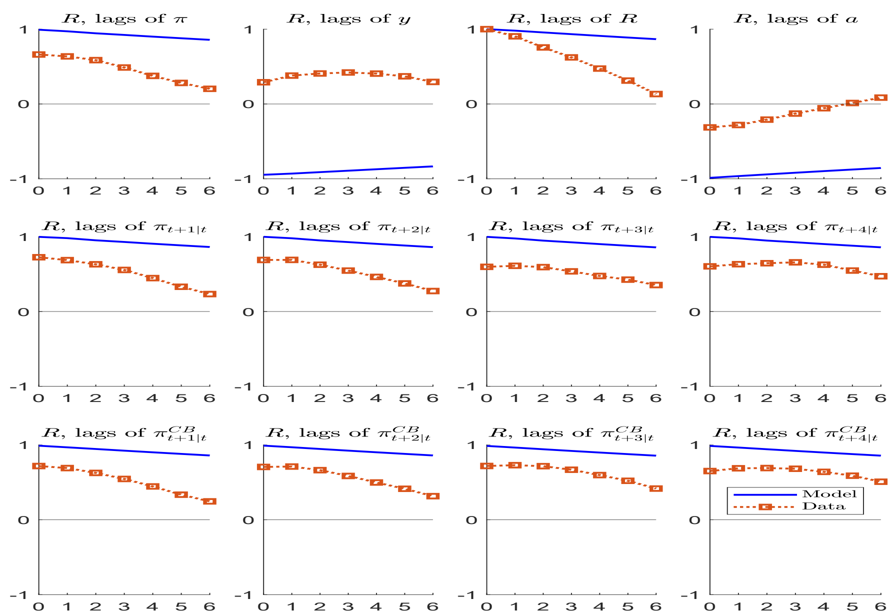


Figure 2.80: Correlations between interest rate and other variables, Melosi calibration.

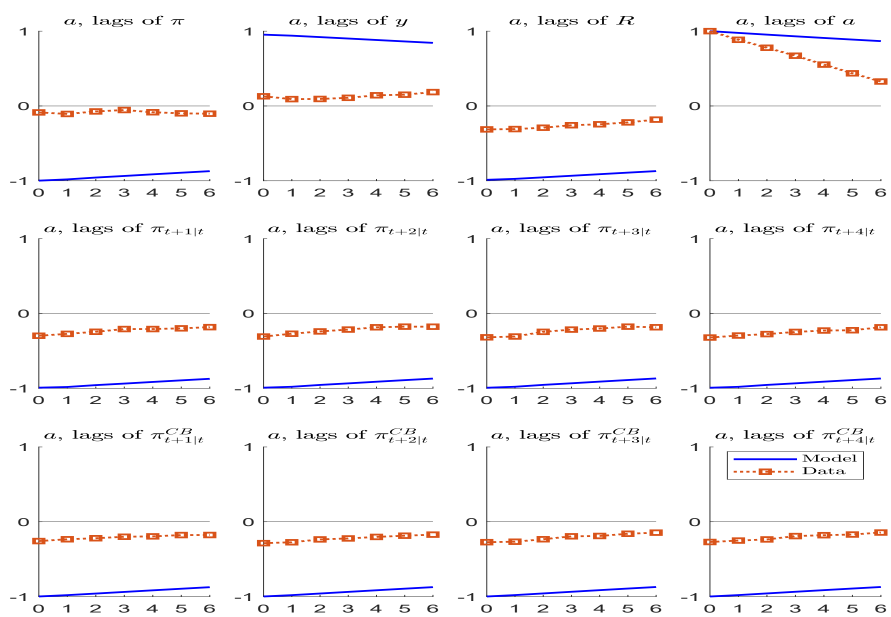


Figure 2.81: Correlations between TFP and other variables, Melosi calibration.

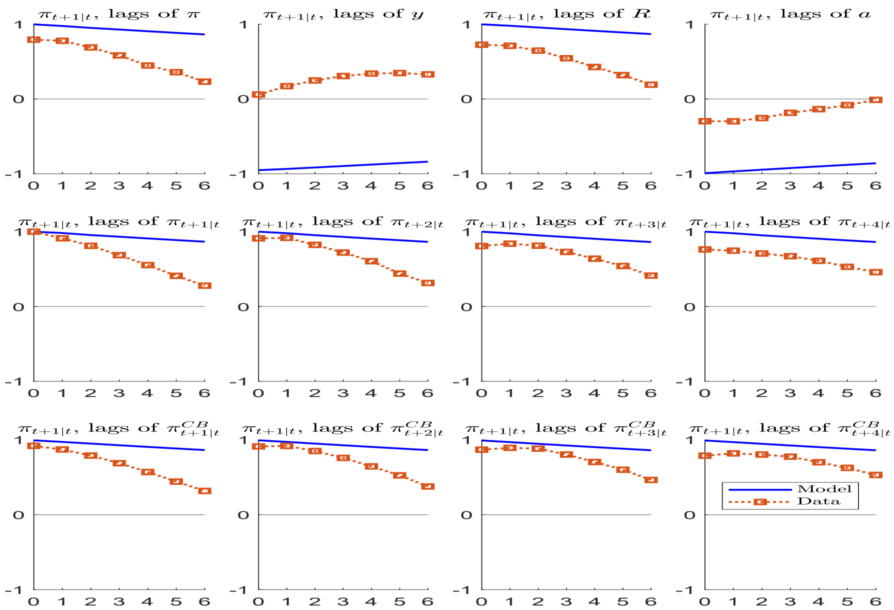


Figure 2.82: Correlations between one-step ahead private inflation expectations and other variables, Melosi calibration.

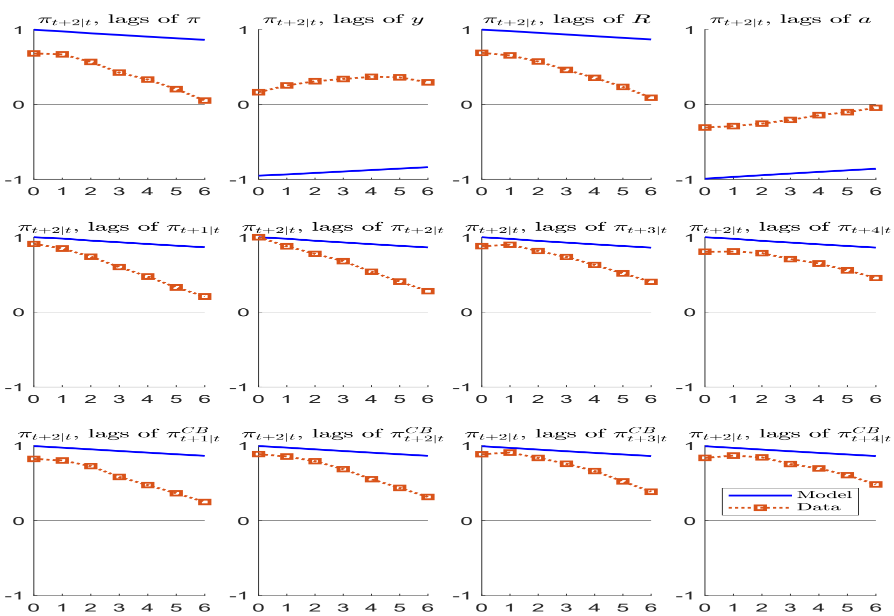


Figure 2.83: Correlations between two step ahead private inflation expectations and other variables, Melosi calibration.

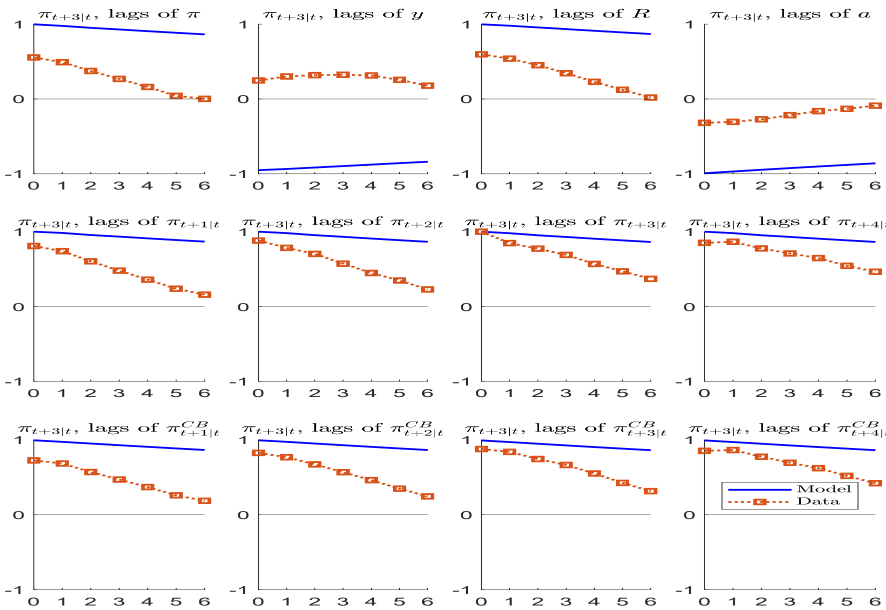


Figure 2.84: Correlations between three step ahead private inflation expectations and other variables, Melosi calibration.

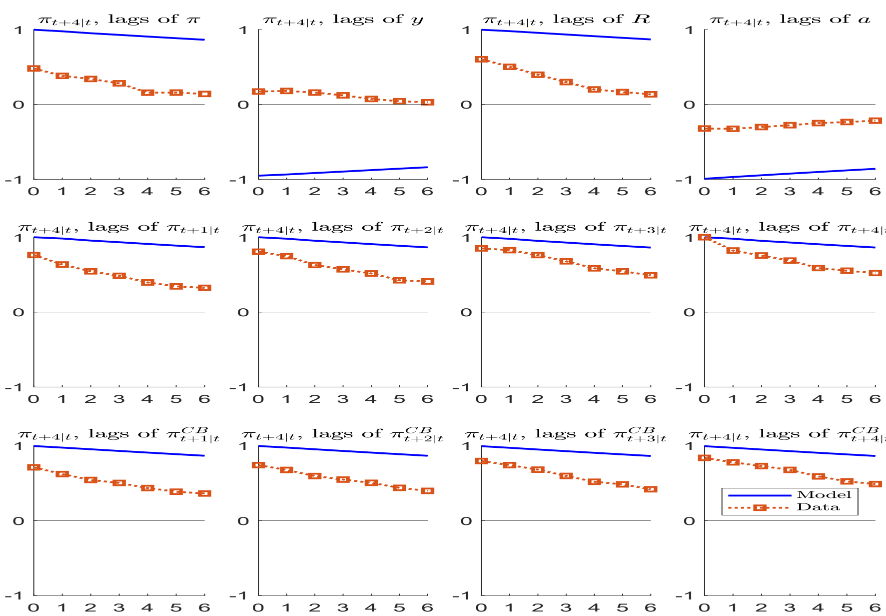


Figure 2.85: Correlations between four step ahead private inflation expectations and other variables, Melosi calibration.

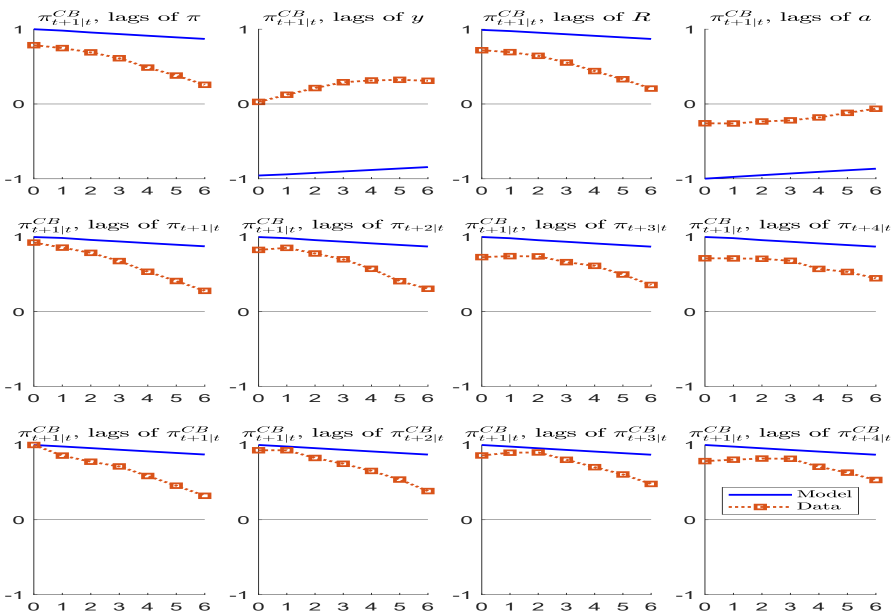


Figure 2.86: Correlations between one step ahead central bank inflation expectations and other variables, Melosi calibration.

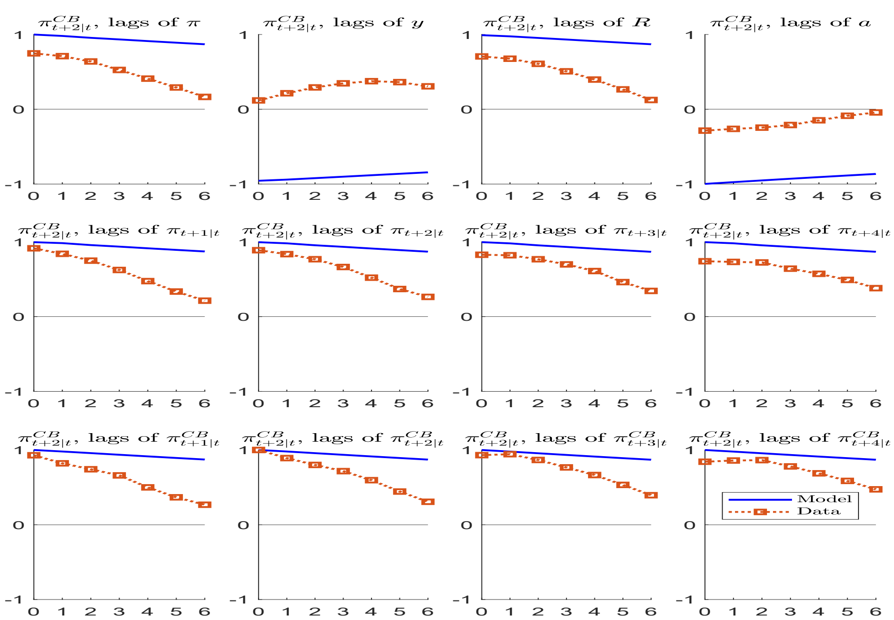


Figure 2.87: Correlations between two step ahead central bank inflation expectations and other variables, Melosi calibration.

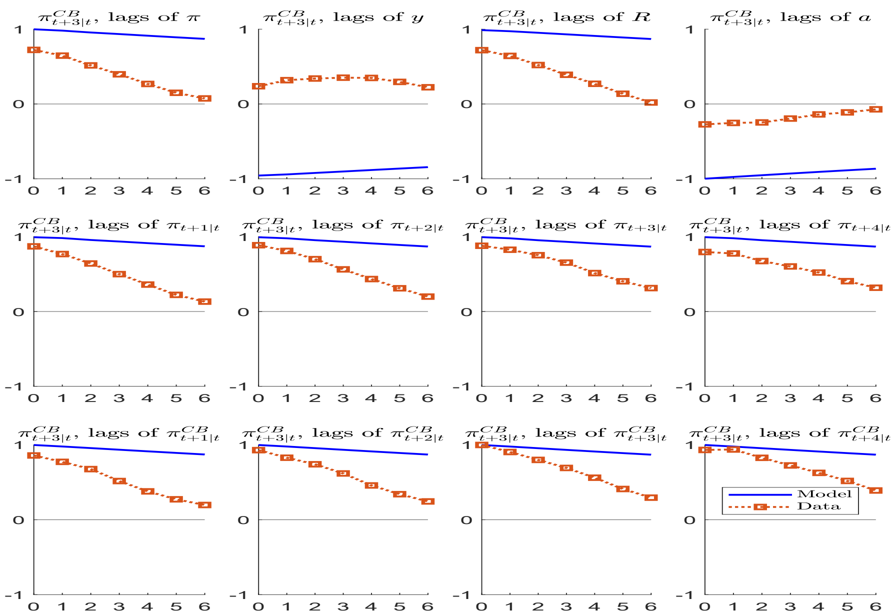


Figure 2.88: Correlations between three step ahead central bank inflation expectations and other variables, Melosi calibration.

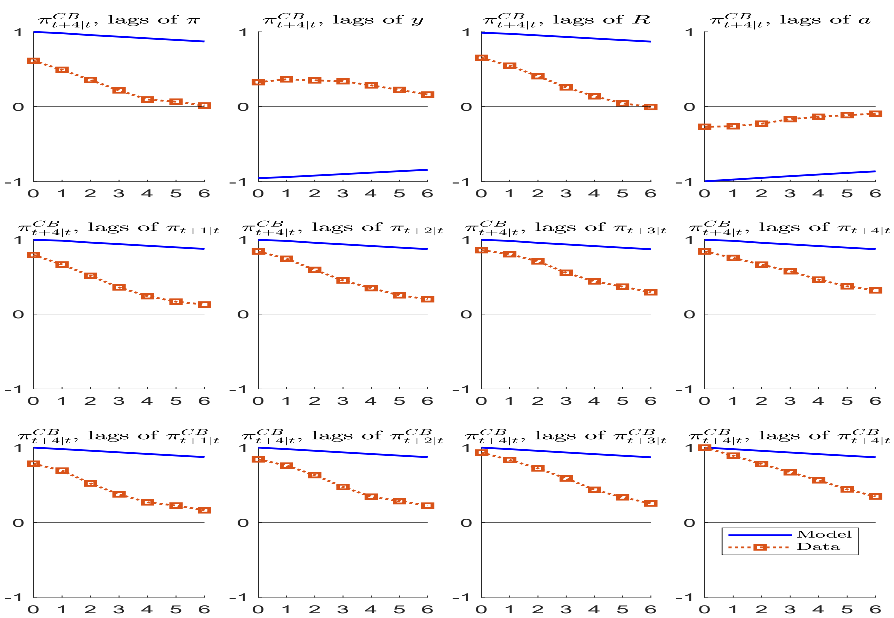


Figure 2.89: Correlations between four step ahead central bank inflation expectations and other variables, Melosi calibration.

2.E.2 Lower persistence of technology

Impulse responses

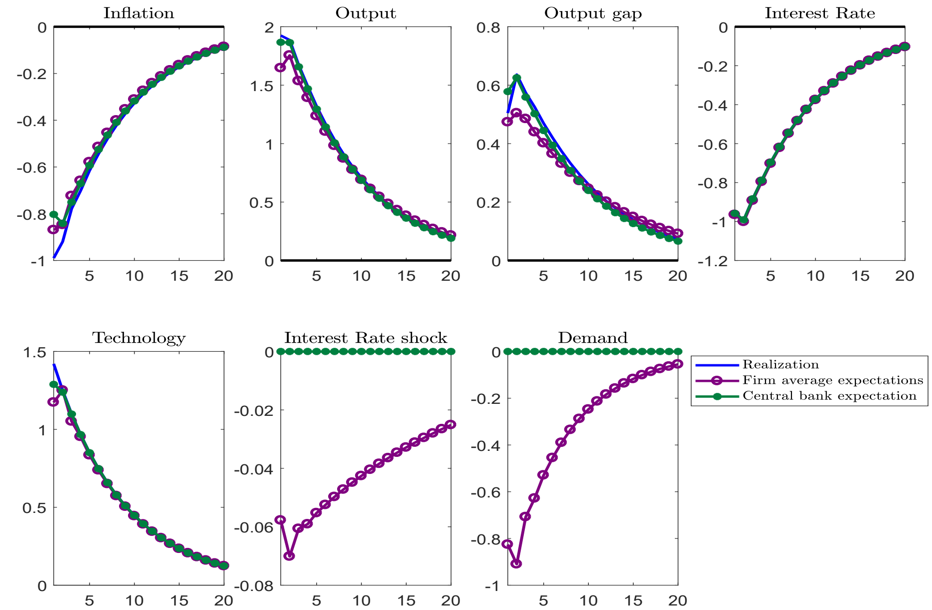


Figure 2.90: Impulse response to TFP shock, Melosi calibration with lower technology persistence.

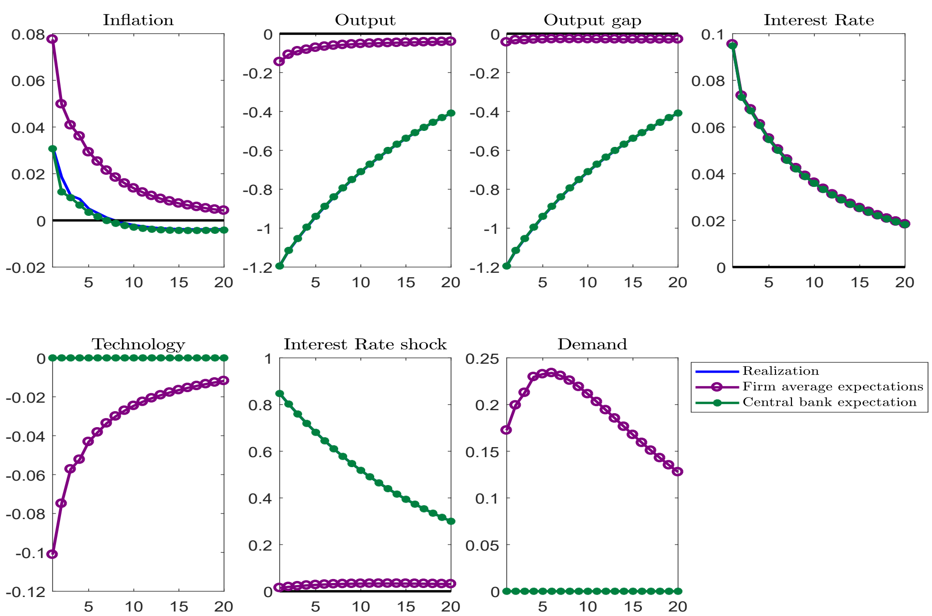


Figure 2.91: Impulse response to monetary policy shock, Melosi calibration with lower technology persistence.

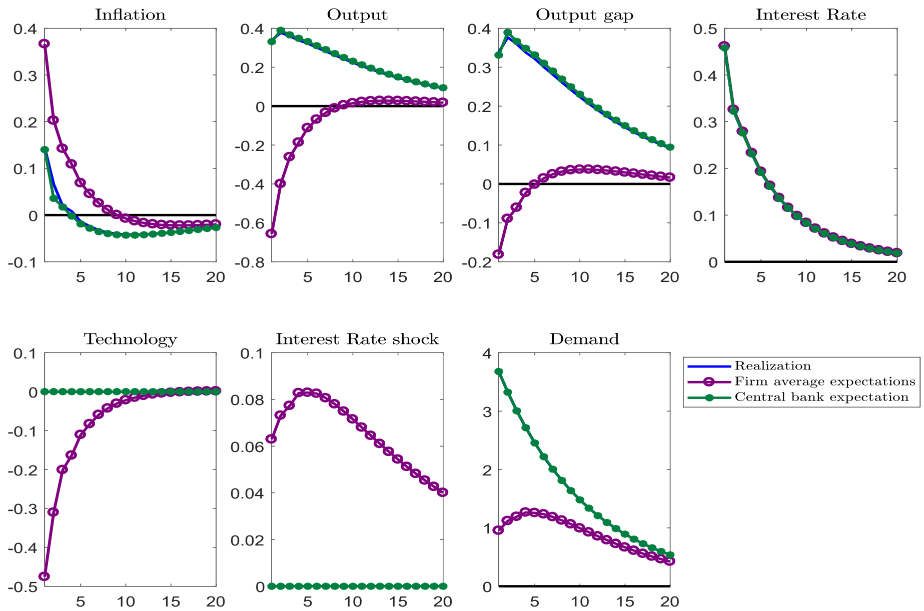


Figure 2.92: Impulse response to demand shock, Melosi calibration with lower technology persistence.

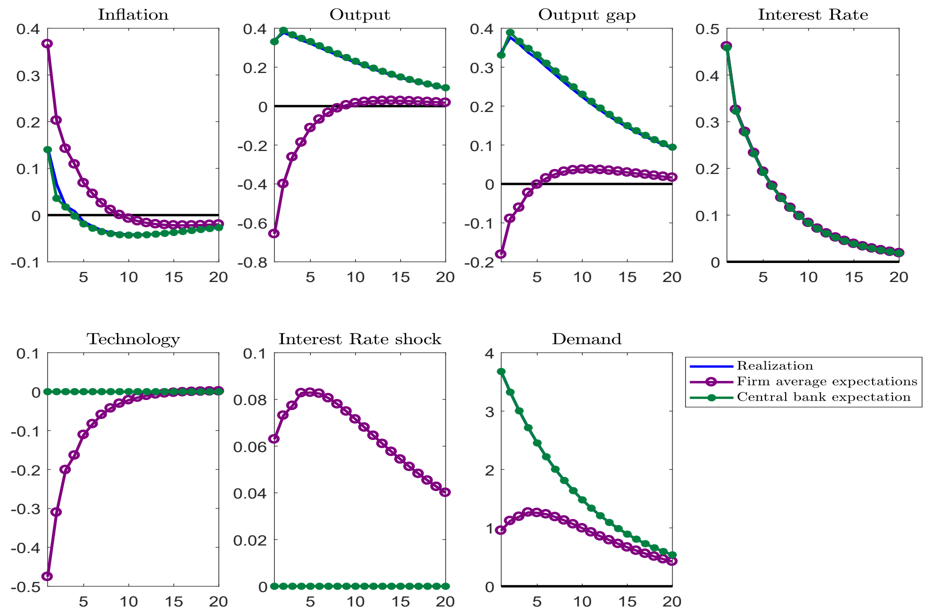


Figure 2.93: Impulse response to demand shock, Melosi calibration with lower technology persistence.

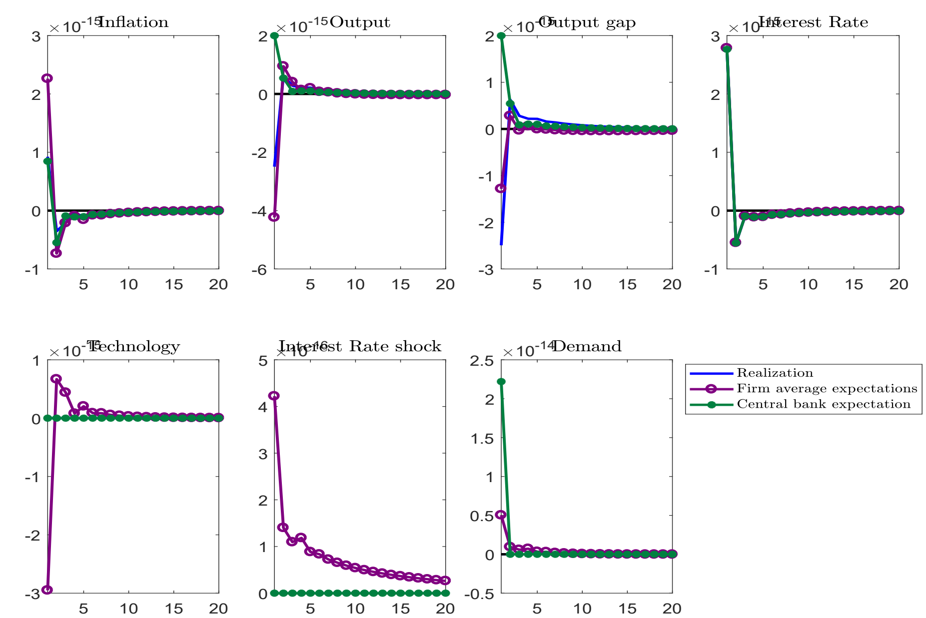


Figure 2.94: Impulse response to central bank demand measurement error shock shock, Melosi calibration with lower technology persistence.

Correlations

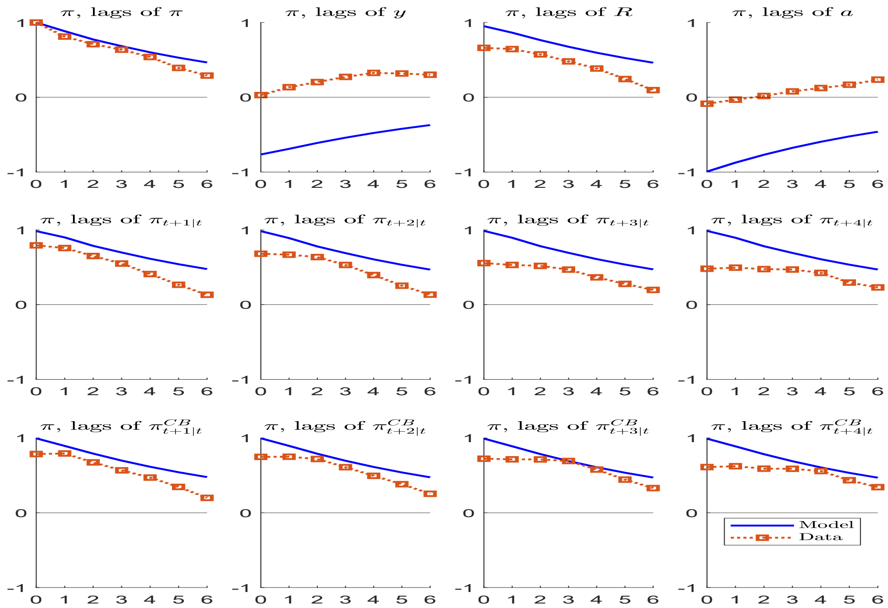


Figure 2.95: Correlation of inflation with other observables, Melosi calibration with lower technology persistence.

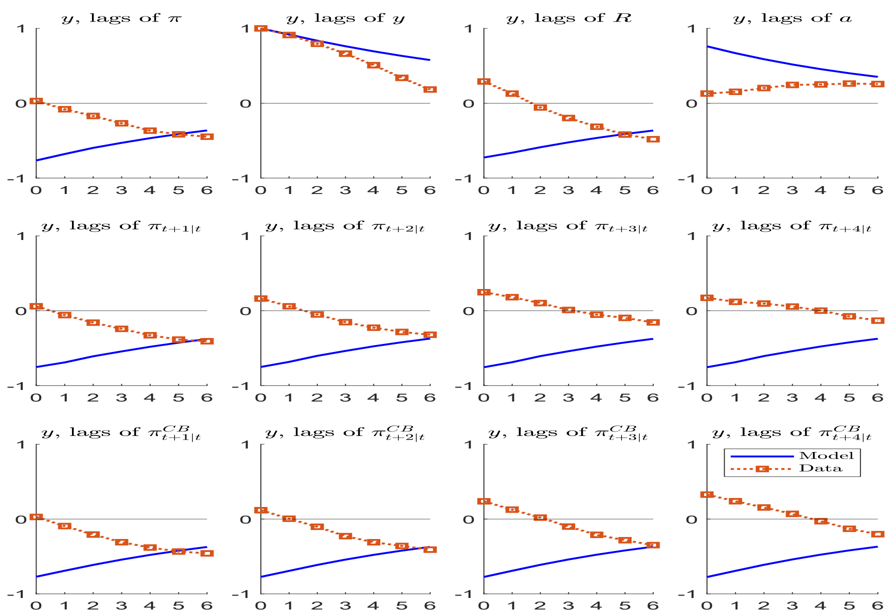


Figure 2.96: Correlation of output with other observables, Melosi calibration with lower technology persistence.

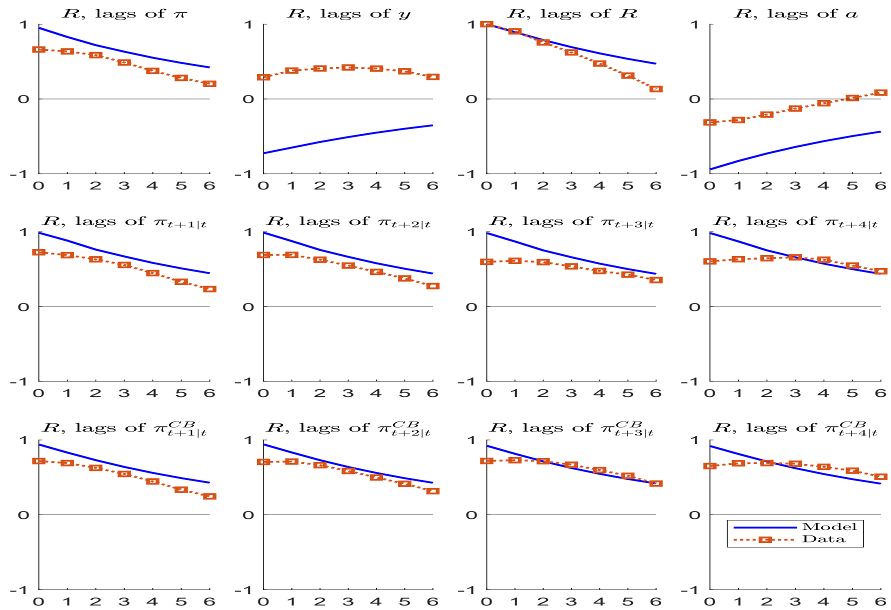


Figure 2.97: Correlations between interest rate and other variables, Melosi calibration with lower technology persistence.

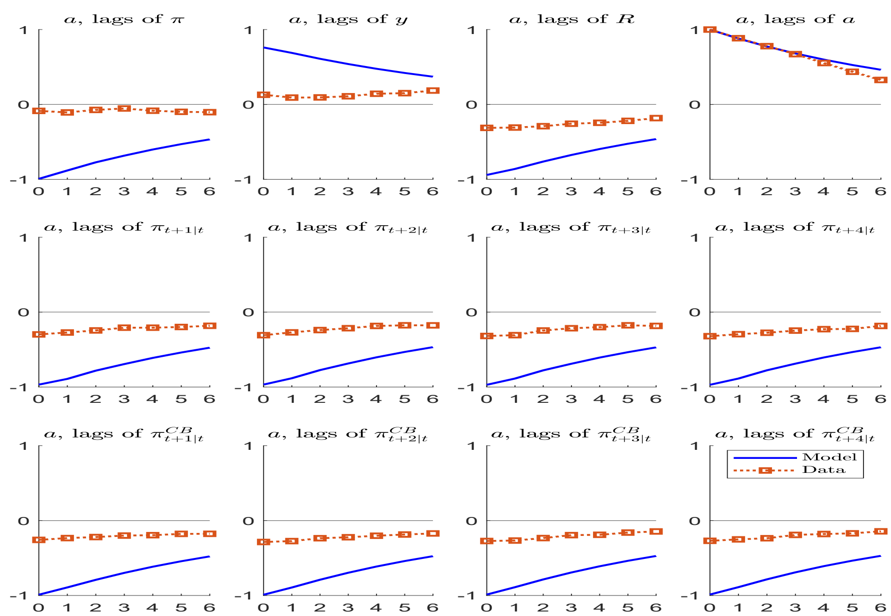


Figure 2.98: Correlations between TFP and other variables, Melosi calibration with lower technology persistence.

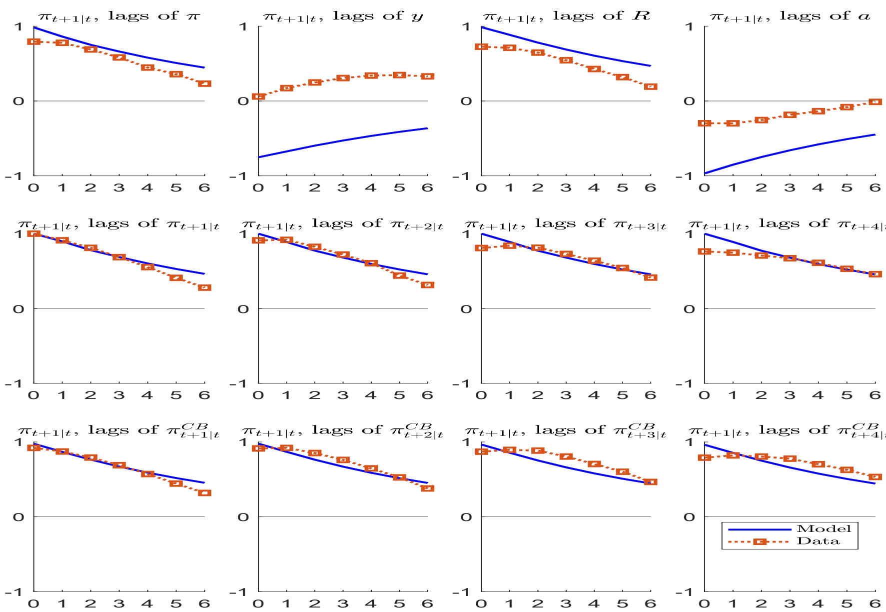


Figure 2.99: Correlations between one-step ahead private inflation expectations and other variables, Melosi calibration with lower technology persistence.

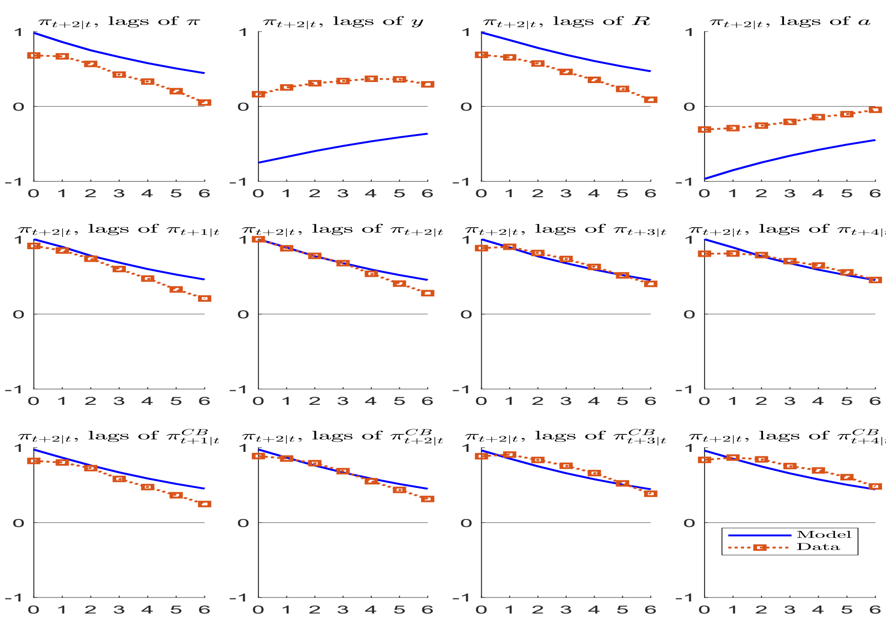


Figure 2.100: Correlations between two step ahead private inflation expectations and other variables, Melosi calibration with lower technology persistence.

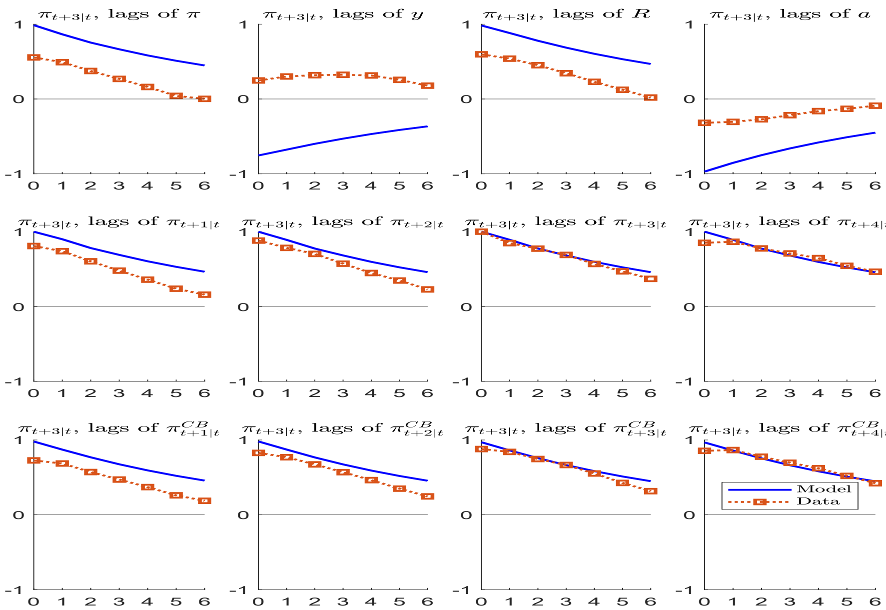


Figure 2.101: Correlations between three step ahead private inflation expectations and other variables, Melosi calibration with lower technology persistence.

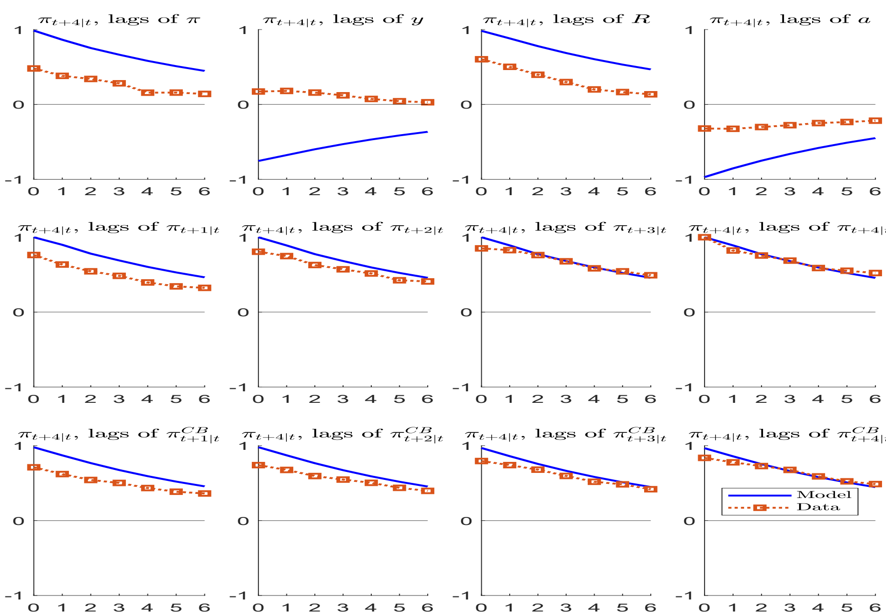


Figure 2.102: Correlations between four step ahead private inflation expectations and other variables, Melosi calibration with lower technology persistence.

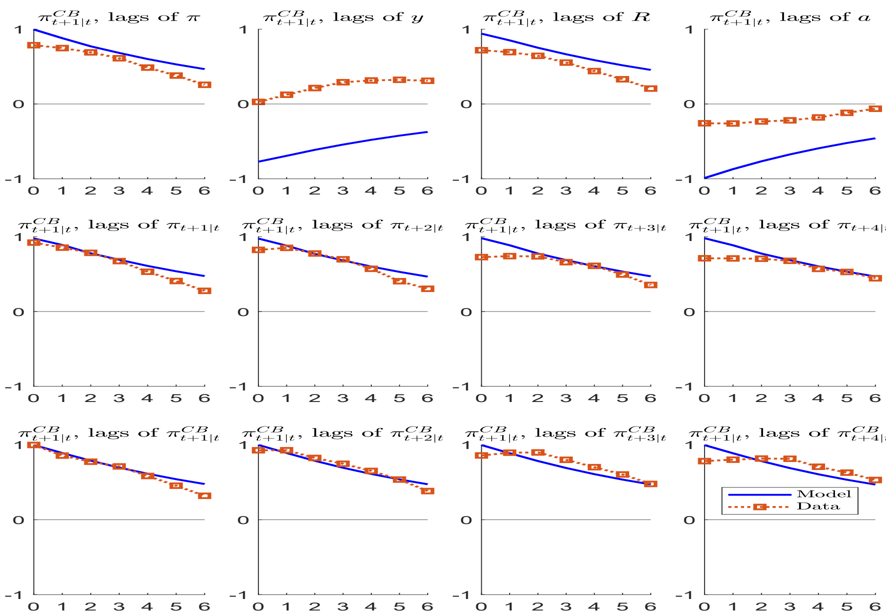


Figure 2.103: Correlations between one step ahead central bank inflation expectations and other variables, Melosi calibration with lower technology persistence.

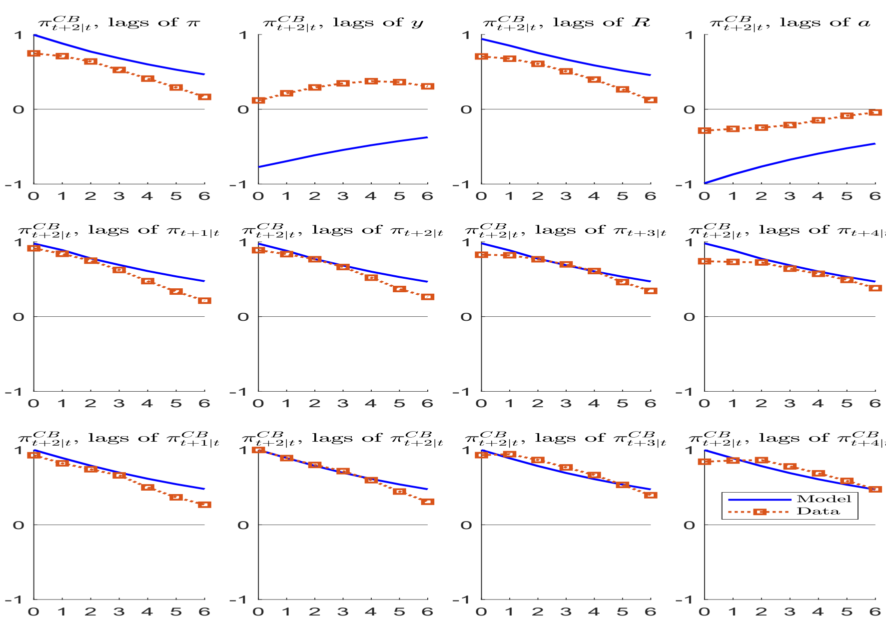


Figure 2.104: Correlations between two step ahead central bank inflation expectations and other variables, Melosi calibration with lower technology persistence.

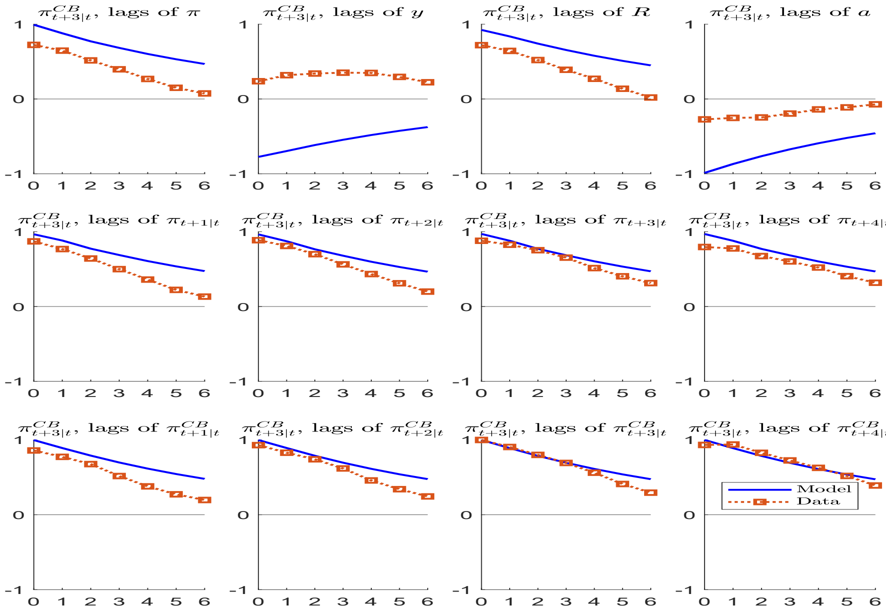


Figure 2.105: Correlations between three step ahead central bank inflation expectations and other variables, Melosi calibration with lower technology persistence.

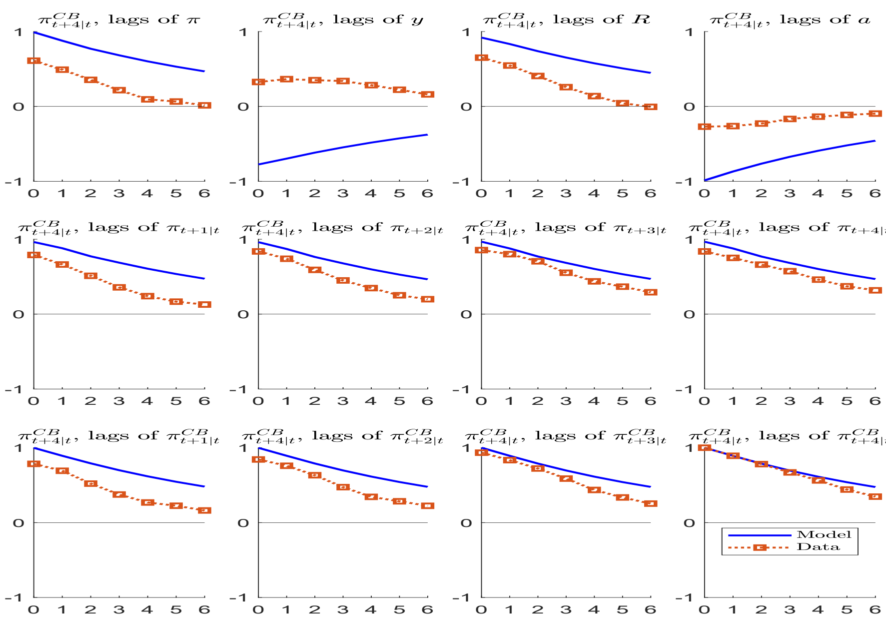


Figure 2.106: Correlations between four step ahead central bank inflation expectations and other variables, Melosi calibration with lower technology persistence.

Chapter 3

Information Investment in a Coordination Game

Abstract

Economic agents may face a trade off between the ability to accurately forecast a signal and having beliefs correlated with other agents. I develop a two stage model where prior to choosing actions in a coordination game, non-atomistic agents decide how to allocate attention across public and private signals. Sufficient conditions for a unique symmetric switching equilibrium in the coordination game are insufficient to support the existence of symmetric information equilibria. Numerical analysis suggests (1) when symmetric equilibria do exist, they are unique (2) Changes to committee size or information quality have small effects on committee welfare but large effects on the accuracy of beliefs

3.1 Introduction

When making binary economic decisions in uncertain environments, agents often have two sets of incentives. One is the incentive to take the correct action given the underlying state of the world. A second is a strategic complementarity motive. For example, members of a committee may be indifferent about a proposition but choose to vote for or against it based on what the majority of the other members seem to want to do, because they do not want to "stick their necks out" or because they value consensus. If they have strong views that differ from the other committee members, however, they may choose to dissent in spite of the coordination motive.

In this environment, the information agents use is important. If agents are relying on the same signals to inform their beliefs, then they will likely act similarly. If they instead rely on idiosyncratic information, there may be greater variation in beliefs and (perhaps) of actions. Often, committees feature information revelation (like debate or expert testimony) before actions are taken. There may be an incentive to invest in research to help guide the committee's decision, even setting aside other strategic motives arising from differences in preferences. On the other hand, understanding information presented by others is costly in terms of time and effort, and data presented to others may be interpreted in idiosyncratic ways. Depending on the trade off perceived by agents between being correct and doing the same thing as others, they may find that paying attention to public information is less worthwhile than improving ones own forecasts or vice versa.

As a concrete example, members of the Federal Open Market Committee (FOMC) have to decide whether to maintain the status quo or adjust policy. The Federal Reserve Board staff prepare an analysis of economic conditions that is circulated prior to the meeting. However, the regional Federal Reserve banks also prepare separate forecasts and analyses of current conditions. Some of these are required and circulated widely, like the Beige Book (the summary of Federal Reserve District conditions). Different Federal Reserve Banks produce statistics about conditions in financial and commodities markets, manufacturing and commerical activity, and also publish model-based forecasts and research briefs on economic trends. However, the Federal Reserve banks do not generally release all their information to the public, and FOMC members present summaries of their views at meetings. These summaries may not reflect all of the information used in forming beliefs. For example, Reserve Bank presidents may request briefings on topics not included in projections reported to the rest of the FOMC. Committee members must decide what indicators to pay attention to when forming their beliefs. Even given a common set of data, they may reach different conclusions based on their readings of the available information. Indeed, FOMC dissents often reflect differences in opinion about what the current stance of monetary policy is and whether it is appropriate given an agreed-upon set of goals (Thornton and Wheelock (2014)).

FOMC members also appear to care about coordinating their actions. For example, writing in the Wall Street Journal, former Federal Reserve Bank of Philadelphia President Charles Plosser noted that the September 2015 FOMC vote to maintain a policy rate at zero had only one hawkish dissent despite three voting members indicating they were ready to increase rates. He argued that the vote might have been lopsided because of

> [...] the desire, inside and outside the Fed, for consensus decision-making. Markets don't like dissent or indecision, goes the argument. Some say that the presidents of the 12 Reserve Banks should quit talking to reporters and giving speeches, or at least refrain from questioning actions taken by the FOMC or the Fed chairman. By the same token, any opposition from the members of the Board of Governors might be seen as signaling a lack of confidence in the chairman. (Plosser (2015)).

Hence, understanding how policymakers form expectations and their desire is thus important for understanding how policy is determined.

In this paper, I develop a stylized model that reflects this intuition about the way members of a committee with dispersed information decide on binary actions. The model consists of a two stage game. In the first stage, agents ("committee members") decide how much attention to pay to a set of signals which are noisy realizations around the true (univariate) state. One set of signals has noise which is partially common across agents, while another signal is purely private. Investing attention is costly, but reduces the idiosyncratic variance in observing that signal. The second stage is a finite player coordination game similar to the "stag hunt" game of Carlsson and van Damme (1993a) and similar in many ways to the global games literature reviewed in Morris and Shin (2003).[1] Agents observe signals whose properties are determined by information choices made in the first stage and, based on those signals, form expectations about the underlying state and others' beliefs. Payoffs in the coordination game depend on agents' own actions, the underlying state, and the actions of others. The coordination game has a unique symmetric switching equilibrium supported by symmetric information choice (Proposition 1).

Given the existence and uniqueness of the coordination equilibrium, I then ask whether a unique, symmetric information choice equilibrium always exists. I find it does not. Investing in better information increases ones ability to correctly forecast the and thus helps the committee members avoid costly mistakes. But the need to forecast the future beliefs of others when deciding what information to observe introduces non-concavity in payoffs; investing in information may have negative marginal benefit even setting aside the costs of attention (Lemma 3). Whether utility is "well behaved" depends in a complicated way on all of the model parameters, and I show using numerical examples how there may not be a benefit to investing in information even when the conditions for an equilibrium in the coordination game are satisfied.[2] This stands in contrast to the beauty contest setting of Myatt and Wallace (2012) which has convex loss function which, combined with convex costs, yields a unique (global) solution.

Given the possible pathologies associated with the information choice game, analytic comparative statics are difficult to come by. In section 3.5, I explore a particular numerical example under quadratic attention costs. In each of the parameterizations I consider, the symmetric equilibrium

[1] Unlike the global games literature, which typically assumes discontinuity of payoffs, I assume payoffs for actions are continuous in the state and in the number of players who take an action.

[2] This is distinct from the non-concavity arising in Radner and Stiglitz (1984) because it is unrelated to the costs of acquiring information.

I find appears to be the unique symmetric equilibrium. In that equilibrium, the majority of forecast weight is on private signals, and the best public signals receive the most weight. Lowering the number of committee members reduces expected payoffs of the members only sightly but has larger deleterious effects on overall forecast accuracy. A single member of the committee improving the quality of her publicly-supplied statistic improves welfare slightly but can lead to noticable improvements in forecast accuracy. Increasing the cost of attention, surprisingly, increases welfare and dramatically improves forecast accuracy. Given the results, I conclude by suggesting some tentative normative lessons of the model and possible positive implications that could be tested using data on monetary policy committee behavior and suggestions for future work.

The paper contributes to the literature on information choice in coordination games. Much of the recent work on game theory models of dispersed information (reviewed in Morris and Shin (2003) and Angeletos and Lian (2016)) takes the information structure as given.[3] Exceptions include Hellwig and Veldkamp (2009), Myatt and Wallace (2008, 2012, 2014) and Chahrour (2014) who all focus on endogenous information structures in the context of many agent "beauty contest" games. The latter two papers, in particular, focus on how a (unitary) monetary policymaker should provide information to a continuum of agents who play a beauty contest game. By contrast, I embed an information choice problem similar to that of Myatt and Wallace (2012) in a finite player coordination game, which is more amenable to examining information choices within, for example, policymaking committees.

This paper is closely related to a smaller literature focused on the information choice foundations of Carlsson and van Damme (1993b)-style global games. Szkup and Trevino (2015) study the decision to invest in precision of a private signal in regime change game. They show the uniqueness of symmetric precision choices in their context and show prior information (which they interpret as a public signal) and private information may be complements. My paper differs from theirs in

[3] Morris et al. (2016) show in coordination games there is a unique rationalizable action played whenever there is "approximate common certainty" of the probabilities players assign to their payoffs being higher than their opponents' payoffs. However, they do not examine information choice per se.

examining the endogenous allocation of attention to both public and private signals. Yang (2015) studies general private information acquisition in a coordination game with an entropy-based cost of precision and shows how multiple equilibria can arise. Morris and Yang (2016) show how a more general cost functional can generate uniqueness of equilibria in coordination games if there are continuous payoffs. Both these papers focus on private signals in a setting with atomistic agents. I impose a somewhat stronger information structure.

The motivating example in the paper and the discussion in section 3.5 is are related to a large literature on how monetary policy committees function in practice. Blinder (2009) reviews a number of features of committee design in central banks around the world. The numerical examples explore some of the same issues, particularly, the optimal committee size and the desirability of consensus decision making. There are many features of monetary policy committees that I set aside to focus on the information choice decision, but which would be interesting to explore as extension in future work. For example, Sibert (2003) examines the incentives for heterogeneous committee members to establish reputations for toughness; Faust (1996) analyzes the design of Federal Reserve as a solution to balance interests for and against opportunistic inflation; Waller (1992) examines partisan appointments to the committee. All three papers focus on heterogeneity among central bankers. By contrast, I assume committee members are ex-ante homogeneous and have an exogenous coordination motive. Malmendier et al. (2017) examine empirically how the life experiences of FOMC members influences their expectations and voting behavior; in the setup here, this could be thought of as the weight on private signals (although here private signals are centered on the truth and I consider symmetric equilibria). Finally, the social choice literature has also examined how committees function as deliberative bodies, emphasizing the effects of voting procedures, and how debate and persuasion can affect beliefs and preferences (Li and Suen (2009) review this literature). I abstract from these aspects of committee decisions.

3.2 General description of the game

Players are members of a committee. There are L members, $l = 1, 2, \ldots, L$. Call $\theta \sim N(0, \sigma_\Theta^2)$ the fundamental.

Players form expectations using $L+1$ signals. The signal structure is similar to Myatt and Wallace (2012).[4] L of those signals are associated with each player of the game, and are "public" in the sense that it is common knowledge that a portion of the noise in the signal is common across agents. For example, each signal can be thought of as a forecast circulated before the meeting). One signal is purely private.

Player l's public signals $\mathbf{x_l} = \{x_{nl}\}_{n=1}^{L}$ are

$$x_{nl} = \theta + \eta_n + \varepsilon_{nl}$$

where

$$\eta_n \sim N(0, \kappa_n^2)$$

$$\varepsilon_{nl} \sim N(0, \zeta_n^2 / z_{nl})$$

The private signal is

$$r_l = \theta + \varepsilon_{rl}$$

where

$$\varepsilon_{rl} \sim N(\zeta_{rl}^2 / z_{rl})$$

The noise in signals is somewhat stylized. For public signals, one could imagine η_n being the underlying noisiness of the signal - for example, sampling error from a statistical release - and ε_{nl}

[4] I do not explicitly refer to a (model consistent) prior. The inclusion of the prior very mildly affects the condition needed for uniqueness in the coordination game, but complicates the notation in the information choice stage significantly. Hence, I follow Myatt and Wallace (2012) in setting aside the prior, or assuming that one of the signals is available without cost and treated as the prior.

being noise coming from imperfections in communication or less-than-perfect attention devoted to receiving the information (closely listening to the person presenting it or paying close attention to the underlying detail of the data release).[5] The l-specific noise can be eliminated by investing in a higher z_{nl}, which I sometimes informally refer to as paying closer attention to the signal (i.e., spending more time examining the data) but the former cannot. Investment in purely private signal precision z_{rl} might be though of as investing effort in collecting and analyzing data privately. A choice of $z_{kl} = 0$ implies a signal has infinite variance; such a signal will be ignored, so choosing to invest no attention to a signal is equivalent to ignoring it entirely. Without loss of generality, I assume the public signals are ordered by "quality" with less noisy signals first, so $\zeta_1^2 < \zeta_2^2 < \ldots < \zeta_L^2$.

The precision of the public and private signals are the inverse of their variance, defined respectively:

$$\psi_{nl} = \frac{1}{\kappa_n^2 + \frac{\zeta_n^2}{z_{nl}}}$$

$$\psi_{rl} = \frac{z_{rl}}{\zeta_{rl}^2}$$

In the first stage of the game, agents choose a vector z_l (the extent to which they want to eliminate idiosyncratic noise in their future signals). I refer to this choice as agents' investment in precision. This investment comes at a cost $C(z_l)$.

After the first stage, nature chooses realizations of the fundamental, the common component of noise in public signals, and the idiosyncratic noise in each players' signals. l's posterior beliefs are formed in a Bayesian manner given her realized signals, so her conditional expectations are

$$E^l(\theta|\mathbf{x_l}, r_l) = \frac{\sum_{n=1}^{L} \psi_{nl} x_{nl} + \psi_{rl} r_l}{\sum_{n=1}^{L} \psi_{nl} + \psi_{rl}} \tag{3.1}$$

Define $w_{k,l} = \frac{\psi_{nl}}{\sum_{n=1}^{L} \psi_{nl} + \psi_{rl}}$. By construction, $\sum_k w_k = 1$. Posterior uncertainty is the inverse of the sum of the precisions, i.e.

[5]The communication interpretation of noise is similar to Prat et al. (2015).

$$\mathrm{Var}(\theta|\mathbf{x_l}, r_l) = \left(\sum_{n=1}^{L}\psi_{nl} + \psi_{rl}\right)^{-1} \tag{3.2}$$

In the second stage, given her beliefs about the fundamental $E^l(\theta|\mathbf{x_n}, r_n)$, members make a binary choice to maximize expected payoffs. The action space in the coordination game is

$$a_l \in \{0, 1\}$$

Payoffs are related to the realization of the fundamental, the choice made in the coordination game, the choices of other agents, and the costs of allocating attention to signals, summarized by the cost function $C(z_l)$.

Call a_{-l} the action profile of players other than l. The payoff of player l is:

$$\begin{aligned} u_l(a_l, z_l, a_{-l}, \theta) =& \alpha a_l(a_l - \bar{a}) + \beta(1 - a_l)(a_l - \bar{a}) \\ &+ \gamma a_l(\bar{a} - \frac{k}{L}) + \delta(1 - a_l)(\frac{k}{L} - \bar{a}) \\ &+ \lambda a_l\theta + \mu(1 - a_l)\theta + \nu a_l + \tau(1 - a_l) \\ &- C(z_l) \end{aligned} \tag{3.3}$$

where $\bar{a} = (1/L)\sum_{n=1}^{L} a_n$.

The payoff function is reduced form but is intended to capture the intuition discussed in the introduction. The payoffs reflect inherent costs and benefits of taking the action, whether their action is similar to everyone else's, and whether the average action is above or below a cutoff value k/L. It also reflects a linear payoff in the fundamental that varies across which actions agents take.[6]

In the next section, I describe the solution to the coordination game conditional on information choice, including restrictions on parameters in the utility function sufficient for a symmetric switching equilibrium. Given this equilibrium, I analyze information choice in section 3.4.

[6] A more realistic applied model would likely include some interaction of the state with whether the average action was above or below the cutoff. This is excluded to simplify the information stage.

3.3 Coordination game

To solve the model, I adapt the solution method developed in Morris and Shin (2003). In particular, I establish some properties of the optimal strategy given the payoff structure (equation (3.3)). The structure of the game, along with some assumptions about the parameters in equation (3.3), gives rise to a switching strategy, which is unique given symmetric information choice.

3.3.1 Payoff properties

Payoffs can be parameterized by a function:

$$\begin{aligned}\pi_l(a_{-l},\theta) = u(1,a_{-l},\theta) - u(0,a_{-l},\theta) =& \alpha\left(1-\frac{1}{L}\right) + \gamma\left(\frac{1-k}{L}\right) - \delta\frac{k}{L} + \nu - \tau \\ &+ \frac{1}{L}(\beta - \alpha + \gamma + \delta)\sum_{i\neq l} a_i \\ &(\lambda - \mu)\,\theta\end{aligned}$$

When $\pi(a_{-l},\theta) > 0$ $a_l = 1$ is preferable, otherwise l should choose $a_l = 0$.

This can be written more compactly as

$$\pi_l(a_{-l},\theta) = \alpha^* + \frac{\beta^*}{L}\sum_{i\neq l} a_i + \lambda^*\theta \tag{3.4}$$

I make the following assumptions about these parameters:

Assumption 1 (Action Monotonicity). Payoffs for $a_l = 1$ are increasing in the number of others who take $a_i = 1, i \neq l$. That is, $\beta^* \equiv \beta - \alpha + \gamma + \delta > 0$.

Assumption 2 (State monotonicity). $a_l = 1$ is more preferable in higher states of the world, all else equal. That is, $\lambda - \mu \equiv \lambda^* > 0$.

These assumptions immediately imply the following lemma, whose proof follows from direct calculation and is included (along with other proofs and derivations) in the appendix:

Lemma 1 (Limit dominance). *There exists $\underline{\theta}, \bar{\theta}$ such that for all $\underset{\sim}{\theta} < \underline{\theta}$ and for all a_{-l}, $\pi(a_{-l},\underset{\sim}{\theta}) < 0$ and for all $\widetilde{\theta} > \bar{\theta}$ and for all a_{-l}, $\pi(a_{-l},\widetilde{\theta}) > 0$.*

3.3.2 Switching strategy

I look for equilibria in the space of symmetric switching strategies.

Definition 2 (Switching strategy)**.** *A switching strategy in the coordination game* $\mathbf{s} : \mathbb{R} \to \{0, 1\}$ *maps observed signals* $\mathbf{x_l}, r_l$ *onto actions.*

In particular, a symmetric switching strategy is a strategy profile such that for all l*,*

$$\boldsymbol{s}(\boldsymbol{x}_l, r_l) = \begin{cases} 1 \text{ if } E^l(\theta|\mathbf{x_l}, r_l) > \theta^* \\ 0 \text{ if } E^l(\theta|\mathbf{x_l}, r_l) \leq \theta^* \end{cases} \tag{3.5}$$

I make the following assumption about precision choices to facilitate exposition:

Assumption 3 (Common knowledge of signal noise)*.* The precision of each players' signals is common knowledge in the coordination stage.

Beliefs about the beliefs of others. With some abuse of notation, denote $E^l(\theta)$ the posterior belief of l. l knows p's expectations are the same as her own. Given signal precisions are common knowledge by assumption 3, her expectation of p's belief is a function of her belief about p's signals. Hence

$$\begin{aligned} E^l(x_{np}) =& E^l(\theta) + E^l(\eta_n) + E^l(\varepsilon_{np}) \\ =& \; E^l(\theta) \end{aligned} \tag{3.6}$$

The *variance* of l's beliefs about p's signals are

$$\begin{aligned} \mathrm{Var}^l(x_{np}) =& \mathrm{Var}^l(\theta) + \mathrm{Var}^l(\eta) + \mathrm{Var}^l(\varepsilon_{np}) \\ =& \left(\sum_n \psi_{nl} + \psi_{rl} \right)^{-1} + \psi_{np}^{-1} \end{aligned} \tag{3.7}$$

Symmetric information choice implies that $p \neq l$, we can write $\psi_{np} = \psi'_n$ and $\psi_{rp} = \psi'_r$. We now have all the pieces needed to prove the existence and uniqueness of a symmetric switching strategy given symmetric information choices.

Proposition 1 (Symmetric switching equilibrium)**.** *Given payoffs described by* (3.3) *and assumptions 1 and 2 and beliefs characterized by* (3.1),(3.2), (3.6), *and* (3.7), *there exists a unique symmetric switching algorithm in the coordination game where the threshold* θ^* *is given by*

$$\theta^* = -\frac{1}{\lambda^*}\left(\alpha^* + \frac{\beta^*}{2}\frac{(L-1)}{L}\right) \tag{3.8}$$

The proof is in appendix 3.A. Briefly, given that everyone forms expectations optimally and is playing the symmetric switching strategy, one can directly solve for a θ^* as the value that would make l indifferent between playing either action, given her beliefs. The assumption of state monotonicity is sufficient in this model to show that at θ^* payoffs are increasing and thus the optimal action does not switch back to $a_l = 0$ for some $\theta > \theta^*$, for instance. The coordination game satisfies standard conditions (Carlsson and van Damme (1993b), Morris and Shin (2003)) which shows that it is the only switching equilibrium that survives the iterated deletion of dominated strategies. The proof of the theorem also includes a comment on how it goes through when we explicitly include prior information.

3.3.3 Two convenient normalizations

Inspection of (3.8) shows for any value of β^* and λ^* we can adjust α^* to generate $\theta^* = 0$. In particular, we can adjust the intrinsic payoffs associated with each action (τ, ν) without affecting the other terms. This normalization does not affect the uniqueness of θ^* or the proof of Proposition 1. Hence, I normalize $\theta^* = 0$ for the remainder of the paper. We also can normalize $\mu = 0$ without loss of generality, since only $\lambda - \mu > 0$ is necessary for uniqueness.

3.4 Information choice

Other than symmetric information, the assumptions made to supporting a unique symmetric switching equilibrium in the coordination stage are fairly weak; they amount to Bayesian updating, strategic complementarity, and defining the payoffs so the payoff to $a_l = 1$ is increasing in the state. Symmetric information choice is widely assumed in the literature on coordination games. In this section, I describe how the information optimization problem can fail to be well behaved.

3.4.1 Preliminaries

Future conditional expectations. Optimal information choice maximizes ex-ante utility given the optimal switching strategy in the coordination stage. It amounts to choosing the variance of the distribution of conditional expectations in the coordination stage to maximize expected payoffs. Player l's future conditional expectation can be written in terms of the θ and the sum of the realizations of noise:

$$\begin{aligned} E^l(\theta) =& \theta + \sum_{n=1}^{L} w_{nl}(\eta_n + \varepsilon_{nl}) + w_{rl}\varepsilon_{rl} \\ =& \theta + Q(l) \end{aligned} \tag{3.9}$$

Note $Q(l) \sim N(0, \sigma^2_{Q(l)})$ where

$$\sigma^2_{Q(l)} = \sum_{n=1}^{L} (w_{nl})^2 \left(\kappa_n + \frac{\zeta_n^2}{z_{nl}} \right) + w_{rl}^2 \frac{\zeta_{rl}^2}{z_{rl}} \tag{3.10}$$

Then $E^l(\theta) = \theta + Q(l) \sim N(0, \sigma^2_\Theta + \sigma^2_{Q(l)})$

Ex-ante utility. Under the symmetric switching strategy, $a_l = 1$ when $E^l(\theta) > \theta^* = 0$ for all l. I look for symmetric information choices as well, which imply other players choose $\sigma^2_{Q(p)} = \sigma^2_{Q'}$ for all p. Ex-ante expected utility is

$$\begin{aligned} E(U_l) &= E(U_l | E^l(\theta) > 0) \times P(E^l(\theta) > 0) + E(U_l | E^l(\theta) \leq 0) \times P(E^l(\theta) \leq 0) \\ &= E(U_l | a_l = 1, \theta + Q) P(\theta + Q > 0) + E(U_l | a_l = 0, \theta + Q) P(\theta + Q \leq 0) \end{aligned} \tag{3.11}$$

I show in appendix 3.B the following lemma:

Lemma 2 (Ex-ante utility)**.** *Ex-ante utility, conditional on* z_l, z' *is*

$$
\begin{aligned}
E(U_l) =& \frac{1}{2}\left(\alpha\frac{L-1}{L} + \gamma\left(\frac{1-k}{L}\right) + \nu + \frac{\delta k}{L} + \tau\right) && \text{(3.12a)}\\
&+ \frac{\lambda}{\sqrt{2\pi}} \frac{\sigma_\Theta^2}{\sqrt{\sigma_\Theta^2 + \sigma_Q^2}} && \text{(3.12b)}\\
&+ \frac{L-1}{L}(\beta - \alpha + \gamma + \delta) \int_{\theta=-\infty}^{\infty} \int_{q=-\theta}^{\infty} \Phi\left(\frac{\beta(z,z')(\theta+q)}{\sqrt{(\sigma_\theta^2 + \sigma_{Q'}^2)(1-\rho^2)}}\right) f_Q(q) f_\Theta(\theta) dq d\theta && \text{(3.12c)}\\
&- C(z) && \text{(3.12d)}
\end{aligned}
$$

Where

$$
\rho = \frac{Cov(\theta+Q, \theta+Q')}{\sqrt{\sigma_\theta^2 + \sigma_{Q'}^2}\sqrt{\sigma_\theta^2 + \sigma_Q^2}} = \frac{\sigma_\theta^2 + \sum_n w_n w_n' \kappa_n^2}{\sqrt{\sigma_\theta^2 + \sigma_{Q'}^2}\sqrt{\sigma_\theta^2 + \sigma_Q^2}} \tag{3.13}
$$

and

$$
\beta(z, z') = \frac{\sigma_\theta^2 + \sum_n w_n w_n' \kappa_n^2}{\sigma_\theta^2 + \sigma_Q^2} \tag{3.14}
$$

and $f_Q(q)$ *and* $f_\Theta(\theta)$ *are the PDFs of* Q *and* θ*.*

What's the intuition for this expression?

- (3.12a) is the constant part of expected payoffs for each action
- (3.12b) is the expected payoff associated with the state. In the absence of noise, the this would just be the expected value of half-normal distribution, but it is adjusted to account for the fact that a player might pick $a_l = 1$ when they (incorrectly) believe $\theta > 0$ due to noise.
- (3.12c) reflects the expected payoff from others' actions, given the switching equilibrium in the second stage.
- Finally, (3.12d) is the cost of attention allocation.

3.4.2 Marginal benefits

Agents allocate attention by paying a cost to receive more precise signals. These benefits come from choosing $a_l = 1$ only when $\theta > 0$ (the "direct effect") and from picking the same action as others (the "coordination effect"). Below, I analyze the direct and coordination benefits of increased attention to a particular signal, i.e., an increase in z_k holding fixed $z_{k'}, k' \neq k$.

The following remark states some results about the effect of paying more attention to a particular signal on weights, noise, and the correlation of the distribution of beliefs (with derivations in appendix 3.B)

Remark 1 (Effects of a change in attention). Increasing attention to a particular signal k (an increase in z_{nk}):

1. Unambiguously increases precision:

$$\frac{\partial \psi_{kl}}{\partial z_{kl}} = \begin{cases} \frac{\zeta_k^2}{(z_{kl}\kappa_k^2 + \zeta_k^2)^2} & k \in 1, 2, \ldots, L \\ \frac{1}{\zeta_r^2} & k = r \end{cases}$$

2. Has weakly diminishing returns on precision: $\frac{\partial^2 \psi_{kl}}{\partial z_{kl}^2} \leq 0$.

3. Increases the weight on the signal paid attention to and decreases the weight on other signals that receive positive attention: $\frac{\partial w_{kl}}{\partial z_{kl}} > 0$, $\frac{\partial w_{k'l}}{\partial z_{kl}} < 0$

4. Causes the overall variance of noise to fall: $\frac{\partial \sigma_Q^2}{\partial z_{kl}} < 0$

5. Changes the distribution of noise: $\frac{\partial f_Q(q)}{\partial z_{kl}} = f_Q(q) \left(\frac{q^2}{\sigma_Q^2} - 1 \right) \frac{\partial \sigma_Q^2}{\partial z_{kl}}$

6. Has ambiguous effect on $\beta(z, z')$ and ρ. $\square$

The punchline of the remark is that increased attention to a signal causes noise to fall but has potentially ambiguous effects on whether posterior beliefs are expected to be more correlated. This is intuitive: Investing in a more precise signal means that signals will be closer to the realization of θ. However, investing in a signal that others are not paying attention to lowers the weight of shared information in forming posterior beliefs and means that those posterior beliefs may be less correlated. Counterbalancing this is the fact that having a more accurate forecast of the fundamental implies a more accurate forecast of the average realization of others' noisy signals.

With these preliminaries out of the way, we can proceed to examining the benefit to a marginal increase in precision.

Direct effect. The direct effect of a marginal increase in attention to a particular signal (an increase in z_{kl}), setting aside the coordination effects:

$$\frac{\partial}{\partial z_{kl}}\left(\lambda\frac{\sigma_\Theta^2}{\sqrt{2\pi}}\frac{1}{\sqrt{(\sigma_\Theta^2+\sigma_Q)^2}}\right)=-\frac{1}{2}\frac{\lambda}{\sqrt{2\pi}}\frac{\sigma_\theta^2\frac{\partial\sigma_Q^2}{\partial z_{kl}}}{(\sigma_\Theta^2+\sigma_Q^2)^{3/2}} \tag{3.15}$$

This is unambiguously positive because investing in precision reduces noise ($\frac{\partial\sigma_Q^2}{\partial z_{kl}}<0$). Investing in precision means you are less likely to make a costly mistake in forecasting the state.

Coordination effect. The coordination effect is more intricate than the direct effect. Suppressing the (positive) coefficient and taking the derivative inside the integral:

$$\begin{aligned}
&\frac{\partial}{\partial z_{kl}}\int_{\theta=-\infty}^{\infty}\int_{q=-\theta}^{\infty}\Phi\left(\frac{\beta(z,z')(\theta+q)}{\sqrt{(\sigma_\theta^2+\sigma_{Q'}^2)(1-\rho^2)}}\right)f_Q(q)f_\Theta(\theta)dqd\theta\\
&=\int_{\theta=-\infty}^{\infty}\int_{q=-\theta}^{\infty}\Bigg\{\phi\left(\frac{\beta(z,z')(\theta+q)}{\sqrt{(\sigma_\theta^2+\sigma_{Q'}^2)(1-\rho^2)}}\right)\frac{(\theta+q)}{\sqrt{(\sigma_\Theta^2+\sigma_{Q'}^2)}}\left[\frac{\beta'(z,z')}{\sqrt{(1-\rho^2)}}+\frac{\beta(z,z')\rho\frac{\partial\rho}{\partial z_{nk}}}{(1-\rho^2)^{3/2}}\right]\\
&\qquad+\Phi\left(\frac{\beta(z,z')(\theta+q)}{\sqrt{(\sigma_\theta^2+\sigma_{Q'}^2)(1-\rho^2)}}\right)\frac{\partial\sigma_Q}{\partial z_{kl}}\left(\frac{q^2}{\sigma_Q^2}-1\right)\Bigg\}f_Q(q)f_\Theta(\theta)dqd\theta
\end{aligned}$$

The first integral is:

$$\frac{\frac{\sqrt{(\sigma_\theta^2+\sigma_Q^2)}}{\sqrt{(\sigma_\theta^2+\sigma_{Q'}^2)}}\left[\frac{\beta'(z,z')}{\sqrt{(1-\rho^2)}}+\frac{\beta(z,z')\rho\frac{\partial\rho}{\partial z_{nk}}}{(1-\rho^2)^{3/2}}\right]}{2\pi\left(1+\left(\frac{\beta(z,z')}{\sqrt{(\sigma_\theta^2+\sigma_{Q'}^2)(1-\rho^2)}}\right)^2(\sigma_\theta^2+\sigma_Q^2)\right)}$$

The sign of this effect is ambiguous, because increased attention has an ambiguous effect on both

ρ and $\beta(z, z')$.

The second is:

$$\frac{\frac{\partial \sigma_Q^2}{\partial z_{kl}} \frac{\beta(z,z')}{\sqrt{(\sigma_\theta^2+\sigma_{Q'}^2)(1-\rho^2)}} \sigma_Q^2}{2\pi \left(1 + \left(\frac{\beta(z,z')}{\sqrt{(\sigma_\theta^2+\sigma_{Q'}^2)(1-\rho^2)}}\right)^2 (\sigma_\theta^2 + \sigma_Q^2)\right) \sqrt{\sigma_Q^2 + \sigma_\Theta^2}}$$

This is strictly negative.

The first term reflects the fact that investing in attention to a particular signal affects the extent to which posterior beliefs are likely to be correlated. The second reflects the fact that, holding $f_\Theta(\theta)$ fixed, shrinking noise means the distribution of possible beliefs is shrinking. To summarize:

Remark 2 (Net coordination effects). The net direction of the coordination effect is ambiguous. □

The following follows immediately:

Lemma 3 (Expected losses from investing in information)**.** *Increasing signal precision may have negative marginal effects on expected utility. Direct comparison of the direct and coordination effects shows that the sign of the benefits side of investing in coordination is*

$$\begin{aligned} Sign\Bigg\{ & \frac{\partial \sigma_Q^2}{\partial z_{kl}} \left[\frac{\beta^*(L-1)\rho\sqrt{1-\rho^2}\sigma_Q^2}{2\pi L\left(\sigma_\theta^2 + \sigma_Q^2\right)} - \frac{\lambda \sigma_\theta^2}{2\sqrt{2\pi}\left(\sigma_\theta^2 + \sigma_Q^2\right)^{3/2}} \right] \\ & + \frac{\beta^*(L-1)\left(\rho^2 \frac{\partial \rho}{\partial z_{kl}} \sqrt{\sigma_\theta^2 + \sigma_{Q'}^2} + \frac{\partial \beta}{\partial z_{kl}}\left(1-\rho^2\right)\sqrt{\sigma_\theta^2+\sigma_Q^2}\right)}{2\pi L\sqrt{(1-\rho^2)\left(\sigma_\theta^2+\sigma_{Q'}^2\right)}} \Bigg\} \end{aligned} \tag{3.16}$$

□

The term in square brackets in (3.16) can be negative when the benefit to coordinating is sufficiently large relative to the direct benefit. Per the last part of remark 1, the second line of (3.16) has ambiguous sign. Intuitively, large β^* relative to λ^* can result in the marginal direct benefit to investing in more accurate signals being smaller than the losses in expected coordination utility.

Table 3.1: Baseline calibration

Parameter	Interpretation	Value
L	Number of players	12
λ	Coefficient on fundamental in utility	1
β^*	Coefficient on sum of others' actions in utility	0.1
σ_θ^2	Variance of fundamental	2
κ_i^2, $i = 1, \ldots, L$	Common component of noise in public signals	4
ζ_1^2	"Receiver" noise in first public signal	0
ζ_i^2, $i = 2, \ldots, L$	"Receiver" noise in public signals	0.05-0.25
ζ_r^2	Receiver noise in private signal	.3
z, z'	Amount of investment in reducing signal noise	0.05/varies

This also depends on the variance of the fundamental and the noise in signals.[7]

This non-concavity is distinct from the non-concavity noted by Radner and Stiglitz (1984), because it is unrelated to the marginal *cost* of acquiring information. Like in their paper, and in Myatt and Wallace (2012), it may also be that in equilibrium some signals are ignored because they are too costly to acquire in the first place. But it turns out that we can generate large regions of investment in signal precision that have negative marginal benefits to investing in attention even setting aside the cost of investment.

Numerical illustration. To illustrate the properties of increased precision on expected benefits of investing in more precise signals, I calibrate the model and calculate marginal utilities, letting the amount of attention paid to *one* signal vary. Details of the baseline calibration are in table 3.1.[8]

Figure 3.1 shows the effect of a marginal change in z_{kl} on the non-cost components of utility, decomposed into the direct effect and coordination effects, varying z_{2l} from 0.001 to 1. The first 12 signals are numbered in order of their quality (i.e., signal 1 has the lowest noise ζ_i^2 has the second lowest, etc) while signal "13" corresponds to the purely private signal. Under this calibration, the direct benefit is large, and the net effect from coordination is also generally positive.

[7] There is some economic intuition behind the result that a small λ leads to negative incentive to invest in information. In the coordination game, with α^* selected to ensure θ^* is zero given β^*, small λ^* implies that the value above which one should always choose $a_l = 1$ regardless of what others are doing $\bar{\theta}$ is very negative. Depending on the the variance of the fundamental, it may be that there is essentially no chance of a θ realization that would result in $a_l = 0$ being optimal.

[8] The parameters chosen are arbitrary in the sense that they are not pegged to a particular real-world variable. $L = 12$ corresponds with the number of voting members on the FOMC (although the committee has 19 members and each report their views.)

Figure 3.2 illustrates the change in σ_Q^2, ρ and $\beta(z, z')$ as z_{2l} increases. For the baseline calibration, the correlation of beliefs is quite high, and increasing as investment in the public signal increases. Figure 3.3 shows the change in precision of signal 2 and the corresponding changes in signal weights. Finally, figure 3.4 shows how a marginal increase in attention paid to each signal (as z_{2l} changes) affects the derivatives discussed in remark 1.

By contrast, if $\sigma_\Theta^2 = 1$, $\lambda = 0.0001$, $\beta^* = 5$, marginal utility is negative at low levels of z_{2l}. This is illustrated in figure 3.5. This is due to the (much smaller) direct effect of investing in signal quality and larger coordination losses.

Returning to the baseline case with the parameters in table 3.1, I illustrate the effect of increased investment in purely private information in figure 3.9. As before, increase investment in private information leads to diminishing (but positive) marginal benefits. However, the marginal returns to investing in additional private information are constant, so as z_{rl} increases, private signal weight continues to increase towards 1. Figure 3.10 shows ρ is declining in increased z_{rl}. This illustrates the ambiguity mentioned at the end of Remark 1.

3.4.3 Cost functions

Even in the case where marginal utilities are well-behaved (in the sense of increasing in z), the properties of the cost function $C(z)$ will generally affect equilibrium existence and uniqueness. In the context of a beauty contest game, Myatt and Wallace (2012) shows in the context of a beauty contest game, nonconvexity of costs (for example, entropy-based costs as in the rational inattention literature) may yield multiple equilibria, but with convex costs and a convex loss function there will be a unique equilibrium. Here, the payoff function is not necessarily concave so a convex cost function is insufficient to guarantee an equilibrium.

3.4.4 Equilibrium

Now that all of the pieces are in place, I can fully characterize the equilibrium of the two stage game.

Definition 3 (Equilibrium)**.** *A pure strategy symmetric Perfect Bayesian Nash Equilibrium is a set of investment choices z_k and a threshold choice θ^* such that for all players $l = 1, 2, \ldots, L$:*

1. *For all $l \in \{1, 2, \ldots, L\}$, the choice of element z_k of $\bar{z}$ is $z_k = max(0, z_k^*)$ where z_k^* solves:*

$$
\begin{aligned}
&-\frac{1}{2}\frac{\lambda}{\sqrt{2\pi}}\frac{\sigma_\theta^2 \frac{\partial \sigma^2_{Q(\bar{z})}}{\partial z_k^*}}{(\sigma_\Theta^2+\sigma^2_{Q(\bar{z})})^{3/2}} + \beta^* \frac{\left[\frac{\beta'(z,z')}{\sqrt{(1-\rho(\bar{z},\bar{z})^2)}} + \frac{\beta(z,z')\rho(\bar{z},\bar{z})\frac{\partial\rho(\bar{z},\bar{z})}{\partial z_{nk}}}{(1-\rho(\bar{z},\bar{z})^2)^{3/2}}\right]}{2\pi\left(1+\left(\frac{\beta(\bar{z},\bar{z})}{\sqrt{(1-\rho(\bar{z},\bar{z})^2)}}\right)^2\right)} \\
&+\beta^* \frac{\frac{\partial \sigma^2_{Q(\bar{z})}}{\partial z_k^*}\frac{\beta(\bar{z},\bar{z})}{\sqrt{(\sigma_\theta^2+\sigma^2_{Q(\bar{z})'})(1-\rho(\bar{z},\bar{z})^2)}}\sigma_Q^2}{2\pi\left(1+\left(\frac{\beta(\bar{z},\bar{z})}{\sqrt{(1-\rho(\bar{z},\bar{z})^2)}}\right)^2\right)\sqrt{\sigma^2_{Q(\bar{z})}+\sigma^2_\Theta}} \qquad = C_k'(\bar{z})
\end{aligned}
\tag{3.17}
$$

   with $z_k^* > 0$ for at least one z_k where $\sigma^2_{Q(\bar{z})}$ is defined as in (3.10)*, $\rho(\bar{z},\bar{z})$ is defined as in* (3.13)*, $\beta(\bar{z},\bar{z})$ is defined as in* (3.14)*, and $C(\bar{z})$ is an increasing, continuous differentiable function.*

2. *Expectations in the coordination game are given by* (3.1) *with variance* (3.2)*.*

3. *For all $l \in \{1, 2, \ldots, L\}$, $a_l = 1$ when $E^l(\theta) > \theta^*$ and $a_l = 0$ when $E_l(\theta) \leq \theta^*$, where θ^* satisfies*

$$\theta^* = -\frac{1}{\lambda^*}\left(\alpha^* + \frac{\beta^*}{2}\frac{(L-1)}{L}\right) \tag{3.18}$$

The first element of definition 3 says, in the attention allocation stage, signals must satisfy a first-order condition of a the player's constrained utility maximization problem, where the constraint is that at least one cannot have $z_k = 0$. The definition allows for corner solutions, in that some signals may have zero precision ($z_k = 0$), but does impose the restriction that at least one signal is observed so the expectation in the coordination game are well-defined, and I restrict the choices to be symmetric. Expectations (which are required to follow Bayes' rule) are the second condition of the definition. The third element restates the switching equilibrium in symmetric strategies.

In section 3.3, I showed that the coordination game satisfies sufficient conditions for a unique switching equilibrium, conditional on symmetric solution to the attention allocation game. It is easy to see how such an equilibrium could fail to arise. For instance, in figure 3.5, marginal utility was negative over a region of z even gross of the costs of investing in information, despite conditions for uniqueness in the coordination game being satisfied. Given positive and increasing costs to investing, an equilibrium may fail to exist because it is too costly to acquire the amount of precision that generates any investment in information, which leads to a failure of expectations being well-defined in the coordination stage.

3.5 Numerical analysis

In the previous section, I derived conditions for a symmetric information allocation equilibrium and shown that there are circumstances where those sufficient conditions can fail to be satisfied. However, I have not demonstrated whether any symmetric equilibrium exists, and if such equilibria are unique. In this section, I explore uniqueness and comparative statics in a computational example.

To completely characterize the model, I need to add a cost function. As Myatt and Wallace (2012) show, in beauty contest games the exact specification of the cost function matters for which signals receive attention and whether there is a unique equilibrium. Since the calibration is somewhat arbitrary, I assume costs are quadratic in the total attention allocated. In particular, the baseline cost function is

$$C(z) = .375 \left(\sum z_k\right)^2 \tag{3.19}$$

Uniqueness, signal weights and signal weights under the baseline baseline. I solve for equilibrium by numerically searching for symmetric attention allocations that satisfy the first order condition in definition (3). I verify uniqueness using a numerical global search method. I find no other solutions which satisfy the first order condition and satisfy the constraints on $\bar{z}$.[9]

The equilibrium precisions for this calibration are shown in the first column of table 3.2. Recall the first signal has zero idiosyncratic noise, and the other signals are ordered by quality; private signals have $\kappa_n^2 = 0$ but $\zeta_r^2 > \zeta_k^2$ for $k = 2, \ldots, L$. Unsurprisingly, more attention is paid to signals have are "higher quality" and the worst signal receives nearly no weight. The private signal receives the most weight, although about 70% of committee members' beliefs are driven by the public signals. This is despite the fact that the private signal does not have diminishing returns to accuracy. Committee members endogenously choose to allocate costly attention to strictly worse (public) information because it helps them coordinate with others. Within the set of public signals, they allocate more attention to the signals that are more accurate, although all signals receive at least a little attention.

With these baseline results, I conduct some simple comparative static exercises. Weights under the different scenarios are shown in the remaining columns of table 3.2. Ex-ante welfare, overall signal accurate, and costs (both in "raw" terms and relative to the baseline) are shown in table 3.3.

A smaller committee. One feature of committee decisions is the number of members. Blinder (2009) documents that the number of members involved in setting monetary policy varies across countries. One trade off when picking committee size might be that when there are many members, many different views are represented, but there may be limits to the extent committee members can absorb information provided by others. In the context of the current example, I set the number of committee members $L = 3$. I choose the "best" three members (those with the lowest idiosyncratic signal variance). As the second column of table 3.2 shows, there is some small re-allocation of

[9]In general, whether the algorithm succeeds or fails to find a symmetric attention allocation is sensitive to the starting value. However, whenever it succeeds, it converges to the same value. While I have found uniqueness in all of the example calibrations considered, it is possible there are regions of the parameter space where multiple equilibria can exist. As the discussion in the previous section suggests, it is straightforward to find parameterizations where no symmetric equilibrium appears to exist.

	Baseline	Smaller L	Higher costs	Smaller κ_2	Smaller σ_θ^2
1	0.147	0.176	0.097	0.138	0.172
2	0.093	0.108	0.064	0.177	0.107
3	0.082	0.095	0.057	0.078	0.095
4	0.073	-	0.051	0.070	0.084
5	0.064	-	0.045	0.061	0.074
6	0.056	-	0.040	0.054	0.064
7	0.048	-	0.035	0.046	0.056
8	0.041	-	0.030	0.040	0.047
9	0.034	-	0.025	0.033	0.039
10	0.028	-	0.020	0.027	0.032
11	0.021	-	0.015	0.020	0.024
12	0.015	-	0.011	0.014	0.017
Private	0.297	0.622	0.510	0.241	0.190

Table 3.2: Weights in numerical examples

attention to the remaining public signals (weights are increased by about 1-3 %).

However, in equilibrium it appears much of the attention previously allocated to the lower-quality public signals is reallocated to the private signal. Moreover, committee members find it less worthwhile to invest in attention generally; the total spent on investment drops by nearly 18% relative to the baseline. Forecast quality suffers; overall noise in beliefs is about 20% higher. Welfare losses are not large, however; welfare only falls by about 3% and when we adjust for the fact that coordination payoffs are affected by the number of agents ($(L-1)/L$ shows up in the ex-ante welfare expression), losses are about 1% relative to the baseline. In short, with fewer committee members it is less optimal to invest in information for coordination reasons, and the committee members' forecasts are worse on average. If social welfare is related to the accuracy of forecasts, then a smaller committee may be deleterious to the public interest. This rationalizes, to an extent, the oft-mentioned justification for having a diversity of views represented on monetary policy committees.

	Ex-ante welfare	Relative Welfare	σ_Q^2	Relative σ_Q^2	$C(Z)$	Relative $C(z)$
Baseline	0.493	1.000	0.589	1.000	0.042	1.000
Smaller L	0.478	0.970	0.704	1.194	0.035	0.826
(Adj. for size)	0.489	0.991	-	-	-	-
Higher costs	0.508	1.031	0.389	0.661	0.048	1.138
Smaller κ_2	0.495	1.004	0.553	0.938	0.044	1.052
Smaller σ_θ^2	0.317	0.642	0.687	1.166	0.026	0.613

Table 3.3: Welfare, accuracy, and total attention expenditure in numeric examples.

Higher cost schedule. The "cost" of attention is stylized; one could imagine different reasons why it might be more or less difficult to increase the precision with which committee members observe a particular signal. To gain some intuition about how the cost function affects attention allocation, I adjust equation (3.19) by increasing the cost schedule:

$$C(z) = .5\left(\sum z_k\right)^2 + 0.1\left(\sum z_k\right) \tag{3.19$'$}$$

Under this adjusted cost function, total costs increase at a faster rate, and a marginal increase in attention now has a constant component. Hence, once the marginal benefit from an increase in precision drops below 0.1, it will no longer be worth it to invest in that signal. Column 3 of table 3.2 shows the weights in this scenario. Less weight is placed on private signals - whose marginal gains in precision are diminishing - and the majority of the weight falls on private information. Perhaps surprisingly, however, welfare is only mildly worse with the higher cost schedule and accuracy is much improved. The latter comes from much higher investment in z (which is about 80% higher, not shown) while total costs are less than 14% higher). This counterintuitive result likely comes from the nonconcavities of the utility function and perhaps reflects partially the Radner and Stiglitz (1984) logic that the net benefit of investing in information is likely negative at low levels of precision; if optimizing committee members equal benefits with costs, they may have to invest more when costs increase.

An increase in the accuracy of a public signal. Although treated as a primitive of the environment, one could imagine the common component of noise in public signals κ_n^2 is a choice variable. For example, prior to the information allocation stage of the game, a committee member could have improved data collection or exerted more effort in producing a forecast. Although the costs of this effort are abstracted from here, we can ask what effect an improvement in the signal has on attention allocation and welfare. Accordingly, column 4 of table 3.2 reflects the weights on signals when κ_2 is reduced from 4 to 2. ζ_i^2 is unchanged. In this scenario, more weight is allocated towards the second signal, mostly at the expense of private information and the first signal. This result is intuitive; agents who care both about accuracy and coordinating with others should put more weight on the best public signals. There are still diminishing returns to investing in the precision of the public signals, however, so it still does not make sense for committee members to allocate all of their attention to the public signal. As the sixth line of table 3.3 shows, the improvement in the fundamental precision of signal 2 does improve the overall accuracy of forecasts, but the overall improvement in welfare is quite small. This suggests, in a richer model where the costs of providing signals is explicitly taken into account and where improving signal quality is costly, equilibrium public signals may be lower quality than would be socially optimal.

Smaller fundamental variance. Finally, I modify the baseline calibration by reducing σ_Θ^2 from 2 to 1. This would reduce welfare in the full-information case because payoffs directly related to θ are truncated below by zero and the mass of the distribution is also close to zero. Here, a similar logic applies; inspection of (3.15) reveals the direct benefit of greater attention is increasing in the variance of θ. Although coordination benefits also depend indirectly on σ_Θ^2, the weights reported in the final column of table 3.2 show a re-allocation of signal weights away the private signal towards public signals is optimal when fundamental variance falls. Given that state-related payoffs are smaller (since the realizations of θ are smaller), committee members invest much less in information overall, and σ_Q^2 is larger. In signal-to-noise-ratio terms, it is much worse since the variance of θ has dropped by half; under the baseline, the signal-to-noise ratio is about 3.4, while in this case it is about 1.5.

Positive implications of the numerical analysis. Although the model is quite stylized, it generates predictions that could be tested against data on how committees with collective information behave in practice. In the model, signals are indexed by committee member. The most accurate signals receive the most attention, and to the extent those signals are associated with particular members, it suggests that members of the committee who produce the most accurate signals should be the most influential (monetary policy should reflect their assessments more often and those members should dissent less.) In the context of the monetary policy example, these predictions could be verified using the individual FOMC projections introduced in Romer (2010) and the FOMC dissent time series data compiled by Thornton and Wheelock (2014).

A second positive prediction is that when the volatility of the fundamental is lower, coordination concerns become more salient and committee members tend to focus more on common public indicators as a result. This rationalizes the observation in Thornton and Wheelock (2014) that dissents were more common in the pre-Great Moderation period when inflation and unemployment volatility were higher: the decline in economic volatility made the unanimity motive more salient for FOMC members. Following the Great Recession, dissents became more common as FOMC members perceived greater uncertainty about the fundamental and relied more on their idiosyncratic signals. This proposition could be more formally tested using US macroeconomic data or macroeconomic data from other countries (and unlike the previous set of predictions about signal quality, does not rely on observing the projections of individual committee members).

Normative implications of the numerical analysis. The model ignores many salient features of the operating environment of monetary policy committees. Nevertheless, the computational experiments suggest two tentative normative lessons. First, reducing committee size may not impact the welfare of committee members, but reduces the accuracy of their projections. To the extent social welfare is more driven by whether the FOMC sets appropriate policy, the smaller committee could make policy worse. This could be taken to justify the role of regional Federal Reserve Banks in forming separate assessments of the economy, as well as institutional arrangements that ensure those views are available to the committee when making decisions. To the extent unanimity or collegiality concerns lead to less investment in accurate economic projections, mechanisms that more closely align committee member payoffs with forecast accuracy. For instance, individual projections of FOMC members are only released in an individually identifiable form ten years after they are made. There may be merit in releasing them sooner to induce reputation incentives to elicit high-quality projections. One might worry about strategic substitutability motives if forecasts were publicly identifiable. However, this risk already exists to the extent that regional Federal Reserve presidents are free to voluntarily disclose their forecasts, excepting media blackout periods near FOMC meetings. And, as Plosser (2015) argues, FOMC members could try to shift their own norms away from unanimity as its own reward.

3.6 Conclusion

In this paper, I have examined whether a stylized finite player coordination game with a unique equilibrium and symmetric information is necessary supported by the information choice of committee members. The answer is no. Conflicting direct and coordination effects in information investment mean that trying to have more precise forecasts may reduce the payoffs of committee members even without taking the costs of information investment into account. The nonconcavity of payoffs is reminiscent of, but distinct, from the results of Radner and Stiglitz (1984). However, I show in numerical examples how unique information equilibria do appear to exist under certain parameterizations, and in computational exercises explore the implications of changes in the economic environment for forecast accuracy and welfare. Crucially for the applied problem of monetary policy committee design, changes that have little impact on welfare of committee members may have larger impacts on the accuracy of the committee's beliefs.

There are a number of directions for future work. Much stronger assumptions (or different distributional assumptions) may make it somewhat easier to generate analytic results for the existence and uniqueness of equilibria (symmetric or otherwise) in the information coordination game. Once the theoretical conditions for uniqueness are more clearly established, it will be easier to modify the model to examine applied questions like optimal monetary policy committee design, or other coordination games that feature finite "informationally large" agents.

An important extension will also to be to understand the motives behind information production in this setting. Research in a committee setting is a public good and thus might be subject to inefficient underinvestment. In this paper, the quality of research associated with each committee member was exogenous. But the amount of resources devoted by committee members to producing research (either public or private) is also a choice. Once conditions for equilibrium in the attention game are understood, it would be interesting to understand how agents decide what information to bring to the meeting, and what the normative implications of those decisions are.

Appendix

3.A Derivations for coordination game

Proof of lemma 1. Given action and state monotonicity, we can characterize $\underline{\theta}, \bar{\theta}$ directly. In particular, the payoff from others' actions is maximized when $\sum_{i \neq l} a_i = L - 1$ and minimized when $\sum_{i \neq l} a_i = 0$. Hence:

$$\underline{\theta} = \sup \theta < -\left(\frac{\alpha^* + \beta^* \frac{L-1}{L}}{\lambda^*} \right)$$

and

$$\bar{\theta} = \inf \theta > \left(\frac{-\alpha^*}{\lambda^*} \right)$$

□

Proof of proposition 1. First, we use the assumptions about how others form expectations to characterize l's belief about p's beliefs. Plugging the signal expectations into (3.1) for p implies:

$$E^l[E^p(\theta)] = \frac{\sum_{n=1}^{L} \psi_{np} + \psi_{rp}}{\sum_{n=1}^{L} \psi_{np} + \psi_{rp}} E^l(\theta) = E^l(\theta)$$

and

$$\mathrm{Var}^l[E^p(\theta)] = (L+1) \left(\sum_n \psi_{nl} + \psi_{rl} \right)^{-1} + \sum_{n=1}^{L} [\psi_{np}]^{-1} + [\psi_{rp}]^{-1}$$

Under the proposed threshold rule, l believes p will take the action if her expectation of p's expectation exceeds the threshold.[10]

$$E^l[E^p(\theta)] > \theta^*$$

This happens when:

$$0 > \frac{\theta^* - E^l(\theta)}{\sqrt{(L+1)\left(\sum_n \psi_{nl} + \psi_{rl}\right)^{-1} + \sum_{n=1}^{L} [\psi_{np}]^{-1} + [\psi_{rp}]^{-1}}}$$

[10] To generalize this to a case where agents potentially have different precisions, this threshold will be specific to p. See Szkup and Trevino (2015).

The probability this occurs is

$$P(a_p = 1|\mathbf{x_l}, r_l) = 1 - \Phi\left(\frac{\theta^* - E^l(\theta)}{\sqrt{(L+1)\left(\sum_n \psi_{nl} + \psi_{rl}\right)^{-1} + \sum_{n=1}^{L}[\psi_{np}]^{-1} + [\psi_{rp}]^{-1}}}\right)$$

where Φ is the CDF of the standard normal.

If we restrict ourselves to symmetric equilibria, then we can write the expected actions of other players as:

$$P\left(\sum_n a_n = \ell\right) = \binom{L-1}{\ell}\left[\Phi\left(\frac{\theta^* - E^l(\theta)}{\sqrt{(L+1)\left(\sum_n \psi_n + \psi_r\right)^{-1} + \sum_{n=1}^{L}[\psi_n']^{-1} + [\psi_r']^{-1}}}\right)\right]^{L-1-\ell}$$
$$\left[1 - \Phi\left(\frac{\theta^* - E^l(\theta)}{\sqrt{(L+1)\left(\sum_n \psi_n + \psi_r\right)^{-1} + \sum_{n=1}^{L}[\psi_n']^{-1} + [\psi_r']^{-1}}}\right)\right]^{\ell}$$

The threshold θ^* is such that

$$\alpha^* + \frac{\beta^*}{L}\sum_{\ell=0}^{L-1} P\left(\sum_n a_n = \ell\right)\ell + \lambda^*\theta^* = 0$$

We would like to show that θ^* is unique. Note that when $E^l(\theta) = \theta^*$

$$P\left(\sum_n a_n = \ell\right) = \binom{L-1}{\ell}\left[\Phi\left(\frac{\theta^* - \theta^*}{\sqrt{(L+1)\left(\sum_n \psi_n + \psi_r\right)^{-1} + \sum_{n=1}^{L}[\psi_n']^{-1} + [\psi_r']^{-1}}}\right)\right]^{L-1-\ell}$$
$$\left[1 - \Phi\left(\frac{\theta^* - \theta^*}{\sqrt{(L+1)\left(\sum_n \psi_n + \psi_r\right)^{-1} + \sum_{n=1}^{L}[\psi_n']^{-1} + [\psi_r']^{-1}}}\right)\right]^{\ell}$$

We can simplify the binomial term:

$$E^l(\pi(\cdot,\cdot)) = \alpha^* + \frac{\beta^*}{L}(L-1)\frac{1}{2} + \lambda^*\theta^* = 0$$

Hence, the threshold is

$$\theta^* = -\frac{1}{\lambda^*}\left(\alpha^* + \frac{\beta^*}{2}\frac{(L-1)}{L}\right)$$

To prove uniqueness of the threshold, we need only show the slope of the utility parameterization is greater than zero at θ^* (otherwise, there are at least three equilibria). This always holds given our assumption that $\lambda^* > 0$.[11]

We have shown that θ^* is the cutoff signal for l given her beliefs about the state and about others' beliefs. We can also show that θ^* is the only one that survives the iterated deletion of dominated strategies. Note that the expected payoff for $a_l = 1$ when $E^l(\theta) = \theta^*$ conditional on all others using a switching strategy at some $\widehat{\theta}$ is

$$\alpha^* + \frac{\beta^*}{L}(L-1)\left(1 - \Phi\left(\frac{\widehat{\theta} - \theta^*}{\text{Var}^l(E^p(\theta))}\right)\right) + \lambda^*\theta^*$$

This is strictly increasing in θ^*, strictly decreasing in $\widehat{\theta}$, continuous, and by Lemma 1 satisfies a "limit dominance" property. Hence, from the exact argument in Appendix A of Morris and Shin (2004), the switching strategy around θ^* is the only one that survives the iterated deletion of dominated strategies. □

3.B Derivations for the information choice game

To simplify the expression for ex-ante utility, it is helpful to use the following result, which (trivially) extends Ellison (1964):

[11] In the case with a model-consistent prior, the above changes somewhat. The probability of another player's signal exceeding the threshold is

$$1 - \Phi\left(\frac{\theta^* - \frac{\sum_n \psi'_n + \psi'_r}{\frac{1}{\sigma_\theta^2} + \sum_n \psi'_n + \psi'_r}\theta^*}{\text{Var}^l(E^p(\theta))}\right)$$

θ^* implicitly solved an appropriately modified version of (3.A). Uniqueness requires

$$\frac{\lambda^* L}{\beta^*(L-1)}\left[\frac{1 - \frac{\sum_n \psi'_n + \psi'_r}{\frac{1}{\sigma_\theta^2} + \sum_n \psi'_n + \psi'_r}}{\text{Var}^l(E^p(\theta))}\right]^{-1} > \frac{1}{\sqrt{2\pi}}$$

.

Lemma 4 (extension of Ellison (1964), Theorem 2, corollary 1)**.**

$$\int_{-\infty}^{\infty} \Phi(\frac{a\theta - b}{c}) f_\Theta(\theta) d\theta = \Phi\left(\frac{-b}{a\sqrt{\frac{c^2}{a^2} + \sigma_\Theta^2}}\right)$$

Proof.

$$\int_{-\infty}^{\infty} \Phi\left(\frac{a\theta - b}{c}\right) f_\Theta(\theta) d\theta = E_\theta\left[\Phi\left(\frac{a\theta - b}{c}\right)\right] = E_\theta\left[\Phi\left(\frac{\theta - b/a}{c/a}\right)\right]$$

Say $X \sim N(b/a, (c/a)^2)$ with PDF $f_X(x)$. With some abuse of notation,

$$\begin{aligned} E_\theta\left[\Phi\left(\frac{\theta - b/a}{c/a}\right)\right] &= \int_{-\infty}^{\infty}\left[\int_{-\infty}^{\frac{\theta - b/a}{c/a}} f_X(x)dx\right] f_\Theta(\theta)d\theta \\ &= E_\theta\left[P_X\left(\frac{X - b/a}{c/a} \le \frac{\theta - b/a}{c/a}\middle|\theta\right)\right] \\ &= P_{X,\Theta}\left(\frac{X - b/a}{c/a} \le \frac{\Theta - b/a}{c/a}\right) \\ &= P(X - \Theta \le 0). \end{aligned}$$

Note that $X - \Theta \sim N(b/a, (c/a)^2 + \sigma_\theta^2)$ and the conclusion follows. $\square$

Proof of lemma 2. Recall the expression for ex-ante utility

$$\begin{aligned} E(U_l) &= E(U_l|E^l(\theta) > 0) \times P(E^l(\theta) > 0) + E(U_l|E^l(\theta) \le 0) \times P(E^l(\theta) \le 0) \\ &= E(U_l|a_l = 1, \theta + Q)P(\theta + Q > 0) + E(U_l|a_l = 0, \theta + Q)P(\theta + Q \le 0) \end{aligned}$$

Breaking up the first term:

$$\begin{aligned} &E(U|E^l(\theta) > 0) \\ &\qquad = \alpha\frac{L-1}{L} + \gamma\left(\frac{1-k}{L}\right) + \lambda E(\theta|E^l(\theta) > 0) + \nu \\ &\qquad + \frac{(\gamma - \alpha)}{L} E\left[\sum_{\ell \ne l} a_\ell | E^l(\theta) > 0\right] \end{aligned}$$

Given the symmetric distribution conjecture,

$$\sum_{\ell \ne l} a_\ell = (L-1)P(a_\ell = 1)$$

The conditional distribution of p's beliefs based on l's beliefs will be:

$$\theta + Q' | \theta + Q \sim N\left(\rho \frac{\sqrt{\sigma_\theta^2 + \sigma_{Q'}^2}}{\sqrt{\sigma_\theta^2 + \sigma_Q^2}}(\theta + Q), (\sigma_\theta^2 + \sigma_{Q'}^2)(1 - \rho^2)\right)$$

where

$$\rho = \frac{Cov(\theta + Q, \theta + Q')}{\sqrt{\sigma_\theta^2 + \sigma_{Q'}^2}\sqrt{\sigma_\theta^2 + \sigma_Q^2}} = \frac{\sigma_\theta^2 + \sum_n w_n w'_n \kappa_n^2}{\sqrt{\sigma_\theta^2 + \sigma_{Q'}^2}\sqrt{\sigma_\theta^2 + \sigma_Q^2}}$$

Call

$$\beta(z, z') = \frac{\sigma_\theta^2 + \sum_n w_n w'_n \kappa_n^2}{\sigma_\theta^2 + \sigma_Q^2}$$

Then rewriting the conditional distribution:

$$\theta + Q' | \theta + Q \sim N\left(\beta(z, z')(\theta + Q), (\sigma_\theta^2 + \sigma_{Q'}^2)(1 - \rho^2)\right)$$

Hence, the probability that another players' beliefs exceed the threshold conditional on any *particular* realization of l's posterior beliefs is

$$1 - \Phi\left(\frac{0 - \beta(z, z')(\theta + Q)}{\sqrt{(\sigma_\theta^2 + \sigma_{Q'}^2)(1 - \rho^2)}}\right) = \Phi\left(\frac{\beta(z, z')(\theta + Q)}{\sqrt{(\sigma_\theta^2 + \sigma_{Q'}^2)(1 - \rho^2)}}\right)$$

Call the PDF of θ $f_\Theta(\theta)$ and the PDF of Q $f_Q(q)$. Then we can write expected utility conditional on l receiving a positive signal as:

$$\int_{\{\theta, q: \theta + q > 0\}} \left(\alpha \frac{L - 1}{L} + \gamma\left(\frac{1 - k}{L}\right) + \nu + \lambda(\theta + q) \right.$$
$$\left. + \frac{(\gamma - \alpha)(L - 1)}{L} \Phi\left(\frac{\beta(z, z')(\theta + q)}{\sqrt{(\sigma_\theta^2 + \sigma_{Q'}^2)(1 - \rho^2)}}\right)\right) f_Q(q) f_\Theta(\theta) dq d\theta$$

This can be re-written

$$\int_{\theta=-\infty}^{\infty} \int_{q=-\theta}^{\infty} \left(\alpha \frac{L-1}{L} + \gamma \left(\frac{1-k}{L} \right) + \nu \right) f_Q(q) f_\Theta(\theta) dq d\theta \tag{3.20a}$$

$$+ \int_{\theta=-\infty}^{\infty} \int_{q=-\theta}^{\infty} \lambda \theta f_Q(q) f_\Theta(\theta) dq d\theta \tag{3.20b}$$

$$+ \frac{(\gamma-\alpha)(L-1)}{L} \int_{\theta=-\infty}^{\infty} \int_{q=-\theta}^{\infty} \Phi \left(\frac{\beta(z,z')(\theta+q)}{\sqrt{(\sigma_\theta^2 + \sigma_{Q'}^2)(1-\rho^2)}} \right) f_Q(q) f_\Theta(\theta) dq d\theta \tag{3.20c}$$

Simplifying (3.20) line by line:

(3.20a) is

$$\begin{aligned}
&\int_{\theta=-\infty}^{\infty} \int_{q=-\theta}^{\infty} \left(\alpha \frac{L-1}{L} + \gamma \left(\frac{1-k}{L} \right) + \nu \right) f_Q(q) f_\Theta(\theta) dq d\theta \\
=&\int_{\theta=-\infty}^{\infty} \left(\alpha \frac{L-1}{L} + \gamma \left(\frac{1-k}{L} \right) + \nu \right) \left(1 - \Phi \left(\frac{\theta}{\sigma_Q} \right) \right) f_\theta(\theta) d\theta \\
=&\left(\alpha \frac{L-1}{L} + \gamma \left(\frac{1-k}{L} \right) + \nu \right) \left(1 - \int_{\theta=-\infty}^{\infty} \Phi \left(\frac{\theta}{\sigma_Q} \right) f_\Theta(\theta) d\theta \right) \\
=&\frac{1}{2} \left(\alpha \frac{L-1}{L} + \gamma \left(\frac{1-k}{L} \right) + \nu \right)
\end{aligned}$$

where we've used lemma 4 at the last line.

For (3.20b):

$$\begin{aligned}
&\int_{\theta=-\infty}^{\infty} \int_{q=-\theta}^{\infty} \lambda \theta f_Q(q) f_\Theta(\theta) dq d\theta \\
=&\lambda \int_{\theta=-\infty}^{\infty} \theta \Phi \left(\frac{\theta}{\sigma_Q} \right) f_\Theta(\theta) d\theta \\
=&\lambda \frac{\sigma_\Theta^2}{\sqrt{2\pi}} \frac{1}{\sqrt{(\sigma_\Theta^2 + \sigma_Q)^2}}
\end{aligned}$$

Unfortunately, there's not a convenient closed form solution for (3.20c).[12].

Expected utility conditional on $E^l(\theta) \leq 0$ is somewhat simpler given the normalizations $\mu = 0$, and can be written:

[12]Conditional on a value for θ, the CDF is the right half of the standard normal CDF multiplied by $f_Q(q)$. Hence, it's bounded somewhere on the interval $(1/2, 1)$, but essentially there's not a nice analytic version of this. See Owen (1980)

$$\frac{1}{2}\left(\frac{\delta k}{L}+\tau\right)-\left(\frac{\beta+\delta}{L}(L-1)\right)\int_{\theta=-\infty}^{\infty}\int_{q=-\infty}^{-\theta}\Phi\left(\frac{\beta(z,z')(\theta+q)}{\sqrt{(\sigma_{\theta}^{2}+\sigma_{Q'}^{2})(1-\rho^{2})}}\right)f_{Q}(q)f_{\Theta}(\theta)dqd\theta$$

Combining all of these we arrive at an expression for expected utility in the text

$$E(U_l)=\frac{1}{2}\left(\alpha\frac{L-1}{L}+\gamma\left(\frac{1-k}{L}\right)+\nu+\frac{\delta k}{L}+\tau\right) \tag{3.21}$$

$$+\frac{\lambda}{\sqrt{2\pi}}\frac{\sigma_{\Theta}^{2}}{\sqrt{\sigma_{\Theta}^{2}+\sigma_{Q}^{2}}} \tag{3.22}$$

$$+\frac{L-1}{L}(\beta-\alpha+\gamma+\delta)\int_{\theta=-\infty}^{\infty}\int_{q=-\theta}^{\infty}\Phi\left(\frac{\beta(z,z')(\theta+q)}{\sqrt{(\sigma_{\theta}^{2}+\sigma_{Q'}^{2})(1-\rho^{2})}}\right)f_{Q}(q)f_{\Theta}(\theta)dqd\theta$$

$$\tag{3.23}$$

$$-C(z) \tag{3.24}$$

□

Derivation of expressions in remark 1. Increasing attention to a particular signal k (an increase in z_{nk}):

1. Unambiguously increases precision:

$$0<\frac{\partial\psi_{kl}}{\partial z_{kl}}=\begin{cases}\frac{\zeta_k^2}{(z_{kl}\kappa_k^2+\zeta_k^2)^2} & k\in 1,2,\ldots,L\\ \frac{1}{\zeta_r^2} & k=r\end{cases}$$

   Note that we can rewrite this as:

$$\frac{\partial\psi_{kl}}{\partial z_{kl}}=\frac{\zeta_k^2}{z_{kl}^2}\psi_{kl}^2>0$$

2. Has weakly diminishing returns on precision:

$$\frac{\partial^2 \psi_{kl}}{\partial z_{kl}^2} = \begin{cases} -\frac{\zeta_k^2 \kappa_k^2}{(z_{kl}\kappa_k^2 + \zeta_k^2)^3} & k \in 1, 2, \ldots, L \\ 0 & k = r \end{cases}$$

3. Increases the weight on the signal paid attention to and decreases the weight on other signals that receive positive attention:

$$\begin{aligned} \frac{\partial w_{kl}}{\partial z_{kl}} &= \frac{\frac{\partial \psi_{kl}}{\partial z_{kl}} (\sum_{n=1}^{L} \psi_{nl} + \psi_{rl} - \psi_{kl})}{(\sum_{n=1}^{L} \psi_{nl} + \psi_{rl})^2} \\ &= \frac{\frac{\partial \psi_{kl}}{\partial z_{kl}}}{\sum_{n=1}^{L} \psi_{nl} + \psi_{rl}} (1 - w_{kl}) = (1 - w_k) \frac{\zeta_k^2}{z_{kl}^2} \psi_{kl} w_k \quad 5 \end{aligned}$$

$$\frac{\partial w_{k'l}}{\partial z_{kl}} = -\, w_{k'l} \frac{\frac{\partial \psi_{kl}}{\partial z_{kl}}}{\sum_{n=1}^{L} \psi_{nl} + \psi_{rl}}$$

4. Causes the overall variance of noise to fall:

Proof. For $z_{kl} > 0$:

$$\begin{aligned} \frac{\partial \sigma_Q^2}{\partial z_{kl}} &= \sum_{n \neq k}^{L} (-2) \frac{\frac{\partial \psi_{kl}}{\partial z_{kl}}}{\sum_n \psi_{nl} + \psi_{rl}} \left(\kappa_n^2 + \frac{\zeta_n^2}{z_{nl}} \right) - 2 w_{rl}^2 \left(\frac{\zeta_{rl}^2}{z_{rl}} \right) \\ &\quad + 2 w_{kl} (1 - w_{kl}) \frac{\frac{\partial \psi_{kl}}{\partial z_{kl}}}{\sum_n \psi_{nl} + \psi_{rl}} \left(\kappa_k^2 + \frac{\zeta_k^2}{z_{nl}} \right) - w_{kl}^2 \frac{\zeta_{kl}^2}{z_{kl}^2} \\ &= \frac{-2}{\sum_{n=1}^{L} \psi_{nl} + \psi_{rl}} \frac{\partial \psi_{kl}}{\partial z_{kl}} \sigma_Q^2 + \frac{2 w_k}{\sum_{n=1}^{L} \psi_{nl} + \psi_{rl}} \frac{\partial \psi_{kl}}{\partial z_{kl}} \psi_k^{-1} - w_{kl}^2 \frac{\zeta_{kl}^2}{z_{kl}^2} \\ &= \frac{-2\sigma_Q^2}{\sum_{n=1}^{L} \psi_{nl} + \psi_{rl}} \frac{\partial \psi_{kl}}{\partial z_{kl}} + 2 w_k^2 \frac{\zeta_{kl}^2}{z_{kl}^2} - w_{kl}^2 \frac{\zeta_{kl}^2}{z_{kl}^2} \\ &= \frac{-2\sigma_Q^2}{\sum_{n=1}^{L} \psi_{nl} + \psi_{rl}} \frac{\partial \psi_{kl}}{\partial z_{kl}} + w_k^2 \frac{\zeta_{kl}^2}{z_{kl}^2} < 0 \end{aligned}$$

The inequality follows from substituting for $\frac{\partial \psi_{kl}}{\partial z_{kl}}$ and using the fact that the weights all sum to 1.

$\square$

5. For a particular q, $\frac{\partial f_Q(q)}{\partial z_{kl}} = f_Q(q) \left(\frac{q^2}{\sigma_Q^2} - 1 \right) \frac{\partial \sigma_Q^2}{\partial z_{kl}}$

6. Has ambiguous effect on $\beta(z, z')$

Proof. Note that $\beta(z, z')$ is bounded between 0 and 1.

$$
\begin{aligned}
\frac{\partial \beta(z,z')}{\partial z_{kl}} =& \frac{\sum_n \frac{\partial w_{nl}}{\partial z_{kl}} w'_k \kappa_k^2 (\sigma_\theta^2 + \sigma_Q^2) - (\sigma_\theta^2 + (\sum_{n=1}^L w_n w'_n \kappa_n^2)) \frac{\partial \sigma_Q^2}{\partial z_{kl}}}{(\sigma_\theta^2 + \sigma_Q^2)^2} \\
=& \frac{\frac{\partial \psi_{kl}}{\partial z_{kl}}}{\sum_n \psi_{nl} + \psi_{rl}} \left[\frac{-\sum_n w_n w'_n \kappa_n^2 + w'_k \kappa_k^2}{\sigma_\theta^2 + \sigma_Q^2} \right] - \frac{\partial \sigma_Q^2}{\partial z_{kl}} \frac{\beta(z,z')}{\sigma_\theta^2 + \sigma_Q^2} \\
=& \frac{\frac{\partial \psi_{kl}}{\partial z_{kl}}}{\sum_n \psi_{nl} + \psi_{rl}} \left[\frac{\sigma_\Theta^2}{\sigma_\theta^2 + \sigma_Q^2} - \beta(z,z') \right] + \frac{\frac{\partial \psi_{kl}}{\partial z_{kl}}}{\sum_n \psi_{nl} + \psi_{rl}} \frac{w'_k \kappa_k^2}{\sigma_\theta^2 + \sigma_Q^2} \\
&+ \frac{2\beta(z,z')\sigma^2}{\sigma_\theta^2 + \sigma_Q^2} \frac{\frac{\partial \psi_{kl}}{\partial z_{kl}}}{\sum_n \psi_{nl} + \psi_{rl}} - w_k^2 \frac{\zeta_{kl}^2}{z_{kl}^2} \frac{1}{\sigma_\theta^2 + \sigma_Q^2} \\
=& \frac{\frac{\partial \psi_{kl}}{\partial z_{kl}}}{\sum_n \psi_{nl} + \psi_{rl}} \left[\frac{\sigma_\theta^2 (1 - \beta(z,z')) + \beta \sigma_Q^2)}{\sigma_\theta^2 + \sigma_Q^2} \right] + \frac{\zeta_{kl}^2}{z_{kl}^2} \frac{w_k w'_k \kappa_k^2}{\sigma_Q^2 + \sigma_\theta^2} - \frac{\zeta_{kl}^2}{z_{kl}^2} \frac{w_k^2}{\sigma_Q^2 + \sigma_\theta^2} \\
=& \frac{\zeta_{kl}^2}{z_{kl}^2} \frac{\psi_{kl} w_k}{\sigma_Q^2 + \sigma_\theta^2} \left\{ \sigma_\theta^2 (1 - \beta(z,z')) + \sigma_Q^2 (\beta(z,z')) + (w'_k - w_k)\kappa_k^2 - w_k \frac{\zeta_k^2}{z_k l} \right\}
\end{aligned}
$$

Clearly this can be zero if $w_k = 0$. The terms involving $\beta(\cdot)$ are positive, the second falls in the interval from ranges from $[-\kappa_k^2, \kappa_k]$ and the third is unambiguously negative if $w_k > 0$. The sign depends on magnitudes of the variance of the noise in signal k, its relative precision, and β, and the variance of the fundamental.

The intuition is that paying to recieve a more accurate signal means their beliefs will be more like yours to the extent that you have a more precise idea what their beliefs are centered around (i.e., the forecast of θ becomes more precise). (In other words, the denominator of $\beta(z, z')$ unambiguously decreases). But it may cause you to pay closer attention to (and thus be exposed to idiosyncratic noise from) signals that others are not paying attention to. Hence the net effect is ambiguous.

□

7. Has an ambiguous effect on ρ:

Proof. Note ρ can be written as $\beta \frac{\sqrt{\sigma_\theta^2 + \sigma_Q^2}}{\sqrt{\sigma_\theta^2 + \sigma_{Q'}^2}}$

$$\frac{\partial \rho}{\partial z_{kl}} = \frac{1}{\sqrt{\sigma_\theta^2 + \sigma_{Q'}^2}} \left[\frac{\partial \beta(z,z')}{\partial z_{kl}} \sqrt{\sigma_\theta^2 + \sigma_Q^2} + \frac{\frac{1}{2}\beta(z,z')\frac{\partial \sigma_Q^2}{\partial z_{kl}}}{\sqrt{\sigma_Q^2 + \sigma_\Theta^2}} \right]$$

□

The first term in square brackets has an ambiguous sign, while the second is unambiguously non-positive.

□

3.C Figures

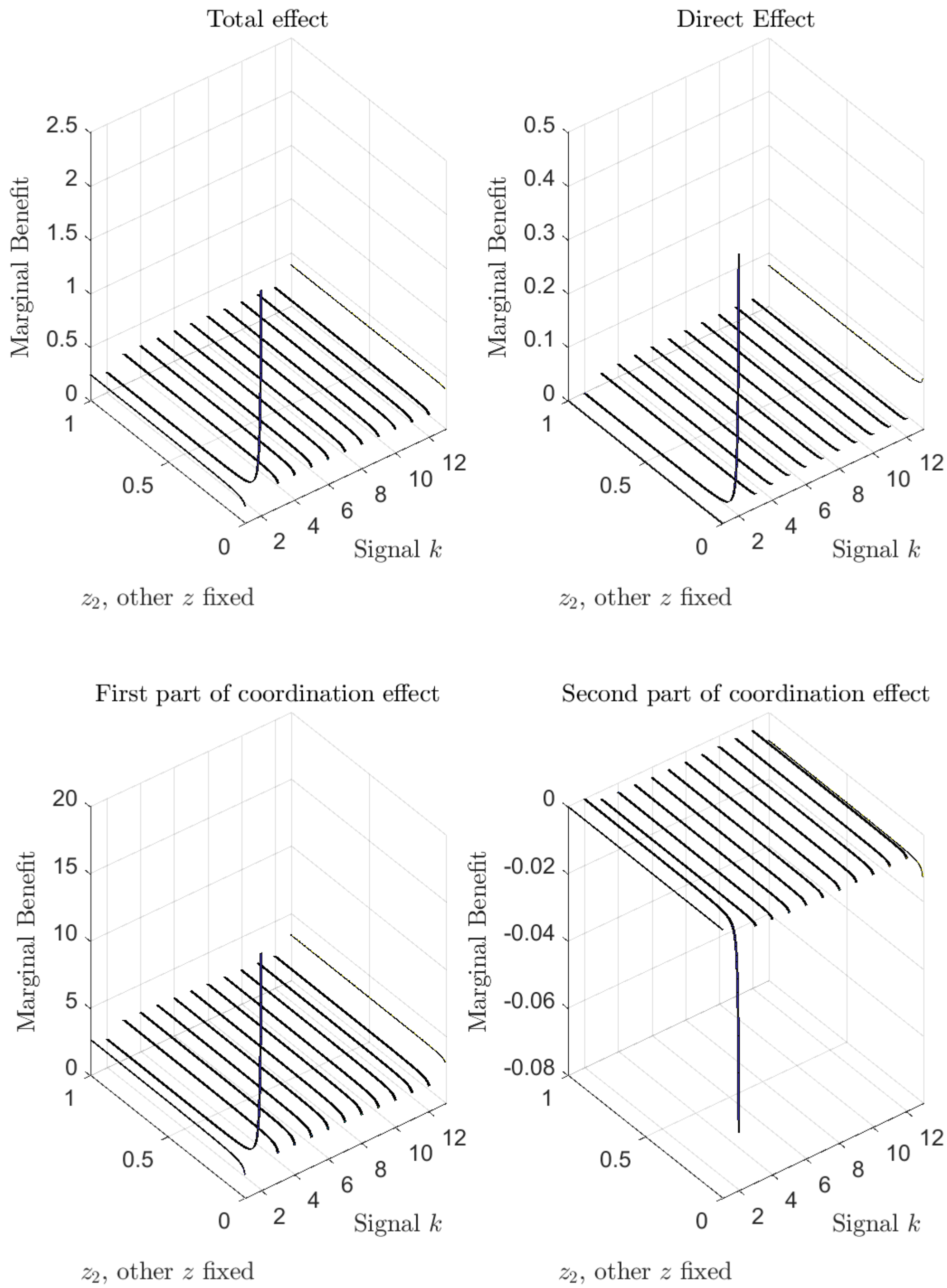


Figure 3.1: Effect of increasing z_{2l} on marginal utilities associated with increases in z_{kl}, decomposed.

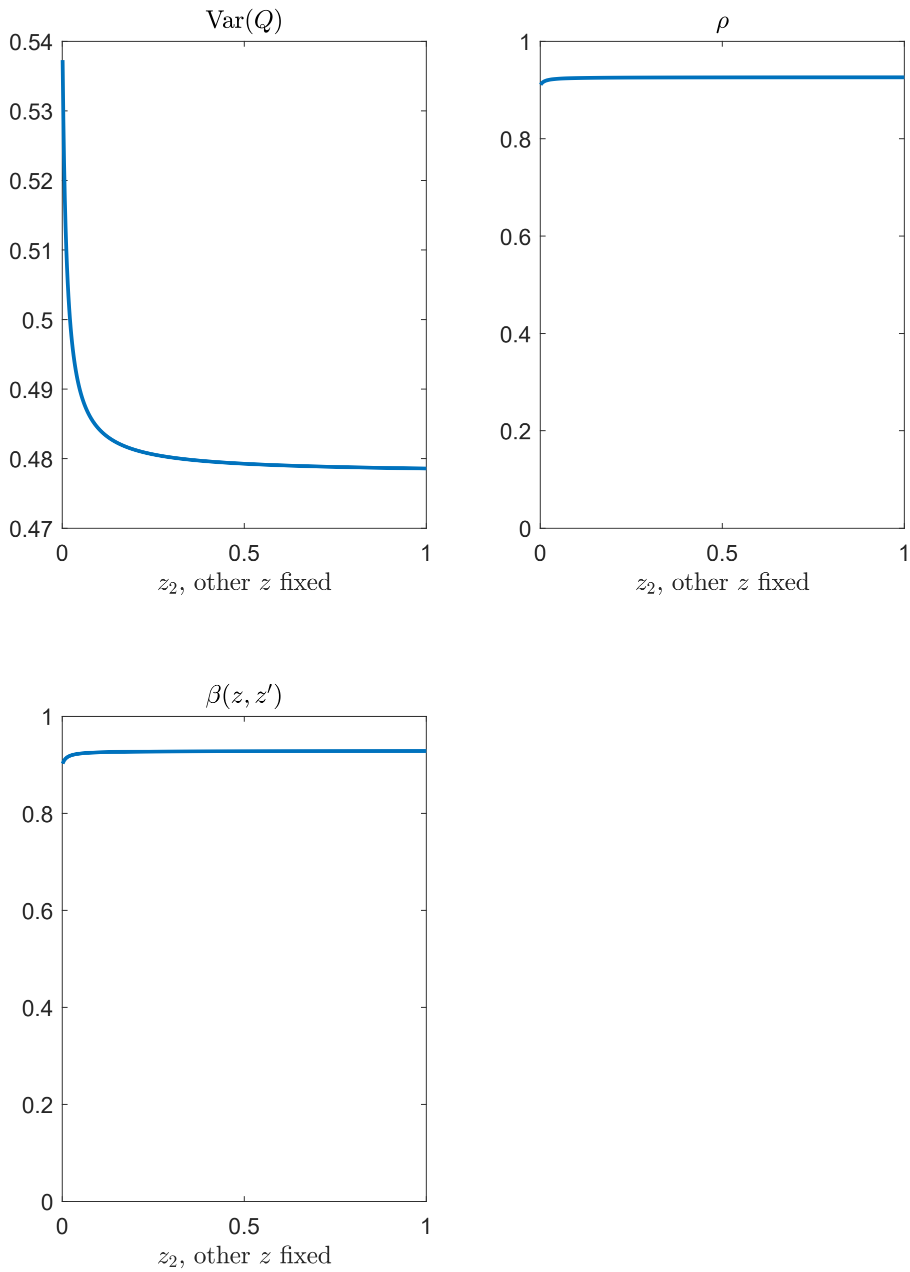


Figure 3.2: Effect of increasing z_{2l} on noise variance, ρ and $\beta(z, z')$.

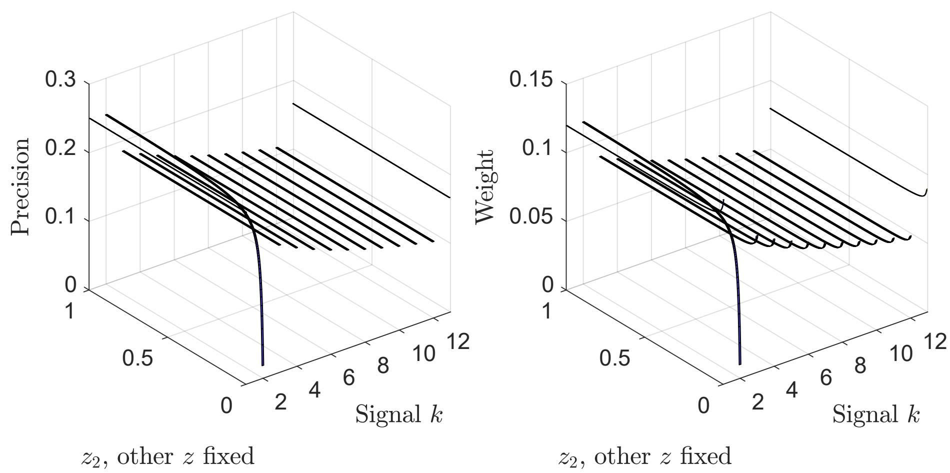


Figure 3.3: Effect of increasing z_{2l} on precisions and signal weights.

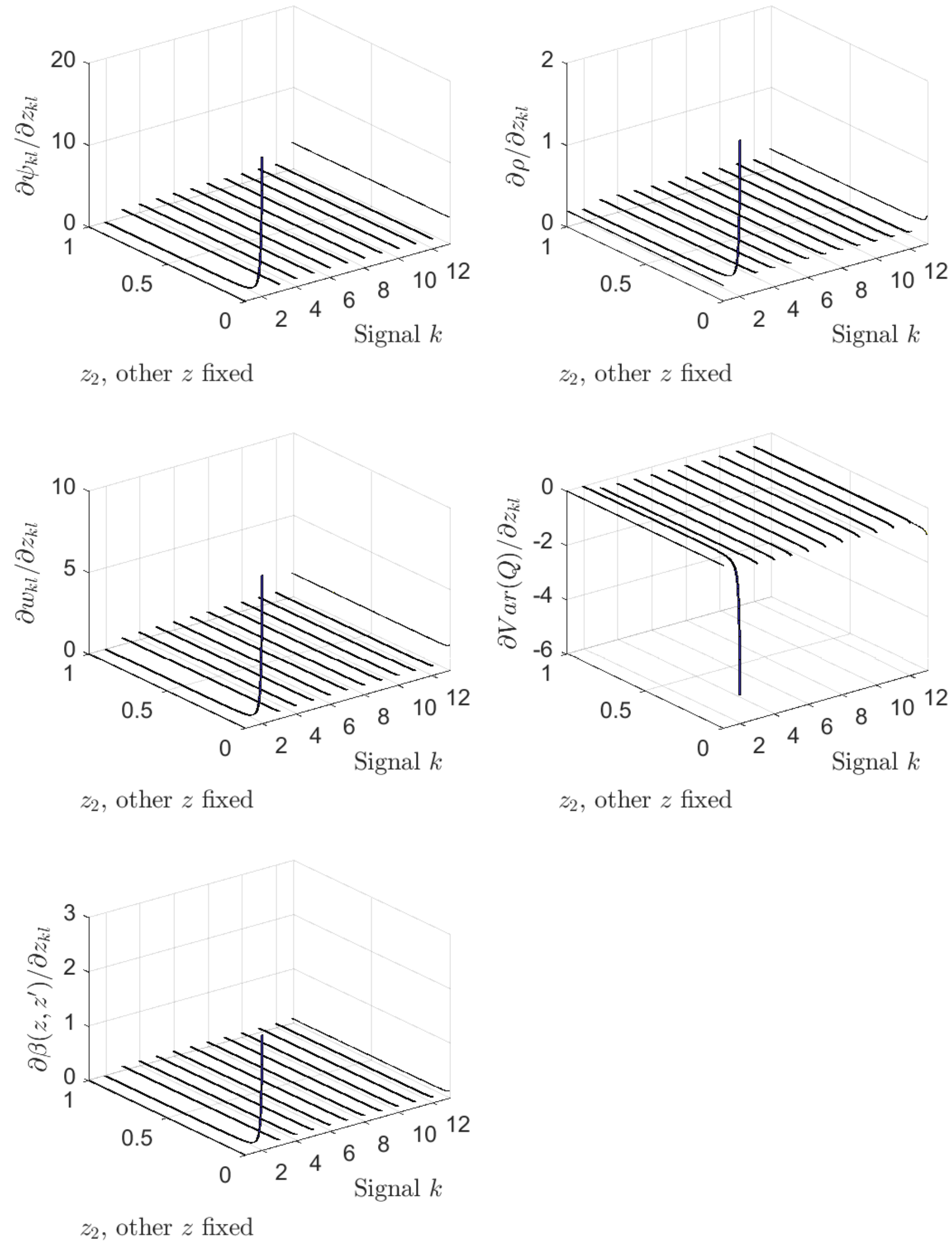


Figure 3.4: Marginal effects of changes in z_l for different values of z_{2l}.

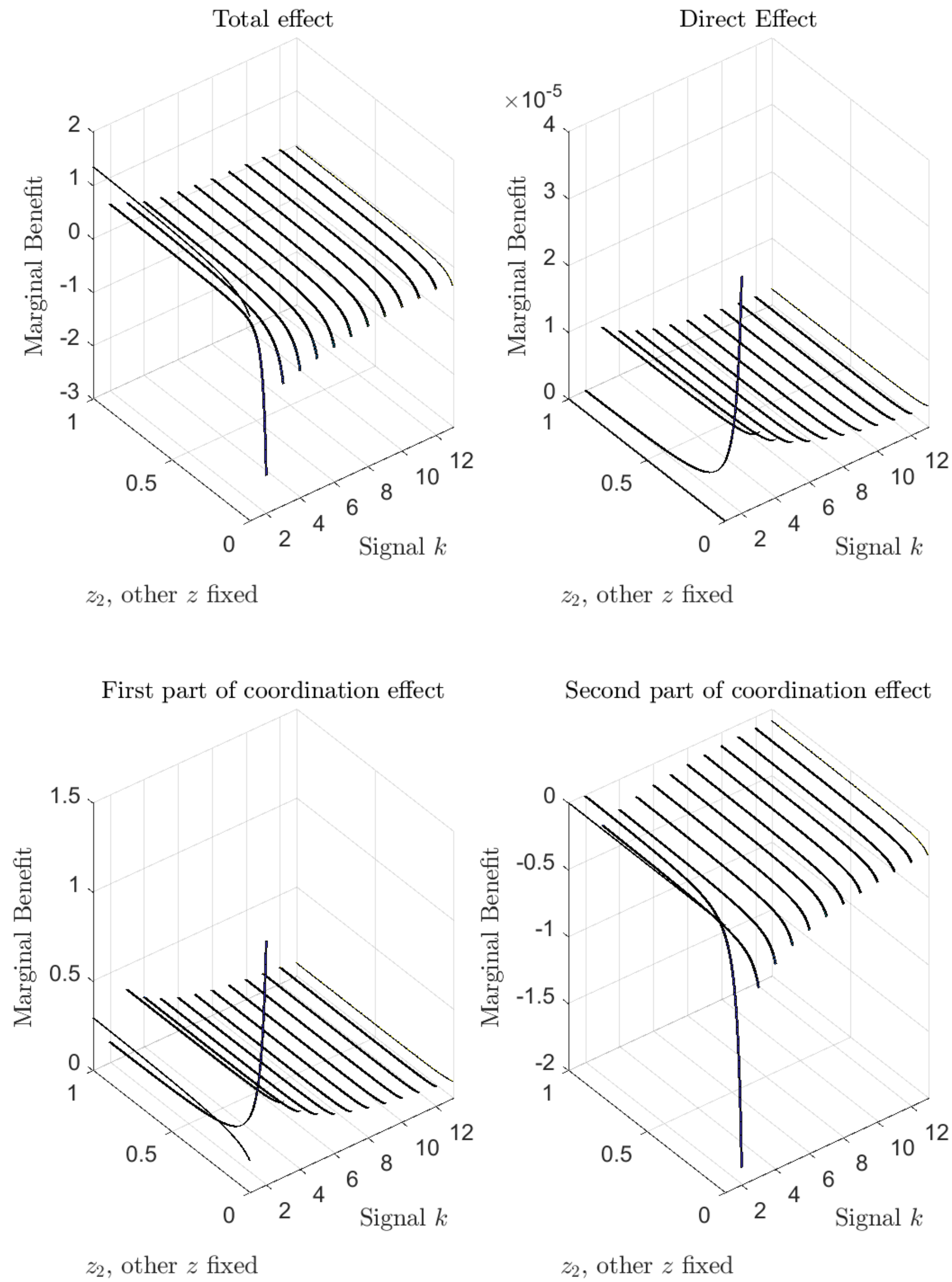


Figure 3.5: Effect of increasing z_{2l} on marginal utilities associated with increases in z_{kl} with $\sigma_\Theta^2 = 1$ $\lambda = 0.0001$, $\beta^* = 5$.

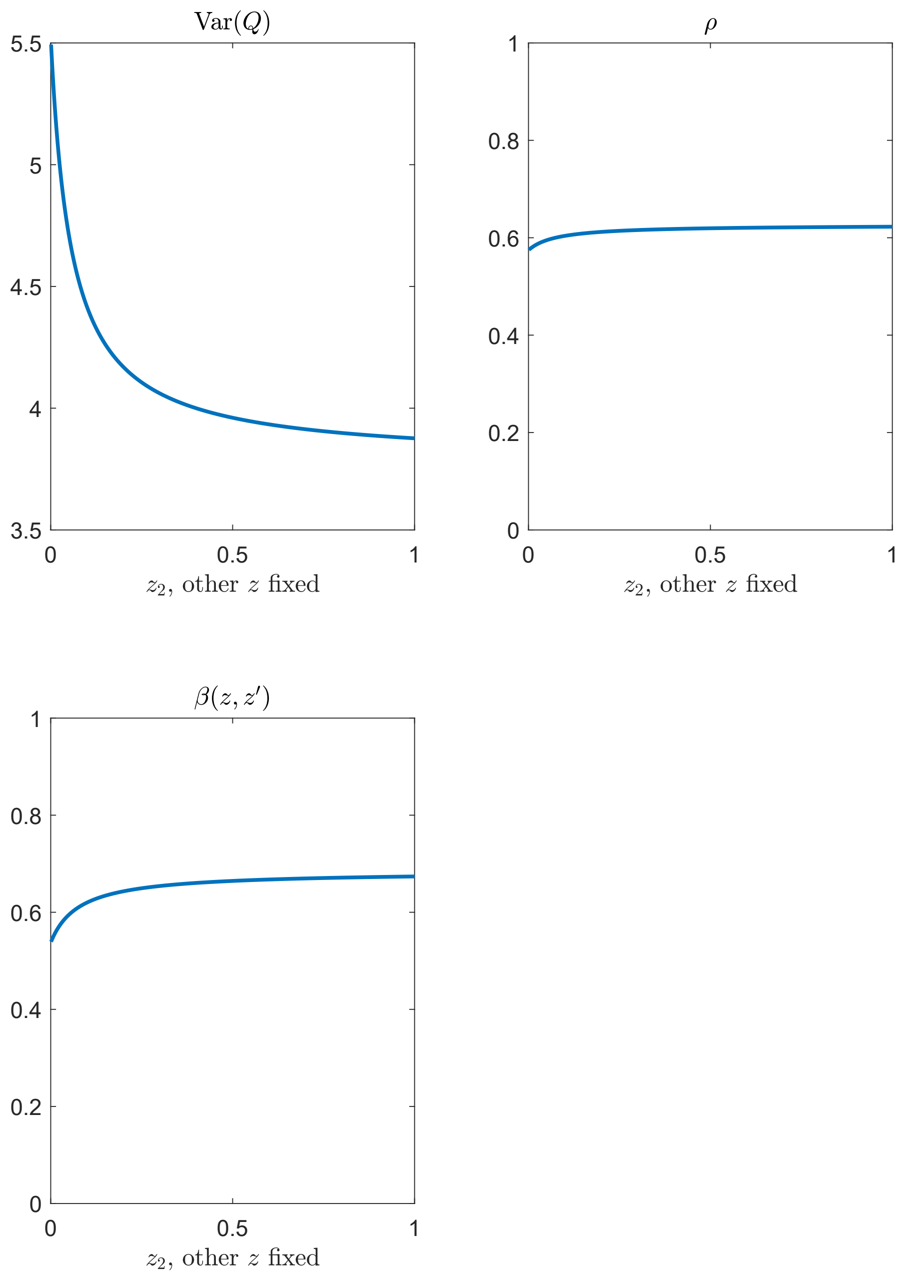


Figure 3.6: Effect of increasing z_{2l} on noise variance, ρ and $\beta(z, z')$, with $\sigma^2_\Theta = 1$ $\lambda = 0.0001$, $\beta^* = 5$.

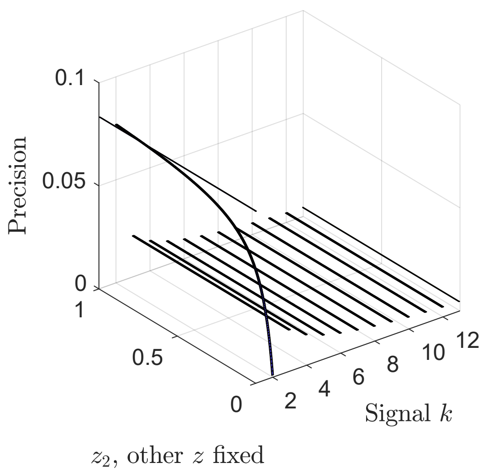


Figure 3.7: Effect of increasing z_{2l} on precisions and signal weights, with $\sigma^2_\Theta = 1$ $\lambda = 0.0001$, $\beta^* = 5$.

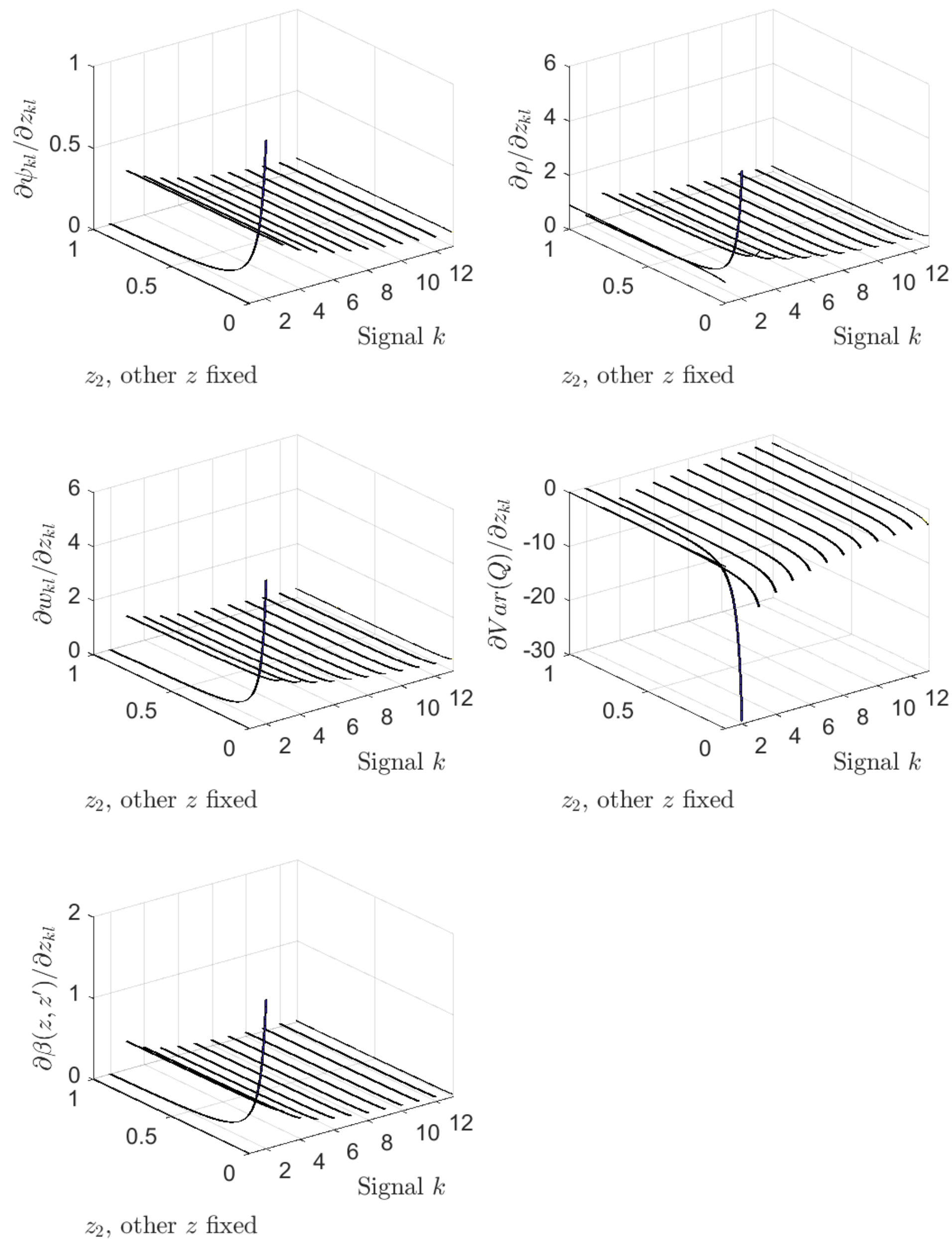


Figure 3.8: Marginal effects of changes in z_l for different values of z_{2l}, with $\sigma_\Theta^2 = 1$ $\lambda = 0.0001$, $\beta^* = 5$.

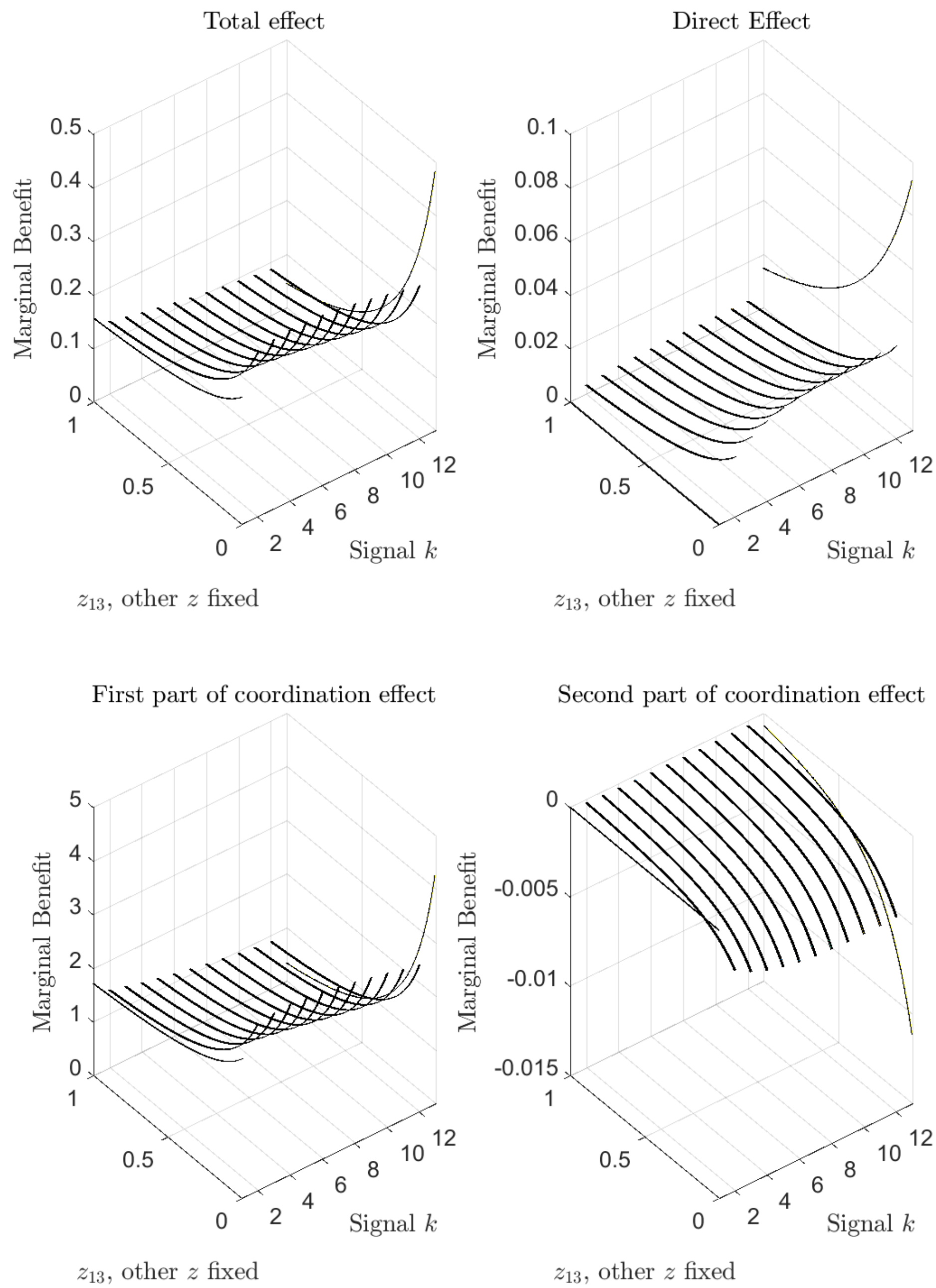


Figure 3.9: Effect of increasing z_r on marginal utilities associated with increases in z_{kl}, decomposed.

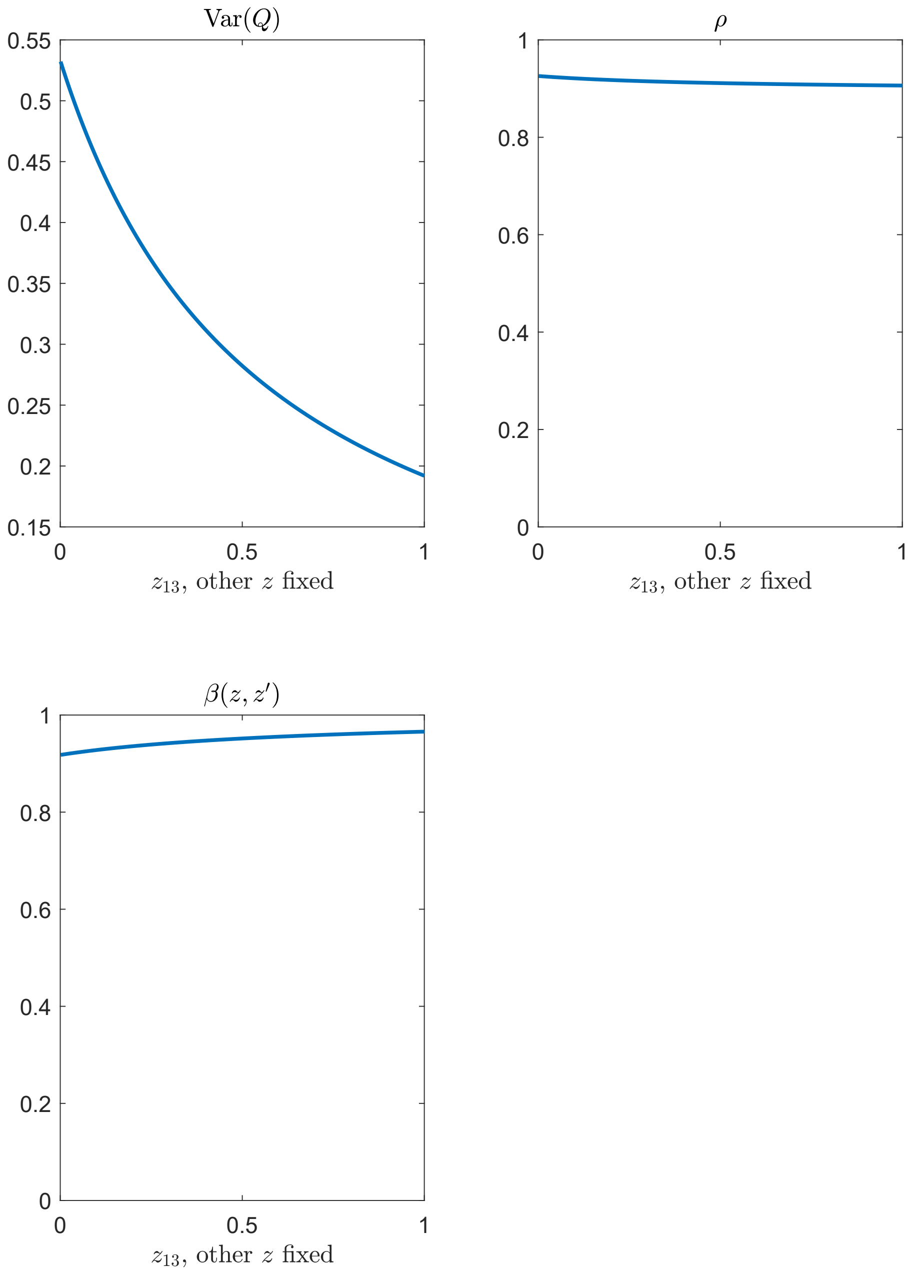


Figure 3.10: Effect of increasing z_r on noise variance, ρ and $\beta(z, z')$.

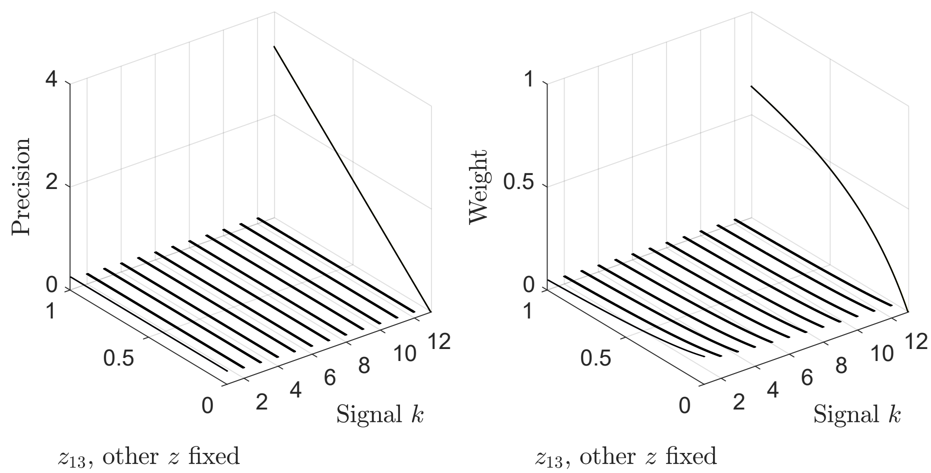


Figure 3.11: Effect of increasing z_r on precisions and signal weights.

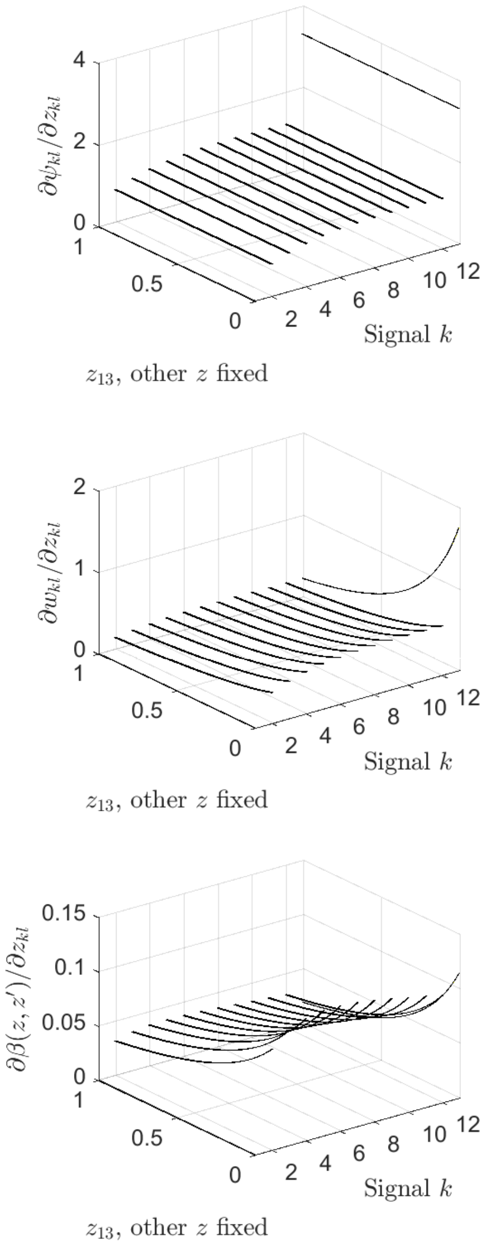


Figure 3.12: Marginal effects of changes in z for different values of z_{rl}.

Bibliography

Admati, Anat R (1985) "A Noisy Rational Expectations Equilibrium for Multi-asset Securities Markets," *Econometrica*, Vol. 53, pp. 629–57.

Allen, Franklin, Stephen Morris, and Hyun Song Shin (2006) "Beauty Contests and Iterated Expectations in Asset Markets," *Review of Financial Studies*, Vol. 19, pp. 719–752.

Andrade, Philippe, Richard K. Crump, Stefano Eusepi, and Emanuel Moench (2014) "Fundamental Disagreement," Staff Reports 655, Federal Reserve Bank of New York.

Ang, Andrew, Sen Dong, and Monika Piazzesi (2007) "No-Arbitrage Taylor Rules," NBER Working Papers 13448, National Bureau of Economic Research.

Angeletos, George-Marios and Chen Lian (2016) "Incomplete Information in Macroeconomics: Accommodating Frictions in Coordination," NBER Working Papers 22297, National Bureau of Economic Research, Inc.

Atolia, Manoj and Ryan Chahrour (2013) "Intersectoral Linkages, Diverse Information, and Aggregate Dynamics in a Neoclassical Model," Working Paper, Boston College Department of Economics.

Aumann, Robert J. (1976) "Agreeing to Disagree," *The Annals of Statistics*, Vol. 4, pp. 1236–1239.

Bacchetta, Philippe and Eric Van Wincoop (2006) "Can Information Heterogeneity Explain the Exchange Rate Determination Puzzle?," *American Economic Review*, Vol. 96, pp. 552–576.

——— (2008) “Higher Order Expectations in Asset Pricing,” *Journal of Money, Credit and Banking*, Vol. 40, pp. 837–866.

Barillas, Francisco and Kristoffer Nimark (2015) “Speculation and the Bond Market: An Empirical No-arbitrage Framework,” Economics Working Papers 1337, Department of Economics and Business, Universitat Pompeu Fabra.

——— (2016) “Speculation and the Term Structure of Interest Rates,” Working Papers 430, Barcelona Graduate School of Economics.

Barsky, Robert B. (2009) “The Japanese Bubble: A ’Heterogeneous’ Approach,” NBER Working Papers 15052, National Bureau of Economic Research, Inc.

Bauer, Michael D. (2016) “Restrictions on Risk Prices in Dynamic Term Structure Models,” *Journal of Business & Economic Statistics*.

Blinder, Alan S. (2009) “Making Monetary Policy by Committee,” *International Finance*, Vol. 12, pp. 171–194.

Boehm, Christoph E. and Christopher L. House (2014) “Optimal Taylor Rules in New Keynesian Models,” NBER Working Papers 20237, National Bureau of Economic Research, Inc.

Campbell, John Y. and Robert J. Shiller (1991) “Yield Spreads and Interest Rate Movements: A Bird’s Eye View,” *Review of Economic Studies*, Vol. 58, pp. 495–514.

Cao, Dan (2011) “Collateral Shortages, Asset Price and Investment Volatility with Heterogeneous Beliefs,” working papers, Georgetown University, Department of Economics.

Carlsson, Hans and Eric van Damme (1993a) “Equilibrium selection in stag hunt games,” Frontiers of Game Theory: MIT Press.

——— (1993b) “Global Games and Equilibrium Selection,” *Econometrica*, Vol. 61, pp. 989–1018.

Chahrour, Ryan (2014) "Public Communication and Information Acquisition," *American Economic Journal: Macroeconomics*, Vol. 6, pp. 73–101.

Chahrour, Ryan and Robert Ulbricht (2017) "Information-driven Business Cycles: A Primal Approach," Boston College Working Papers in Economics 925, Boston College Department of Economics.

Cochrane, John H. (2005) *Asset Pricing*: Princeton University Press.

Cochrane, John H. and Monika Piazzesi (2005) "Bond Risk Premia," *American Economic Review*, Vol. 95, pp. 138–160.

Cochrane, John and Monika Piazzesi (2008) "Decomposing the Yield Curve," 2009 Meeting Papers 18, Society for Economic Dynamics.

Coibion, Olivier and Yuriy Gorodnichenko (2012a) "What Can Survey Forecasts Tell Us about Information Rigidities?," *Journal of Political Economy*, Vol. 120, pp. 116 – 159.

——— (2012b) "What Can Survey Forecasts Tell Us about Information Rigidities?," *Journal of Political Economy*, Vol. 120, pp. 116 – 159.

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

Colacito, Riccardo, Eric Ghysels, Jinghan Meng, and Wasin Siwasarit (2016) "Skewness in Expected Macro Fundamentals and the Predictability of Equity Returns: Evidence and Theory," *Review of Financial Studies*, Vol. 29, pp. 2069–2109.

Collin-Dufresne, Pierre, Michael Johannes, and Lars A. Lochstoer (2016) "Parameter Learning in General Equilibrium: The Asset Pricing Implications," *American Economic Review*, Vol. 106, pp. 664–98.

Cover, Thomas M. and Joy A. Thomas (2006) *Elements of Information Theory*: Wiley-Interscience, 2nd edition.

Crump, Richard K., Stefano Eusepi, and Emanuel Moench (2016) "The term structure of expectations and bond yields," Staff Reports 775, Federal Reserve Bank of New York.

Dewachter, Hans, Leonardo Iania, and Marco Lyrio (2014) "Information in the Yield Curve: A Macro-Finance Approach," *Journal of Applied Econometrics*, Vol. 29, pp. 42–64.

Dewachter, Hans and Marco Lyrio (2008) "Learning, Macroeconomic Dynamics and the Term Structure of Interest Rates," in John Y. Campbell ed. *Asset Prices and Monetary Policy*: National Bureau of Economic Research.

Doh, Taeyoung (2012) "What Does the Yield Curve Tell Us about the Federal ReserveâĂŹs Implicit Inflation Target?" *Journal of Money, Credit and Banking*, Vol. 44, pp. 469–486.

Duffee, Gregory R. (2002) "Term Premia and Interest Rate Forecasts in Affine Models," *Journal of Finance*, Vol. 57, pp. 405–443.

Ellison, Bob E. (1964) "Two theorems for inferences about the normal distribution with applications in acceptance sampling," *Journal of the American Statistical Association*, Vol. 59, pp. 89–95.

Faust, Jon (1996) "Whom can we trust to run the Fed? Theoretical support for the founders' views," *Journal of Monetary Economics*, Vol. 37.

Faust, Jon and Jonathan H. Wright (2013) "Forecasting Inflation," in *Handbook of Economic Forecasting*, Vol. 2: Elsevier, Chap. Chapter 1, pp. 2–56.

Fernald, John (2012) "A quarterly, utilization-adjusted series on total factor productivity," Working Paper Series 2012-19, Federal Reserve Bank of San Francisco.

Giacoletti, Marco, Kristoffer T. Laursen, and Kenneth J. Singleton (2015) "Learning, Dispersion of

Beliefs, and Risk Premiums in an Arbitrage-free Term Structure Model," working paper, Graduate School of Business, Stanford University.

Graham, Liam and Stephen Wright (2010) "Information, heterogeneity and market incompleteness," *Journal of Monetary Economics*, Vol. 57, pp. 164–174.

Grossman, Sanford J and Joseph E Stiglitz (1980) "On the Impossibility of Informationally Efficient Markets," *American Economic Review*, Vol. 70, pp. 393–408.

Gurkaynak, Refet S., Brian Sack, and Eric Swanson (2005) "The Sensitivity of Long-Term Interest Rates to Economic News: Evidence and Implications for Macroeconomic Models," *American Economic Review*, Vol. 95, pp. 425–436.

Gurkaynak, Refet S., Brian Sack, and Jonathan H. Wright (2007) "The U.S. Treasury yield curve: 1961 to the present," *Journal of Monetary Economics*, Vol. 54, pp. 2291–2304.

Hamilton, James D. (1994a) *Time Series Analysis*: Princeton University Press.

——— (1994b) *Time Series Analysis*: Princeton University Press.

——— (2017) "Why You Should Never Use the Hodrick-Prescott Filter," Working Papers, University of California - San Diego Dept. of Economics.

Harrison, J. Michael and David M. Kreps (1978) "Speculative Investor Behavior in a Stock Market with Heterogeneous Expectations," *The Quarterly Journal of Economics*, Vol. 92, pp. 323–336.

Harsanyi, John C. (1968) "Games with Incomplete Information Played by 'Bayesian' Players, Part III. The Basic Probability Distribution of the Game," *Management Science*, Vol. 14, pp. 486–502.

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

Hassan, Tarek A. and Thomas M. Mertens (2014) "Information Aggregation in a DSGE Model," NBER Working Papers 20193, National Bureau of Economic Research, Inc.

Hellwig, Christian and Laura Veldkamp (2009) “Knowing What Others Know: Coordination Motives in Information Acquisition,” *Review of Economic Studies*, Vol. 76, pp. 223–251.

Hellwig, Martin F. (1980) “On the aggregation of information in competitive markets,” *Journal of Economic Theory*, Vol. 22, pp. 477–498.

Hilscher, Jens, Alon Raviv, and Ricardo Reis (2014) “Inflating Away the Public Debt? An Empirical Assessment,” NBER Working Papers 20339, National Bureau of Economic Research, Inc.

Hubert, Paul (2014) “Revisiting the greenbook’s relative forecasting performance,” Sciences Po publications info:hdl:2441/35kgubh40v9, Sciences Po.

Huo, Zhen and Naoki Takayama (2014) “Rational Expectations Models with Higher Order Beliefs,” Unpublished Job Market Paper, University of Minnesota, November.

Ireland, Peter N. (2015) “Monetary policy, bond risk premia, and the economy,” *Journal of Monetary Economics*, Vol. 76, pp. 124–140.

Jermann, Urban J. (1998) “Asset pricing in production economies,” *Journal of Monetary Economics*, Vol. 41, pp. 257–275.

Kasa, Kenneth, Todd B. Walker, and Charles H. Whiteman (2014) “Heterogeneous Beliefs and Tests of Present Value Models,” *Review of Economic Studies*, Vol. 81, pp. 1137–1163.

Kohlhas, Alexandre (2015) “Learning-by-Sharing: Monetary Policy and the Information Content of Public Signals,” 2015 Meeting Papers 57, Society for Economic Dynamics.

Kohlhas, Alexandre N. (2014) “The Informativeness of Prices: Dispersed Asymmetric Information and Higher-Order Beliefs,” Working Paper.

Li, Hao and Wing Suen (2009) “Viewpoint: Decision-making in committees,” *Canadian Journal of Economics*, Vol. 42, pp. 359–392.

Makarov, Igor and Oleg Rytchkov (2012) "Forecasting the forecasts of others: Implications for asset pricing," *Journal of Economic Theory*, Vol. 147, pp. 941–966.

Malmendier, Ulrike, Stefan Nagel, and Zhen Yan (2017) "The Making of Hawks and Doves: Inflation Experiences on the FOMC," NBER Working Papers 23228, National Bureau of Economic Research, Inc.

Mankiw, N. Gregory and Ricardo Reis (2006) "Pervasive Stickiness," *American Economic Review*, Vol. 96, pp. 164–169.

Mankiw, N. Gregory, Ricardo Reis, and Justin Wolfers (2004) "Disagreement about Inflation Expectations," in *NBER Macroeconomics Annual 2003, Volume 18*: National Bureau of Economic Research, Inc, pp. 209–270.

McCallum, Bennett (2005) "Monetary policy and the term structure of interest rates," *Economic Quarterly*, pp. 1–21.

Melosi, Leonardo (2014) "Estimating Models with Dispersed Information," *American Economic Journal: Macroeconomics*, Vol. 6, pp. 1–31.

——— (2017) "Signalling Effects of Monetary Policy," *The Review of Economic Studies*, Vol. 84, p. 853.

Miao, Jianjun, Jieran Wu, and Eric Young (2017) "Macro-Financial Volatility under Dispersed Information," working paper, Boston University.

Mondria, Jordi and Climent Quintana-Domeque (2012) "Financial contagion and Attention Allocation," *The Economic Journal*, Vol. 123.

Morris, Stephen and Hyun Song Shin (2002) "Social Value of Public Information," *American Economic Review*, Vol. 92, pp. 1521–1534.

——— (2003) "Global Games: Theory and Applications," Advances in Economics and Econometrics (Proceedings of the Eighth World Congress of the Econometric Society): Cambridge University Press.

——— (2004) "Coordination risk and the price of debt," *European Economic Review*, Vol. 48, pp. 133–153.

Morris, Stephen, Hyun Song Shin, and Muhamet Yildiz (2016) "Common belief foundations of global games," *Journal of Economic Theory*, Vol. 163, pp. 826–848.

Morris, Stephen and Ming Yang (2016) "Coordination and Continuous Choice," working papers, Princeton University, Department of Economics, Econometric Research Program.

Myatt, David P. and Chris Wallace (2008) "On the Sources and Value of Information: Public Announcements and Macroeconomic Performance," Economics Series Working Papers 411, University of Oxford, Department of Economics.

——— (2012) "Endogenous Information Acquisition in Coordination Games," *Review of Economic Studies*, Vol. 79, pp. 340–374.

——— (2014) "Central bank communication design in a Lucas-Phelps economy," *Journal of Monetary Economics*, Vol. 63, pp. 64–79.

Nakamura, Emi and Jon Steinsson (2013) "High Frequency Identification of Monetary Non-Neutrality: The Information Effect," NBER Working Papers 19260, National Bureau of Economic Research, Inc.

Nimark, Kristoffer (2007) "Dynamic higher order expectations," Economics Working Papers 1118, Department of Economics and Business, Universitat Pompeu Fabra.

——— (2008) "Dynamic pricing and imperfect common knowledge," *Journal of Monetary Economics*, Vol. 55, pp. 365–382.

Orphanides, Athanasios and Simon van Norden (2002) “The Unreliability of Output-Gap Estimates in Real Time,” *The Review of Economics and Statistics*, Vol. 84, pp. 569–583.

Orphanides, Athanasios and John C. Williams (2009) “Imperfect Knowledge and the Pitfalls of Optimal Control Monetary Policy,” in Klaus Schmidt-Hebbel, Carl E. Walsh, Norman Loayza (Series Editor), and Klaus Schmidt-Hebbel (Series Editor) eds. *Monetary Policy under Uncertainty and Learning*, Vol. 13: Central Bank of Chile, 1st edition, Chap. 4, pp. 115–144.

Owen, D.B. (1980) “A table of normal integrals,” *Communications in Statistics*, Vol. 9, pp. 389 – 419.

Patton, Andrew J. and Allan Timmerman (2010) “Why do forecasters disagree? Lesson from the term structure of cross-sectional dispersion,” *Journal of Monetary Economics*, Vol. 57, pp. 803–820.

Piazzesi, Monika (2009) “Affine Term Structure Models,” in *Handbook of Financial Econometrics*, Vol. I: Elsevier, Chap. 12, pp. 691–766.

Piazzesi, Monika, Juliana Salomao, and Martin Schneider (2013) “Trend and Cycle in Bond Premia,” working paper, Stanford University.

Piazzesi, Monika and Martin Schneider (2007) “Equilibrium Yield Curves,” in *NBER Macroeconomics Annual 2006, Volume 21*: National Bureau of Economic Research, Inc, pp. 389–472.

Plosser, Charles I. (2015) “The Veneer of Consensus at the Fed,” *The Wall Street Journal*.

Prat, Andrea, Joan de MartÃŋ, and Antoni CalvÃş-Armengol (2015) “Communication and influence,” *Theoretical Economics*, Vol. 10.

Radner, Roy and Joseph Stiglitz (1984) “A nonconcavity in the value of information,” in M. Boyer and R.E. Kihlstrom eds. *Bayesian Models of Economic Theory*.

Romer, David (2010) "A New Data Set on Monetary Policy: The Economic Forecasts of Individual Members of the FOMC," *Journal of Money, Credit and Banking*, Vol. 42, pp. 951–957.

Romer, David H. and Christina D. Romer (2000) "Federal Reserve Information and the Behavior of Interest Rates," *American Economic Review*, Vol. 90, pp. 429–457.

Rudebusch, Glenn D. and Tao Wu (2008) "A Macro-Finance Model of the Term Structure, Monetary Policy and the Economy*," *The Economic Journal*, Vol. 118, pp. 906–926.

Sibert, Anne (2003) "Monetary Policy Committees: Individual and Collective Reputations," *Review of Economic Studies*, Vol. 70, pp. 649–665.

Simon, Dan (2006) *Optimal State Estimation: Kalman, H Infinity, and Nonlinear Approaches*: Wiley-Interscience.

Sims, Christopher A. (2002) "The Role of Models and Probabilities in the Monetary Policy Process," *Brookings Papers on Economic Activity*, Vol. 33, pp. 1–62.

——— (2003) "Implications of rational inattention," *Journal of Monetary Economics*, Vol. 50, pp. 665–690.

Simsek, Alp (2013) "Belief Disagreements and Collateral Constraints," *Econometrica*, Vol. 81, pp. 1–53.

Sinha, Arunima (2016) "Learning and the Yield Curve," *Journal of Money, Credit and Banking*, Vol. 48, pp. 513–547.

Stone, Amey (2015) "Bond Yields Jump on Signs of Consumer Strength," `http://blogs.barrons.com/incomeinvesting/2015/09/15/bond-yields-jump-on-signs-of-consumer-strength/`, Published September 15, 2015.

Svensson, Lars E. O. and Michael Woodford (2004a) "Indicator variables for optimal policy under asymmetric information," *Journal of Economic Dynamics and Control*, Vol. 28, pp. 661–690.

——— (2004b) “Indicator variables for optimal policy under asymmetric information,” *Journal of Economic Dynamics and Control*, Vol. 28, pp. 661–690.

Szkup, Michal and Isabel Trevino (2015) “Information acquisition in global games of regime change,” *Journal of Economic Theory*, Vol. 160, pp. 387 – 428.

Tan, Fei and Todd B. Walker (2015) “Solving generalized multivariate linear rational expectations models,” *Journal of Economic Dynamics and Control*, Vol. 60, pp. 95–111.

Tang, Jenny (2013) “Uncertainty and the signaling channel of monetary policy,” Working Papers 15-8, Federal Reserve Bank of Boston.

——— (2015) “Uncertainty and the Signaling Channel of Monetary Policy,” Working Paper, Federal Reserve Bank of Boston.

Taylor, John B. (1993) “Discretion versus policy rules in practice,” *Carnegie-Rochester Conference Series on Public Policy*, Vol. 39, pp. 195–214.

Thornton, Daniel L. and David C. Wheelock (2014) “Making sense of dissents: a history of FOMC dissents,” *Federal Reserve Bank of St. Louis Review*, Vol. 96, pp. 213–227.

Townsend, Robert M (1983) “Forecasting the Forecasts of Others,” *Journal of Political Economy*, Vol. 91, pp. 546–88.

Veldkamp, Laura L. (2011) *Information Choice in Macroeconomics and Finance*: Princeton University Press.

Venkateswaran, Venky and Christian Hellwig (2009) “Setting The Right Prices for the Wrong Reasons,” 2009 Meeting Papers 260, Society for Economic Dynamics.

Waller, Christopher (1992) “A bargaining model of partisan appointments to the central bank,” *Journal of Monetary Economics*, Vol. 29, pp. 411–428.

Woodford, Michael (2001) "Imperfect Common Knowledge and the Effects of Monetary Policy," NBER Working Papers 8673, National Bureau of Economic Research, Inc.

Wright, Jonathan H. (2011) "Term Premia and Inflation Uncertainty: Empirical Evidence from an International Panel Dataset," *American Economic Review*, Vol. 101, pp. 1514–34.

Wu, Jing Cynthia and Fan Dora Xia (2014) "Measuring the Macroeconomic Impact of Monetary Policy at the Zero Lower Bound," NBER Working Papers 20117, National Bureau of Economic Research, Inc.

Wu, Tao (2001) "Macro factors and the affine term structure of interest rates," Working Paper 2002-06, Federal Reserve Bank of San Francisco.

Yang, Ming (2015) "Coordination with flexible information acquisition," *Journal of Economic Theory*, Vol. 158, pp. 721–738.