Identification of Rational Expectations Models Under Information Frictions

Julian F. Ludwig*
Texas Tech University

September 19, 2019

Link to the most recent version: `www.julianfludwig.com`

Abstract

Identification of full information rational expectations (FIRE) models suffers from Manski's (1993) reflection problem. I extend the standard rational expectations (RE) model to allow for a more general information structure and introduce a new framework to identify the generalized model with forecaster data. Identification is no longer subject to the reflection problem when two changes are made to the information structure: the addition of news shocks and imperfect information. News shocks provide additional variation in expectations about the future. Imperfect information provides changes in beliefs about past states, through which the feedback between expectations and decisions goes only in one direction. Expectations data are consistent with both. An application to Greenbook forecasts illustrates the importance of both news shocks and learning about the past. When I apply this framework to a Blanchard and Quah (1989) decomposition, I reach qualitatively new results. For example, expansionary supply shocks decrease unemployment. Supply shocks are also particularly subject to both news and information rigidities, so relaxing the information structure is key to correctly identifying these shocks.

Keywords: Expectations, Information Rigidity, News Shocks, Survey Forecasts
JEL Codes: C2, C3, C5, E3, E5, E6, E7

---

*Ludwig: Assistant Professor, Department of Economics, Texas Tech University, 253 Holden Hall, Lubbock TX 79409, U.S.A. Email: julian.ludwig@ttu.edu. I am grateful to Olivier Coibion, Saroj Bhattarai and Haiqing Xu for their invaluable guidance and support. I also benefited from conversations with Chi Zhang, Stefano Eusepi, Christoph Boehm, Tatevik Sekhposyan, Jessie Coe, Kevin Kuruc, Cooper Howes, Choongryul Yang, Andrés Méndez Ruiz, and Shaofei Jiang as well as with many seminar participants from the University of Texas at Austin, and Texas A&M University.

1 Introduction

In modern micro-founded macro models, the decisions of economic agents are inherently forward-looking and therefore depend on their expectations about the future. But if their expectations are also formed based on the current state, as is generally assumed in full-information models, it becomes difficult to determine to what extent expectations affect actions and vice-versa. I relate this simultaneity to Manski's (1993) reflection problem and introduce a new way to deal with its implications for identification.

The importance of expectations has long been emphasized in rational expectations (RE) models (see e.g. Lucas, 1972, 1976; Kydland and Prescott, 1982). This paper provides a new methodology to identify the parameters of RE models using data on the expectations of economic agents at multiple horizons. Specifically, I relax the full-information assumption and allow for information to diffuse to agents both before and after shocks are realized, which allows me to match both forecasts and backcasts in the data. The resulting variation in expectations is no longer proportional to current actions so that identification is possible. This flexible information structure brings together the literature on news shocks (see Beaudry and Portier, 2004, 2006), where information arrives before impact, and the literature on information frictions (see Mankiw and Reis, 2002; Woodford, 2003; Sims, 2003), where agents gather relevant information only after the period is realized. Combining both dimensions is the key to identifying the parameters of the model.

I relate the identification issue to Manski (1993), who describes the role of expectations about competitors and peers in social interaction models. His well-known *reflection problem* states that when expectations are modelled simply as the average among individuals, the researcher cannot distinguish whether expectations change individual behaviour or if they simply reflect behaviour without causing it. I extend his proposition to show that full information rational expectations (FIRE) models suffer from the same type of reflection problem as well: since information is only gathered in one period, expectations about the future are proportional to realizations today. The researcher can therefore not distinguish the direct effect of a shock on the economy from the indirect effect of observing that shock. The FIRE literature successfully bypasses this reflection problem by imposing a set of parameter restrictions on the FIRE model typically derived from a micro-founded model. I provide a new way to deal with this identification issue by incorporating data on expectations at multiple horizons and by relaxing the full information assumption.

There are many ways to relax full information (FI). Mankiw and Reis (2002) introduce a *sticky information* approach where agents update their expectations infrequently, but when they update their expectations, they fully observe the state. Woodford (2003) models information rigidities with a *noisy information* approach where agents receive noisy signals about the state of the economy. Sims (2003) and Maćkowiak and Wiederholt (2009) make acquiring information costly so that agents need to choose whether they want to pay the cost to get information. The solution of this problem is characterized by *rational inattention*, where agents choose to deviate from full information. Models with such information rigidities allow for information to arrive *after* the shocks are realized so that agents no longer fully observe the current state of the economy. In contrast, news shock models allow for information to arrive *before* the shock hits the economy. Beaudry and Portier (2004) introduce news shocks as noisy signals about future developments of the economy, while Beaudry and Portier (2006), Davis (2007), and Christiano et al. (2010) model news shocks as future shocks that are fully observed today. My paper combines the two branches of the literature to formalize a much more general information structure than what is allowed in standard macroeconomic models.

The core identification property is that future outcomes only have an effect on today's economy if agents observe these outcomes in advance. Hence, any fluctuations that are unobserved cannot be caused by future outcomes. These unobserved fluctuations are therefore fully backward-looking and can be used to identify the effect of current on future outcomes. Changes in expectations about the previous period, defined as *backcast revisions*, collect these unobserved fluctuations. Hence, data on expectations about both the contemporaneous and the previous period are sufficient to identify one direction of the relationship between outcomes and expectations. The other direction is identified when data on expectations about the next period is available. Given the identified effect of current outcomes on expectations of future outcomes, the remaining variation in expectations of the next period must come from additional information that is obtained today about the future. The effect of the remainder on current outcomes thus identifies the effect of expectations so that the forward- and backward-looking components of the RE model are identified. Key for identification is the timing assumption that expectations can only depend on information obtained up until today, but not on information obtained tomorrow. Timing restrictions on information sets are common in the production function literature (see Olley and Pakes, 1996; Blundell and Bond, 2000; Levinsohn and Petrin, 2003). To my knowledge, this is the first paper that uses the variation generated by relaxing full information to separately identify the parameters governing the backward- and forward-looking dynamics of RE models.

The proposed strategy identifies the forward- and backward-looking components simultaneously, without imposing additional restrictions on the model equations. Instead of imposing a particular structural model, this approach nests all models that have the form of RE models with a flexible information structure. Hence, estimation is less subject to model specifications other than the choice which variables to include, the number of lags and leads, and how the shocks are orthogonalized. Identification instead relies on the assumption that data on expectations across horizons, both future and past, is available and that this data correctly captures beliefs of agents. Moreover, identification requires a positive variance of backcast revisions, and noncollinearity between now- and forecasts, features which appear to be consistent with the data. Hence, I impose a completely different set of assumptions to identify the RE model than what is common in the literature.

I implement the identification strategy using data on expectations from the Greenbook, a collection of forecasts from the U.S. Federal Reserve. In the first application, I estimate an unrestricted multivariate RE model with output, consumption, and investment growth, as well as inflation. I find that about half the information about shocks is gathered before and during realization, while the other half is obtained only after the shock materializes. This provides evidence for information rigidities, where agents don't fully observe the current state of the economy. Moreover, around one fourth of the information is collected before realization, which provides support for the news shock literature. I then produce counterfactuals by shifting arrival of information to simulate a full information environment, where everything is observed on impact. In this counterfactual, persistence of all variables significantly declines. Hence, the information structure seems to be a significant driver of the persistence in macro variables.

In the second application, I estimate a RE model with output growth and changes in unemployment, and orthogonalize the shocks following Blanchard and Quah (1989): demand shocks are assumed to be transitory shocks, while supply shocks are the only shocks with a long run impact on output. The results indicate that demand shocks are much better observed than supply shocks, hence, agents seem to be better informed about the demand side of the economy, and less informed about the production side. Moreover, I find on average a significant response to supply shocks before impact, which is line with the news shock literature modelling anticipated supply side shocks. Overall, both demand and supply shocks increase output and decrease unemployment which is consistent with standard real business cycles (RBC) models. The finding that supply shocks decrease unemployment is different from the conclusion of Blanchard and Quah (1989), who find increased unemployment in response to supply shocks. This illustrates the importance of properly controlling for timing of information arrival and processing when identifying economic shocks.

This paper relates to the literature on structural vector autoregressions (SVARs), where economic shocks are identified with reduced form models (see Sims, 1980). Similar to SVARs, my identification strategy does not require one to specify the structural equations of the RE model. However, while SVARs identify the reduced form version of the RE model, this paper identifies the structural version directly, where expectations about the future and dependence on past are identified separately.

My paper builds on and contributes to the literature on news shocks. Beaudry and Portier (2004, 2006) introduce the notion of TFP news shocks as changes in current expectations about future productivity, and they find that these news shocks have real effects today, even though they only materialize in the future. Beaudry and Portier (2006, 2014), Barsky and Sims (2011, 2012), Beaudry, Dupaigne and Portier (2011), Barsky, Basu and Lee (2015) and others identify TFP news shocks in SVARs as changes in current expectations, measured for example in terms of stock prices or consumer confidence, that are orthogonal to current but contribute to future TFP. Today's effects of the news shocks are then considered as effects of expectations, while the shock only realizes in the future when TFP increases. These papers therefore rely on structural assumptions about the dynamic effects of news shocks on economic variables. My approach relies on a different set of assumptions, yet yields results that confirm the importance of news shocks for macroeconomic dynamics.

This paper also contributes to the literature on VAR invertibility, which refers to the ability to rewrite the RE model as a reduced form VAR of observables (see Fernández-Villaverde et al., 2007). Non-invertibility occurs if there are unobserved state variables causing a bias in VAR estimation (see Watson, 1986). Models with news shocks are particularly vulnerable to non-invertibility, because VAR variables might not be able to capture the information used to predict news shocks (see Leeper, Walker and Yang, 2013). Watson (1986), Sims and Zha (2006), and Sims (2012) address the invertibility issue by including forward-looking variables. I introduce an alternative way to avoid non-invertibility by estimating the reduced form model (VAR) with backcast revisions, as described above.

An essential feature in Manski (1993) is that beliefs are modelled rather than observed. Whether true expectations are indeed equal to modelled expectations cannot be tested using data on realizations alone. Manski (2004) thus recommends usage of data on expectations, namely: "econometric analysis of decision making with information cannot prosper on choice data alone" (see Manski, 2004, p. 1330). Macroeconomists have access to rich data on expectations on macro indicators at multiple horizons, which is provided by the Federal Reserve Bank of Philadelphia, the University of Michigan and other institutions, and there is a growing literature that makes use of this data.

Coibion and Gorodnichenko (2015) use survey data on expectations to test for both, full information (FI) and rational expectations (RE). They reject FIRE, and find that rejection most likely reflects deviations from FI rather than RE. This is in line with Adam and Padula (2011), who show that survey expectations satisfy the law of iterated expectations, a law that is satisfied under RE. Based on these findings and in line with the literature on information frictions and news shocks, I only relax FI and keep the RE assumption.

There is already a growing literature using data on expectations in RE models. Del Negro and Eusepi (2011) and Del Negro and Schorfheide (2013) use data on both realizations as well as on expectations to model inflation expectations in DSGE models, while Nguyen and Miyamoto (2014), Hirose and Kurozumi (2012), and Milani and Rajrhandari (2012) include data on expectations in DSGE models to measure the effect of TFP news shocks, which are anticipated productivity shocks. Fuhrer (2017) embeds survey forecasts in a standard DSGE model and demonstrates that much of the persistence in aggregate data can be attributed to slow moving expectations. Precursors of this literature estimate univariate models, for example, Roberts (1995) uses survey measures to estimate the Philips curve. Instead of using survey data directly, Galí and Gertler (1999) first regress future variables on information sets to model expectations, and then they estimate the Philips curve based on these projections. While I relax the information structure to match data on expectations, these models either rely on measurement errors, or deviations from rationality (RE), as these models recognize that observed expectations do not always match model predictions.

Section 2 illustrates the reflection problem in rational expectations models. Section 3 provides intuition for the new identification method with a univariate model for inflation. Section 4 generalizes the information structure and identifies the model. Section 5 discusses identifying restrictions to extract economic shocks. Section 6 examines forecaster data from the Federal Reserve and uses this data to estimate the model. Section 7 concludes.

2 Manski's Reflection Problem in RE Models

Consider a standard linear rational expectations model:

$$\mathbf{y}_t = \mathbf{\Gamma}\mathbf{y}_{t-1} + \mathbf{\Phi}E[\mathbf{y}_{t+1}|\mathcal{F}_t] + \mathbf{\Pi}\mathbf{u}_t, \tag{1}$$

where $\mathbf{y}_t$ is an N-dimensional vector of variables and $\mathbf{u}_t$ is an N-dimensional vector of mutually exclusive shocks with zero-mean, i.e. $E[\mathbf{u}_t] = 0$, identity variance-covariance matrix, i.e. $E[\mathbf{u}_t\mathbf{u}_t'] = I$, and no correlation over time, i.e. $E[\mathbf{u}_t\mathbf{u}_{t+k}'] = 0$, $k \neq 0$. Moreover, $\mathcal{F}_t$ is all the information the agent obtains up to period t.[1] $\mathbf{\Gamma}$, $\mathbf{\Phi}$, and $\mathbf{\Pi}$ are $N \times N$ matrices of parameters, which are typically derived by linearising a micro-founded model.

The trademark of rational expectations models is the role of expectations about the future, which makes the model not just backward-looking, but also forward-looking. The workhorse macroeconomic models (e.g. RBC or New Keynesian models) would satisfy the following assumptions:

Rational Expectations (RE) *Expectations are consistent with model.*
The notion of rational expectations is originally defined by Muth (1961), who suggests to model expectations the same way as predictions of the underlying model. This notion is introduced to macroeconomics by Lucas (1972), promoting RE in his later called Lucas's (1976) critique. In particular, he points out that expectations should not be modelled arbitrarily in a static way, instead, macro models should allow for expectations to change over time, and they should be consistent within the model.

[1] Formally, $\mathcal{F}_t$ is denoted by a collection of events (i.e. σ-algebra) that can be assigned probabilities by using all the information from the past. If the information represented by a random variable w_s belongs to $\mathcal{F}_t$ (i.e. the σ-algebra generated by the random variable w_s), then we have $E[w_s|\mathcal{F}_t] = w_s$.

Full Information (FI) *Shocks are only and fully revealed at realization:*

$$E[\mathbf{u}_t|\mathcal{F}_{t+s}] = \begin{cases} E[\mathbf{u}_t], & s < 0, \qquad (i) \\ \mathbf{u}_t, & s \geq 0, \qquad (ii) \end{cases}$$

and initial conditions are known: $E[\mathbf{y}_0|\mathcal{F}_0] = \mathbf{y}_0$.

FI means *all* the information about shocks $\mathbf{u}_t$ is gathered at time of realization. Specifically, FI.i states that no information is gathered before realization, and FI.ii states that no information is gathered after realization, as realized shocks are fully observed at the moment they are realized. Recall that expectations of shocks are normalized to be zero, i.e. $E[\mathbf{u}_t] = 0$. Thus, FI.i is also denoted as the zero conditional mean (ZCM) condition. FI implies that the agent observes all realized variables and shocks, $\mathbf{y}_{t-s}$ and $\mathbf{u}_{t-s}$, for all $s \geq 0$, and predictions about the future are only based on these realizations. Models that satisfy both FI and RE are standard and are sometimes referred to as full information rational expectations (FIRE) models (see e.g. Coibion and Gorodnichenko, 2012).

Under RE and FI.i, the shocks disappear when taking expectations of future variables so that for all $h \geq 1$:

$$E[\mathbf{y}_{t+h}|\mathcal{F}_t] = \mathbf{\Gamma} E[\mathbf{y}_{t+h-1}|\mathcal{F}_t] + \mathbf{\Phi} E[\mathbf{y}_{t+h+1}|\mathcal{F}_t]. \tag{2}$$

Inspection of (2) reveals the following:

Lemma 1 (Invertibility of Future) *Under RE and FI.i, expectations about the future are invertible,* $E[\mathbf{y}_{t+h}|\mathcal{F}_t] = \mathbf{A} E[\mathbf{y}_{t+h-1}|\mathcal{F}_t]$, $h \geq 1$, *where* $\mathbf{A}$ *is the fixed point solving* $\mathbf{A} = (I - \mathbf{\Phi}\mathbf{A})^{-1}\mathbf{\Gamma}$.

Proof: see appendix A.1.

Lemma 1 states that it is sufficient to know expectations about realized variables, $E[\mathbf{y}_t|\mathcal{F}_t]$, to infer expectations about the future, $E[\mathbf{y}_{t+h}|\mathcal{F}_t]$, for all $h \geq 1$. Namely, $E[\mathbf{y}_{t+h}|\mathcal{F}_t] =$

$\mathbf{A}^h E[\mathbf{y}_t|\mathcal{F}_t]$. Under RE and FI.i, lemma 1 is a sufficient condition to identify $\mathbf{A}$ and residuals that are independent of past information, $\mathcal{F}_{t-1}$, which is referred to invertability or fundamentalness in the VAR literature (see Watson, 1986; Fernández-Villaverde et al., 2007; Leeper, Walker and Yang, 2013; Sims, 2012, and discussion of proposition 1).

Lemma 1 means in particular that for $h = 1$ and under FI, $E[\mathbf{y}_{t+1}|\mathcal{F}_t] = \mathbf{A}\mathbf{y}_t$. There is therefore no independent variation in expectations in system (1), which prevents the separate identification of forward- and backward-looking components in (1). More formally, the rational expectations model suffers from Manski's reflection problem, which is a wel-known issue for identification in social interaction models (see Manski, 1993).[2] Proposition 1 uses a similar econometric structure as Manski (1993) and extends the reflection problem from a cross-sectional environment to rational expectations models:

Proposition 1 (Reflection Problem) *(a) Under RE and FI, parameters $\mathbf{\Phi}$ and $\mathbf{\Gamma}$ are not identified separately. Moreover, (b) under RE and FI.i, the composite parameter $\mathbf{A} = (I - \mathbf{\Phi}\mathbf{A})^{-1}\mathbf{\Gamma}$ and residuals $E[\mathbf{y}_{t+k}|\mathcal{F}_t] - E[\mathbf{y}_{t+k}|\mathcal{F}_{t-1}]$ are identified by regressing $E[\mathbf{y}_{t+k}|\mathcal{F}_t]$ on $E[\mathbf{y}_{t+k-1}|\mathcal{F}_{t-1}]$ for any $k \geq 0$.*

Proof: see appendix A.2.

Proposition 1 shows that in the FIRE model, data reveals how the economy propagates (i.e. $\mathbf{A}$) but cannot tell whether this persistence is caused by forward-looking or backward-looking behaviour. Macroeconomists therefore rely on a series of parameter restrictions to extract both $\mathbf{\Gamma}$ and $\mathbf{\Phi}$. For instance, if one assumes $\mathbf{\Phi} = 0$, then $\mathbf{\Gamma} = \mathbf{A}$. This identifying restriction implies that expectations don't matter, and dynamics in the model are entirely backward-looking. While unusual in modern macro models, this is precisely the environment used in empirical models like Laubach and Williams (2003) which is the workhorse approach

[2] Manski (1993) discovers the reflection problem in models with endogenous social effects, which describe how individuals' behaviour is influenced by the behaviour of a group. The reflection problem arises in Manski (1993) when data cannot identify whether average behaviour affects the behaviour of individuals, or whether the average simply reflects individuals' behaviour without causing it. Manski (1993) compares the problem with observing the simultaneous movements of a person with her reflection in the mirror, as this observation does not reveal whether the mirror causes the movements or reflects them.

used by central banks to estimate the natural rate of interest (see e.g. Bullard, 2018). The opposite approach is to set $\mathbf{\Gamma} = 0$ so that $\mathbf{\Phi} = \mathbf{A}^{-1}$. In this case, dynamics are entirely forward-looking, as is the case for example in the workhorse New Keynesian model (see e.g. Clarida, Galí and Gertler, 2000). This implies that information is very important, as without it, the economy is simply a white noise process. More generally, structural macroeconomic models imply a set of zero restrictions on $\mathbf{\Gamma}$ and $\mathbf{\Phi}$. However, different models imply different restrictions and there is little a priori reason to favor one approach over another.

A different branch of literature started by Sims (1980) identifies economic shocks and their responses with vector autoregressions (VARs), which are reduced-form models. Under FI, the regression in proposition 1.b is a VAR estimated with realizations $\mathbf{y}_t$. Proposition 1.b shows that this VAR identifies composite parameter $\mathbf{A}$, as well as residuals, i.e. $E[\mathbf{y}_t|\mathcal{F}_t] - E[\mathbf{y}_t|\mathcal{F}_{t-1}] = \mathbf{\Pi}\mathbf{u}_t$. If these residuals can be linked to economic shocks with restrictions on the impact matrix $\mathbf{\Pi}$, the VAR can produce impulse responses thereof. The property that these residuals are not contaminated by past shocks so that they are only function of current shocks makes the VAR invertible or fundamental (see Watson, 1986; Fernández-Villaverde et al., 2007; Leeper, Walker and Yang, 2013; Sims, 2012).

Blanchard, L'Huillier and Lorenzoni (2013) discuss VARs when relaxing full information FI.ii. To illustrate their point, consider the VAR estimated with real-time data, $E[\mathbf{y}_t|\mathcal{F}_t]$, across t. This VAR identifies composite parameter $\mathbf{A}$, as well as residuals $E[\mathbf{y}_t|\mathcal{F}_t] - E[\mathbf{y}_t|\mathcal{F}_{t-1}]$ according to proposition 1. Note that these residuals are uncorrelated to past information, $\mathcal{F}_{t-1}$, by law of iterated expectations, which means the VAR satisfies the invertibility condition. Blanchard, L'Huillier and Lorenzoni (2013) however show that these residuals can only be linked to economic shocks under FI, where VAR residuals are equal to the residuals of the model, $E[\mathbf{y}_t|\mathcal{F}_t] - E[\mathbf{y}_t|\mathcal{F}_{t-1}] = \mathbf{\Pi}\mathbf{u}_t$. Once FI is relaxed, true economic shocks can no longer be separated from noise. Hence, when relaxing FI, identification of composite parameter $\mathbf{A}$ and residuals $E[\mathbf{y}_t|\mathcal{F}_t] - E[\mathbf{y}_t|\mathcal{F}_{t-1}]$ might not provide economic meaning per se.

Estimation of regression in proposition 1.b does not require data on expectations at multiple horizons, i.e. $E[\mathbf{y}_{t+h}|\mathcal{F}_t]$, for several h. Even when data on expectations is available at multiple horizons, it is unclear how to incorporate this data in FIRE models. In particular, proposition 1.a shows that a researcher could still only identify a combination of $\mathbf{\Gamma}$ and $\mathbf{\Phi}$.

Full information is inconsistent with data on expectations. In particular, data rejects FI.ii as expectations are not equal to realized values, $E[\mathbf{y}_{t-k}|\mathcal{F}_t] \neq \mathbf{y}_{t-k}$, for some t and $k \geq 0$. Moreover, FI.i does not hold because forecasts are noncollinear: there is no matrix that satisfies lemma 1, $E[\mathbf{y}_{t+h}|\mathcal{F}_t] = \mathbf{A}E[\mathbf{y}_{t+h-1}|\mathcal{F}_t]$, for all t and $h \geq 1$. FI.ii is relaxed in models with information frictions such as sticky and noisy information models (see Mankiw and Reis, 2002; Woodford, 2003; Sims, 2003). FI.i is relaxed in models with news shocks (see Beaudry and Portier, 2004, 2006). This paper uses the variation generated by relaxing FI to match data on expectations at all horizons, and to separately identify $\mathbf{\Gamma}$ and $\mathbf{\Phi}$.

3 A Simple Example

Before I formally introduce the model with flexible information structure, I use a simple example to illustrate the new framework. Consider the following univariate version of model (1) governing inflation:

$$\pi_t = \gamma\pi_{t-1} + \phi E[\pi_{t+1}|\mathcal{F}_t] + \sigma u_t, \tag{3}$$

with *iid* shocks $u_t \sim N(0, 1)$, a forward-looking and a backward-looking component, ϕ and γ respectively. Model 3 represents the solution to a central bank's problem (see appendix B.1).

Under FI, expected future inflation fully reflects current inflation, $E[\pi_{t+1}|\mathcal{F}_t] = a\pi_t$, where a is the solution to the fixed point in $a = (1 - \phi a)^{-1}\gamma$, according to lemma 1. By proposition 1, additional restrictions on parameters are necessary to identify the model.

For example, Coibion and Gorodnichenko (2015) describe inflation with a backward-looking AR(1) process and thus implicitly assume $\phi = 0$ so that $\gamma = a$ is identified by assumption. In this environment, expectations have no impact on inflation, which means that whether agents are well informed or not does not affect the inflation rate.

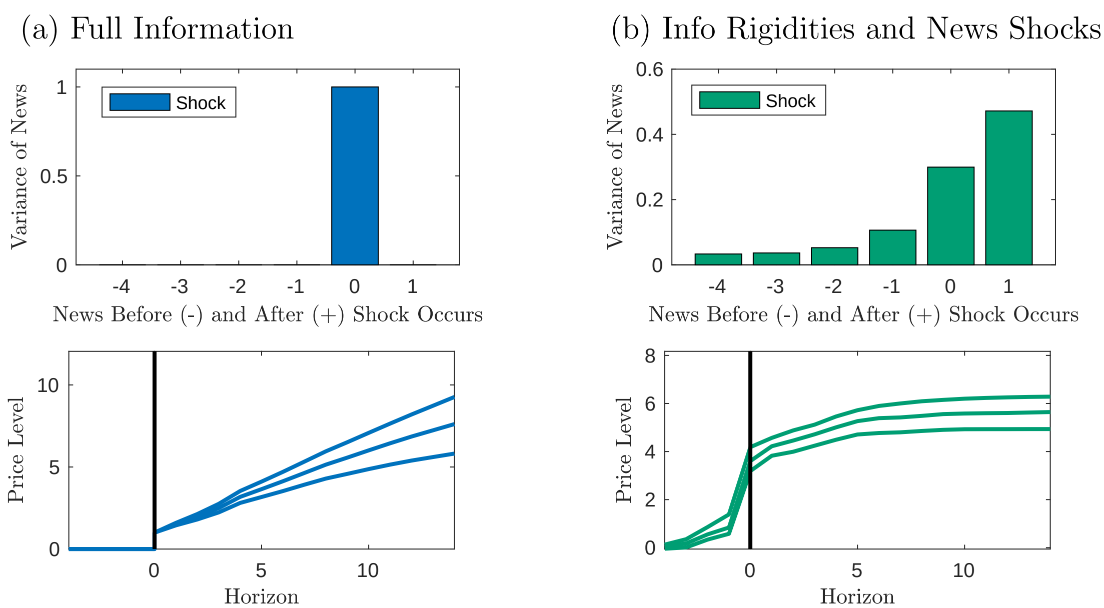
Note that FI imposes strong model restrictions that can be rejected when data on expectations is observed. The Federal Reserve Bank of Philadelphia provides data on Fed's expectations about past, current, and future inflation. This Greenbook data shows that about 14% of inflation's variance is not observed in real time, which rejects the full information assumption FI.ii (see figure 2 of section 6). Moreover, the forecasts of inflation are noncollinear, which rejects FI.i according to lemma 1. Motivated by these observations, I relax the full information assumption and allow for information to diffuse slowly over time. Specifically, the agent obtains news about future shocks ahead of time, while some information about the shock is delayed. The top left plot of figure 1 illustrates how information is concentrated under FI, and the top right plot shows an example of how information about the shock diffuses slowly starting before realization $(t-4)$ up until one period after realization $(t+1)$.

Consider the regression of proposition 1.b when relaxing FI.i:

$$\begin{aligned}
E[\pi_t|\mathcal{F}_t] &= \gamma E[\pi_{t-1}|\mathcal{F}_t] + \phi E[\pi_{t+1}|\mathcal{F}_t] + \sigma E[u_t|\mathcal{F}_t], \\
&= aE[\pi_{t-1}|\mathcal{F}_t] + \sum_{s=0}^{\infty}(b_1)^s b_0 \sigma E[u_{t+s}|\mathcal{F}_t], \qquad (4)\\
&= aE[\pi_{t-1}|\mathcal{F}_{t-1}] + a(E[\pi_{t-1}|\mathcal{F}_t] - E[\pi_{t-1}|\mathcal{F}_{t-1}]) + \sum_{s=0}^{\infty}(b_1)^s b_0 \sigma E[u_{t+s}|\mathcal{F}_t], \qquad (5)
\end{aligned}$$

where a is the fixed point in $a = (1-\phi a)^{-1}\gamma$, $b_1 = (1-\phi a)^{-1}\phi$, and $b_0 = (1-\phi a)^{-1}$. Regression of $E[\pi_t|\mathcal{F}_t]$ on $E[\pi_{t-1}|\mathcal{F}_{t-1}]$ is biased, as for example yesterday's beliefs about tomorrow's shock, $E[u_{t+1}|\mathcal{F}_{t-1}]$, shift the left- and right-hand side variable simultaneously: $E[\pi_{t-1}|\mathcal{F}_{t-1}]$ shifts by $(b_1)^2 b_0 \sigma$, and $E[\pi_t|\mathcal{F}_t]$ shifts by $b_1 b_0 \sigma$. This inability to identify reduced form

Figure 1: Information Diffusion in RE Model



Notes: This figure compares (a) the full information rational expectations model (3) with (b) the same model when FI is relaxed. The top panel displays the share of variance of the shock that is learned on average h quarters after realization, respectively, -h quarters before realization, where h is plotted on the x-axis (see (28) for a formal definition of the metric). The lower panel displays the average impulse response to a shock. (a) The FI impulse response is the estimated response of an AR(5) process using real-time data. (b) The impulse response of the model without FI is estimated based on a backcast, a nowcast, and five forecasts, where the RE model in (3) is extended to include five lag and four lead terms. 68% confidence bands are estimated using bootstrap method. The model is estimated with quarterly Greenbook data, 1967Q2 - 2011Q4, and price level is defined as GDP price deflator. Units of the bottom plots are cumulated annualized percentage points.

parameters under foresight is known as non-invertibility problem (see Fernández-Villaverde et al., 2007; Leeper, Walker and Yang, 2013; Sims, 2012).

Data on expectations at multiple horizons solve this identification issue. The idea is that inflation does not *directly* depend on the future in the RE model, it only depends *indirectly* on the future through current expectations. Hence, any variation that is unobserved cannot depend on the future, because expectations are formed based on observed variation, only, i.e. variation in $\mathcal{F}_t$. Variation that is unobserved today, but it is observed tomorrow is captured in tomorrow's *backcast revision*, which is the change in expectations about the previous period. Again, this backcast revision cannot depend on the future, because the variation that causes a backcast revision is unobserved when it is realized. Formally, the forward-looking term drops out when calculating the backcast revision:

$$
\begin{aligned}
&E[\pi_{t-1}|\mathcal{F}_t] - E[\pi_{t-1}|\mathcal{F}_{t-1}] \\
&\quad = \gamma(E[\pi_{t-2}|\mathcal{F}_t] - E[\pi_{t-2}|\mathcal{F}_{t-1}]) + \phi\left(E\big[E[\pi_t|\mathcal{F}_{t-1}]\big|\mathcal{F}_t\big] - E\big[E[\pi_t|\mathcal{F}_{t-1}]\big|\mathcal{F}_{t-1}\big]\right) \\
&\qquad + \sigma\left(E[u_{t-1}|\mathcal{F}_t] - E[u_{t-1}|\mathcal{F}_{t-1}]\right), \\
&\quad = \gamma(E[\pi_{t-2}|\mathcal{F}_t] - E[\pi_{t-2}|\mathcal{F}_{t-1}]) + \sigma\left(E[u_{t-1}|\mathcal{F}_t] - E[u_{t-1}|\mathcal{F}_{t-1}]\right), \qquad (6)
\end{aligned}
$$

as by law of iterated expectations, $E[E[\pi_t|\mathcal{F}_{t-1}]|\mathcal{F}_t] = E[\pi_t|\mathcal{F}_{t-1}]$. Note how the backcast revision in 6 only depends on beliefs about past shocks, but not on beliefs about current and future shocks, which are responsible for the omitted variable bias (OVB) in estimating a. In particular, equation (5) together with (6) show that regression of the nowcast $E[\pi_t|\mathcal{F}_t]$ on backcast revision $(E[\pi_{t-1}|\mathcal{F}_t] - E[\pi_{t-1}|\mathcal{F}_{t-1}])$ does not suffer from OVB, as the left- and right-hand side variables depend on different unobservables. This regression requires expectations at two horizons, expectations about today and expectations about the previous period. This is different from the regression of proposition 1 of today's nowcast $E[\pi_t|\mathcal{F}_t]$ on yesterday's nowcast $E[\pi_{t-1}|\mathcal{F}_{t-1}]$, where only one horizon is necessary.

Is the above regression unbiased? There is one more catch, while past shocks are uncorrelated with future shocks, it is still possible that *beliefs* about shocks are correlated across time, i.e. $Corr(E[u_{t+k}|\mathcal{F}_t], E[u_{t+h}|\mathcal{F}_t]) \neq Corr(u_{t+k}, u_{t+h}) = 0$, for some $k \neq h$. This is not an issue under FI, where expectations are either equal to zero or equal to the true shock so that correlations of beliefs are always equal to zero, anyways. Intuitively, beliefs about uncorrelated shocks are correlated when there is uncertainty about timing. For example, the central bank receives information about a future tax cut, but it doesn't know whether it will occur in one or in two years. The belief of experiencing a tax cut is thus positively correlated between one and two years from now, even though tax cuts are uncorrelated over time. Identification of feature a needs the following restriction: there is no more uncertainty about timing, once the shock is realized. Hence, once the tax cut occurs, the central bank is no longer uncertain whether it occurs today or another time, while there might still be uncertainty about the size of the tax cut. This restriction on uncertainty about timing is denoted as *revealed timing at realization* (RTR) assumption: $Cov(E[u_t|\mathcal{F}_{t+k}], E[u_{t+h}|\mathcal{F}_{t+k}]) = Cov(u_t, u_{t+h}) = 0$, $\forall h \geq 1$, $\forall k \geq 0$.

The above exercise shows how we can still identify the composite parameter a, despite having relaxed FI.i and FI.ii. It turns out, data on expectations can do even more than that. The reason is that the model is no longer subject to Manski's (1993) reflection problem, as relaxing FI.i provides additional variation in expectations that breaks the reflection. Data on expectations about the future can make use of this variation to separately identify ϕ and γ. In particular, shift equation (4) forward:

$$E[\pi_{t+1}|\mathcal{F}_t] = aE[\pi_t|\mathcal{F}_t] + \sum_{s=0}^{\infty}(b_1)^s b_0 \sigma E[u_{t+1+s}|\mathcal{F}_t], \tag{7}$$

and use identified a to extract the remainder, i.e. $E[\pi_{t+1}|\mathcal{F}_t] - aE[\pi_t|\mathcal{F}_t]$, which is a weighted sum of all future shocks. Moreover, extract the remainder of the nowcast in (4), as well, i.e. $E[\pi_t|\mathcal{F}_t] - aE[\pi_{t-1}|\mathcal{F}_t]$, and then regress it on the remainder of (7) to identify b_1. This is apparent in equation (4), after taking the current shock out of the sum:

$$
\begin{aligned}
E[\pi_t|\mathcal{F}_t] - aE[\pi_{t-1}|\mathcal{F}_t] &= \sum_{s=0}^{\infty}(b_1)^s b_0 \sigma E[u_{t+s}|\mathcal{F}_t], \\
&= b_1 \left(\sum_{s=0}^{\infty}(b_1)^s b_0 \sigma E[u_{t+1+s}|\mathcal{F}_t] \right) + b_0 \sigma E[u_t|\mathcal{F}_t], \\
&= b_1 \left(E[\pi_{t+1}|\mathcal{F}_t] - aE[\pi_t|\mathcal{F}_t] \right). + b_0 \sigma E[u_t|\mathcal{F}_t]. \qquad (8)
\end{aligned}
$$

This regression gives an unbiased estimate of b_1 under RTR. RTR is required because the weighted sum of beliefs of future shocks should not be correlated with the belief of the current shock, i.e. $E[u_t|\mathcal{F}_t]$, as it is the error term in regression (8). Remember that composite parameters are invertible functions of the model parameters, $a = (1-\phi a)^{-1}\gamma$, and $b_1 = (1-\phi a)^{-1}\phi$, hence, once a and b_1 are identified, ϕ and γ are identified, as well as $b_0 = (1-\phi a)^{-1}$.

To summarize so far, if data on backcasts, nowcasts, and forecasts are available, the parameters γ and ϕ of the rational expectations model (3) are identified. Identification is possible without a parametric structure for how agents gather information. This flexible information structure allows to match data on expectations at different horizons, without relying on measurement errors or deviations from rationality. The necessary rank conditions are that backcast revisions, nowcasts, and forecasts are noncollinear. These conditions are testable and are satisfied for inflation expectations. The necessary zero conditional mean (ZCM) conditions are that beliefs about realized shocks are uncorrelated with beliefs about future and past shocks (RTR), while beliefs across unrealized shocks can still be correlated.

Under FI, parameters γ and ϕ of model (3) and the underlying process π_t are sufficient to extract the shocks, i.e. $\sigma u_t = \frac{1}{b_0}(\pi_t - a\pi_{t-1})$, to identify the persistence of the process, i.e. $\rho(\pi_t, \pi_{t-1}) = a$, and to identify predictions of the agents at all horizons, i.e. $E[\pi_{t+h}|\mathcal{F}_t] = a^h \pi_t$, $\forall h > 0$, and $E[\pi_{t-k}|\mathcal{F}_t] = \pi_{t-k}$, $\forall k \geq 0$. Without FI, identified parameters γ and ϕ quantify economic relationships, but they say nothing about realized shocks, u_t, nothing about how inflation propagates over time, and they do not reveal predictions of the agents beyond the expectations used for estimation. To identify these characteristics, the researcher needs to identify how much information is gathered by the agents before, during, and after the shocks are realized. One extreme is that agents receive most information about the shocks far ahead of time so that agents respond to this information in advance through ϕ, generating a high persistence in inflation. Another extreme is when most information is gathered after realization so that once the shocks are observed, they are no longer relevant for today's economy so that expectations barely matter, despite a large ϕ.

I introduce a non-parametric way to identify the information structure, in order to capture the characteristics described above. For this I need to assume a finite *information diffusion interval* (IDI): all the information about a shock diffuses within the interval of H periods before until K periods after the shock is realized. This means that shocks are fully revealed after K periods, i.e. $E[u_t|\mathcal{F}_{t+K}] = u_t$, and there is no information available $H+1$ periods before impact, i.e. $E[u_t|\mathcal{F}_{t-H-1}] = E[u_t] = 0$. The shocks can then be decomposed into $H+1+K$ revisions:

$$u_t = E[u_t|\mathcal{F}_{t+K}] = \sum_{k=-H}^{K} \Big(E[u_t|\mathcal{F}_{t+k}] - E[u_t|\mathcal{F}_{t+k-1}] \Big). \tag{9}$$

These revisions are uncorrelated by law of iterated expectations. The variances of the different shock components therefore add up to one:

$$Var(u_t) = \sum_{k=-H}^{K} Var\Big(E[u_t|\mathcal{F}_{t+k}] - E[u_t|\mathcal{F}_{t+k-1}] \Big),$$

where $Var(u_t) = 1$. The variance of a shock component, i.e. $Var\left(E[u_t|\mathcal{F}_{t+k}] - E[u_t|\mathcal{F}_{t+k-1}]\right)$, thus reflects how much of total variance the agent learns k periods after the shock is realized, respectively, $-k$ periods before it is realized. The top right plot of figure 1 shows this measure for shocks to inflation expectations under the assumption of $K = 1$ and $H = 4$. About half of the shock's variance is learned before or when it is realized, and the remaining half after. Identification of this result will be discussed shortly. The top left panel shows that under FI, all the variance is learned on impact.

The variances of the components don't fully capture the information structure, yet. Remember there is uncertainty about timing so that beliefs about different shocks can be correlated. This is captured in the covariances, which add up to zero for $h > 0$:

$$Cov(u_t, u_{t+h}) = \sum_{k=-H}^{K+h} Cov\Big(E[u_t|\mathcal{F}_{t+k}] - E[u_t|\mathcal{F}_{t+k-1}], E[u_{t+h}|\mathcal{F}_{t+k}] - E[u_{t+h}|\mathcal{F}_{t+k-1}]\Big),$$

where $Cov(u_t, u_{t+h}) = 0$. RTR states that the expression is also equal to zero when taking the sum only from $k = -H$ to $k = 0$. Again, let's consider the example where the agent receives information about a future tax cut but it is unclear whether it occurs in one or in two years. The agent revises her expectations of both shocks in the same direction so that the revisions about $t+1$ and $t+2$ shocks are positively correlated. Let's say the tax cut occurs in one year at $t+1$. When the tax cut occurs, the agent observes that a tax cut is realized and thus reinforces her previous prediction. Moreover, she revises her incorrect prediction of a $t+2$ tax cut, as it turns out the previous information did not refer to that period, after all. On average, the agent readjusts her beliefs of $t+1$ and $t+2$ tax cuts to an extent where they become uncorrelated, because the agent knows that there is no correlation over time. Hence, the covariances of time t and time $t+1$ revisions add up to zero.

The variance-covariance matrix of the shock components describes how information arrives, which together with the model parameters provide sufficient information to characterize the first and second moments of the data, as well as predictions of the agents. Identification of the different shock components requires data on expectations for the horizons for which agents receive information about shocks. Hence, the IDI assumptions on H and K dictate the necessary forecast horizons for identification. Formally, the shock components can be collected with backcast, nowcast, and forecast revisions according to model (3):

$$
\begin{aligned}
& E[\pi_{t-k}|\mathcal{F}_t] - E[\pi_{t-1}|\mathcal{F}_{t-1}] \\
& \qquad = \gamma(E[\pi_{t-k-1}|\mathcal{F}_t] - E[\pi_{t-k-1}|\mathcal{F}_{t-1}]) + \sigma\left(E[u_{t-k}|\mathcal{F}_t] - E[u_{t-k}|\mathcal{F}_{t-1}]\right), \ k \geq 1, \qquad (10) \\
& E[\pi_{t+h}|\mathcal{F}_t] - E[\pi_{t+h}|\mathcal{F}_{t-1}] \\
& \qquad = \gamma(E[\pi_{t+h-1}|\mathcal{F}_t] - E[\pi_{t+h-1}|\mathcal{F}_{t-1}]) + \phi(E[\pi_{t+h+1}|\mathcal{F}_t] - E[\pi_{t+h+1}|\mathcal{F}_{t-1}]) \\
& \qquad\quad + \sigma\left(E[u_{t+h}|\mathcal{F}_t] - E[u_{t+h}|\mathcal{F}_{t-1}]\right), \ h \geq 0. \qquad (11)
\end{aligned}
$$

After extracting the shock revisions, they can be summed up according to (9) which identifies the shock σu_t, where σ is the standard deviation of that expression. The variance-covariance matrix of the shock components then identifies the information structure so that the model is identified.

Let's assume that information about a shock is gathered four periods ahead of time, up to one period after realization, i.e. $H = 4$ and $K = 1$. I include additional lag and lead terms, as well as a trend in model (3) and estimate the model using quarterly Greenbook data. For identification, I need $K = 1$ backcast, as well as a nowcast and $H+1 = 5$ forecasts of inflation. The top right panel of figure 1 plots the variances of the components, which are informative about how much the agent learns on average before and after the shock is realized. Note that the agent learns more and more about the variance with time which is not an assumption: it is possible for example that the agent learns more at $t-1$ than at $t+1$. The low amount of learning four quarters ahead of time supports the IDI assumption that there is no learning about shocks far ahead of time. Strikingly, the high variance of the revision one quarter after realization shows that a significant share of shocks to inflation are only observed in retrospect. Remember that this is assumed to be zero in standard FIRE models. Hence, the top right panel of figure 1 provides economically significant evidence that the assumption of full information does not hold for modelling inflation; a finding that is consistent with the empirical literature on information frictions (see e.g. Coibion and

Gorodnichenko, 2015).

In an environment with flexible information structure, impulse responses to shocks change over time, as they depend on what information the agent receives about each shock. The lower right panel of figure 1 therefore plots the *average* impulse response. Interestingly, on average, inflation responds to a shock even before it is realized. Since the future can only affect today's economy through expectations, this means that expectations about the future matter. This is exactly the story of the news shock literature, where today's economy is affected by a future TFP shock, because expectations thereof change current behaviour (see Beaudry and Portier, 2004, 2006).

This section shows how data on backcasts, nowcasts, and forecasts fully identify a univariate rational expectations model, when full information is relaxed. The findings of this section might be a result of having a highly stylized model, which is why the next section extends the analysis to a multivariate rational expectations model.

4 RE Model with Generalized Information Structure

This section identifies model (1) and its information structure. The foundation for identification is data on expectations at multiple horizons. Building on the literature on information frictions (see Mankiw and Reis, 2002; Woodford, 2003; Sims, 2003) and news shocks (see Beaudry and Portier, 2004, 2006), I relax the full information (FI) assumption to an extent where the model can replicate data on expectations without relying on measurement errors or deviations from rational expectations (RE). Specifically, relaxing FI allows to generate all the moments of the data on expectations other than the ones restricted by the law of iterated expectations, a law that holds under RE. This generalized information structure permits rich variation in expectations that is used to separately identify the forward-looking and backward-looking components of the model.

Section 4.1 relaxes FI, by allowing for information to be spread out across time so that information arrives before, at, and after shocks are realized. In line with the news shocks literature, and the literature on sticky and noisy information, I model information exogenously, as opposed to models with rational inattention, where information is chosen endogenously. Instead of introducing a parametric model for the information process, I characterize arrival of information with the moments of the data. In particular, I use the moments of conditional expectations of the shocks to describe the information structure.

Section 4.2 solves model (1), without having a parametric expression for diffusion of information. Solving models without FI is challenging, as when expectations matter, the way shocks propagate depends on the available information. The solution thus needs to keep track of the state of the information set, $\mathcal{F}_t$, which increases the number of state variables significantly. However, since information is exogenous, I can solve the model without keeping track of information. The idea is that the rational expectations model only depends on two different information sets, the information set when the economy is realized, $\mathcal{F}_t$, through $E[\mathbf{y}_{t+1}|\mathcal{F}_t]$, and the information set when everything is revealed, through $\mathbf{y}_{t-1}$ and $\mathbf{u}_t$. I thus split up the solution into two parts, the part that is observed at realization, $E[\mathbf{y}_t|\mathcal{F}_t]$, and the part that is observed after realization, $\mathbf{y}_t - E[\mathbf{y}_t|\mathcal{F}_t]$. These two parts can then be solved separately, where solution can be found by conditioning on the same information. As information is fix for each part, solving the two parts is similar to solving perfect foresight models. Solution is then expressed in terms of conditional expectations of shocks, which are exogenously determined by the information structure according to section 4.1.

Section 4.3 identifies model parameters $\mathbf{\Gamma}$ and $\mathbf{\Phi}$, the shocks of the model and the information structure described in section 4.1. Identification is possible as the model is no longer subject to Manski's (1993) reflection problem, which occurs when the econometrician cannot separate the direct effect of a shock from the indirect effect of observing that shock (see section 2). The variation coming from information about the future breaks Manski's (1993) reflection problem, as expectations no longer solely reflect changes in the current economy. New information about the past identifies the backward-looking component, which then helps to extract the variation that solves Manski's (1993) reflection problem. The necessary assumption for identification is that there is no uncertainty about when a shock hits the economy, once it is realized. Identification of shocks and information structure requires the assumptions that information about shocks diffuses within a finite interval. Information structure is characterized by the variance-covariance matrix of changes in expectations about past, current, and future shocks. Identification of shocks also requires identifying restrictions on the impact matrix of the shocks, i.e. $\mathbf{\Pi}$, which are discussed in detail in section 5.

Changing information structure might alter the form of the rational expectations model (1), depending on how it is micro-founded. Two generalizations are considered. First, if there is missing information about the past, beliefs about the past, $E[\mathbf{y}_{t-1}|\mathcal{F}_t]$, might have different effects than the actual past, $\mathbf{y}_{t-1}$, a case discussed in appendix C.1. Second, if there is information about future shocks, expectations of those shocks, $E[\mathbf{u}_{t+h}|\mathcal{F}_t]$, might enter the system separately, a case discussed in appendix C.2. The findings of the paper are robust to both generalizations. The model estimated in section 6.2 extends the number of lag and lead terms of model (1), and introduces a trend.

4.1 Information Structure

Consider a standard linear rational expectations model as in (1) under rational expectations (RE), but without imposing full information (FI). In this generalized framework, a representative agent forms expectations about current, future, and past variables $\mathbf{y}_{t+s}$, as well as shocks $\mathbf{u}_{t+s}$, $\forall s$. The information structure is expressed in terms of shocks only, $\mathbf{u}_{t+s}$, without loss of generality. The assumption of full information (FI) is abandoned among the following dimensions:

Information Diffusion Interval (IDI) *Information about shocks arrive H periods before* *until K periods after realization:*

$$E[\mathbf{u}_t|\mathcal{F}_{t+s}] = \begin{cases} E[\mathbf{u}_t], & s < -H, \qquad (i) \\ \mathbf{u}_t, & s \geq K. \qquad (ii) \end{cases}$$

IDI states that new information may arrive before, during, and after shocks are realized, but information is only gathered within $H + K + 1$ periods. FI is a degenerated case of IDI, where all the information arrives on impact. Specifically, FI.i implies $H = 0$, and FI.ii implies $K = 0$. Identification of forward- and backward-looking parameters $\mathbf{\Gamma}$ and $\mathbf{\Phi}$ does not require K and H to be finite. Finite K and H are necessary for identification of shocks and the underlying information structure.

Incomplete Information and News about the Future (IIN) *Agents update beliefs about shocks before, at, and after realization: $Var(E[\mathbf{u}_t|\mathcal{F}_{t+s}] - E[\mathbf{u}_t|\mathcal{F}_{t+s-1}]) > 0$, for $s = 0$, (i) for some $s < 0$, and (ii) for some $s > 0$.*

IIN is a rank condition necessary for identification of parameters. IIN violates FI, where agents only update on impact, i.e. $Var(E[\mathbf{u}_t|\mathcal{F}_t] - E[\mathbf{u}_t|\mathcal{F}_{t-1}]) = Var(\mathbf{u}_t)$, and $Var(E[\mathbf{u}_t|\mathcal{F}_{t+s}] - E[\mathbf{u}_t|\mathcal{F}_{t+s-1}]) = 0$, for $s \neq 0$. IIN is testable if data on expectations is available at multiple horizons. In particular, agents update their beliefs about past shocks if expectations are not equal to realized values, $E[\mathbf{y}_{t-k}|\mathcal{F}_t] \neq \mathbf{y}_{t-k}$, for some t and $k \geq 0$. Agents update their beliefs about future shocks if there is no matrix that satisfies lemma 1, $E[\mathbf{y}_{t+h}|\mathcal{F}_t] = \mathbf{A}E[\mathbf{y}_{t+h-1}|\mathcal{F}_t]$, for all t and $h \geq 1$, respectively, if now- and forecasts are noncollinear. IIN.i is satisfied in models with news shocks (see Beaudry and Portier, 2004, 2006), while IIN.ii is satisfied in models with information frictions (see Mankiw and Reis, 2002; Woodford, 2003; Sims, 2003).

As discussed in section 3, relaxing FI may cause *beliefs* about shocks to be correlated across time, because of *uncertainty about timing*. The following zero conditional mean (ZCM) assumption restricts this uncertainty for realized shocks.

Revealed Timing at Realization (RTR) *There is no uncertainty about timing after shocks are realized:* $Cov\big(E[\mathbf{u}_t|\mathcal{F}_{t+k}], E[\mathbf{u}_{t+h}|\mathcal{F}_{t+k}]\big) = Cov\big(\mathbf{u}_t, \mathbf{u}_{t+h}\big) = 0,\ \forall h \geq 1,\ \forall k \geq 0.$

RTR states that uncertainty about timing, $Cov\big(E[\mathbf{u}_t|\mathcal{F}_{t+k}], E[\mathbf{u}_{t+h}|\mathcal{F}_{t+k}]\big) \neq 0$, is still allowed among future shocks, but it is no longer allowed for current and past shocks. This means that uncertainty about *when* the shock occurs disappears the moment the shock is realized. Note that there might still be uncertainty about sign, magnitude, and type of shock. In other words, current information on $\mathbf{u}_t$ is no longer correlated with information that is used to predict future and past shocks, $\mathbf{u}_{t+h}$, $h \neq 0$, but information on $\mathbf{u}_t$ might still be incomplete. For example, a tax cut occurs a year from now, but people don't know whether it occurs in one or in two years. Once the tax cut is realized, there is no more uncertainty on whether tax rate changes that same year or the following year. There is however still uncertainty by how much the tax rate changes, and it might even be unclear whether the shock is indeed a tax cut.

RTR is the ZCM assumption necessary to identify parameters $\mathbf{\Gamma}$ and $\mathbf{\Phi}$. RTR is satisfied in models with news shocks as in Beaudry and Portier (2004, 2006); Davis (2007); Christiano et al. (2010), as these models don't allow for uncertainty about timing, at all. In this literature, changes in expectations of future shocks are modelled as independent to changes in expectations about other shocks.

Relaxing FI together with IDI, IIN and RTR produce the following five properties:

(1) *News shocks about the future:* $E[\mathbf{u}_t|\mathcal{F}_{t-h}] \neq E[\mathbf{u}_t]$, *for some* t *and* $1 \leq h \leq H$. This is the counterpart of FI.i and follows from IIN.i, and IDI.i restricts H. News shocks are shocks that are partially or fully observed before they are realized. This means the agent receives information not just about today but also about the future. Information about the future is available, for example, when the government announces future military spending in advance.

(2) *Incomplete information:* $E[\mathbf{u}_t|\mathcal{F}_{t+k}] \neq \mathbf{u}_t$, *for some* t *and* $0 \leq k < K$. This is the

counterpart of FI.ii and follows from IIN.ii, and IDI.ii restricts H. Without FI.ii, information might be missing about shocks, even after they are realized. Information is imperfect, for example, when the government lacks some information about current GDP, because it does not collect taxes and therefore does not have the necessary information until after the calender year ends.

Property (1) is standard in the news shock literature, while property (2) is standard in the literature on information frictions. The next three properties are non-standard, and describe how *beliefs* about shocks can be correlated, even though realized shocks are uncorrelated. Note that properties (3), (4), and (5) describe possible relationships, but they are not assumptions necessary for identification.

(3) *Uncertainty about timing among future shocks:* $Cov\big(E[\mathbf{u}_{t+k}|\mathcal{F}_t], E[\mathbf{u}_{t+h}|\mathcal{F}_t]\big) \neq Cov\big(\mathbf{u}_{t+k}, \mathbf{u}_{t+h}\big) = 0$, *is possible for* $h > k \geq 1$. Note that under FI, this property does not exist, because $Cov\big(E[\mathbf{u}_{t+k}|\mathcal{F}_t], E[\mathbf{u}_{t+h}|\mathcal{F}_t]\big)$ can either be equal to $Cov\big(\mathbf{u}_{t+k}, \mathbf{u}_{t+h}\big)$, $Cov\big(0, \mathbf{u}_{t+h}\big)$, $Cov\big(\mathbf{u}_{t+k}, 0\big)$, or $Cov\big(0, 0\big)$, which are all equal to zero. Timing is uncertain, for example, when information about a future tax cut is available, but it is unclear whether the tax cut occurs in one or in two years. The belief that taxes will be reduced in one year is thus correlated with the belief that taxes will be cut in two years, even though tax reductions themselves are uncorrelated over time. This property is an explanation why expected variables have different persistence than realized variables. For example, Jain (2017) and Ryngaert (2017) find evidence that expected inflation has different persistence than actual inflation. While they argue that perceived inflation persistence might deviate due to missperception of forecasters, this could instead be rationalized with uncertain timing.

(4) *Uncertainty about type:* $Cov\big(E[u_{it}|\mathcal{F}_{t+s}], E[u_{jt}|\mathcal{F}_{t+s}]\big) \neq Cov\big(u_{it}, u_{jt}\big) = 0$, *is possible for* $i \neq j$. Type is uncertain, for example, when there is information about a shock, but it is unclear whether it is a fiscal policy shock or a monetary policy shock. This

property is an explanation why expected variables have different correlations among each other than realized variables.

(5) *Uncertainty about both timing and type:* $Cov\big(E[u_{it+k}|\mathcal{F}_{t+s}], E[u_{jt+h}|\mathcal{F}_{t+s}]\big) \neq Cov\big(u_{it+k}, u_{jt+h}\big) = 0$, *is possible for* $i \neq j$, *and* $h > k \geq 1$. This property occurs when there is both types of uncertainty, for example, when it is unclear whether a signal refers to a monetary policy shock in two years or to a fiscal policy shock in one year.

I don't impose a parametric structure on the information structure. This is desirable because of two reasons. First, it keeps the information structure as general as possible so that the structure nests existing models of the literature. Second, additional assumptions would restrict the relationships of conditional expectations, $E\big[E[\mathbf{y}_{t+k}|\mathcal{F}_t]E[\mathbf{y}_{t+h}|\mathcal{F}_t]\big]$ across k and h. Deviations in observed relationships from these restrictions could then only be explained by either relaxing the RE assumption, or by introducing mistakes in data collection. Instead, this paper matches the data by adjusting information sets only. Specifically, the estimated model of section 6.2 will be exactly identified. In particular, multiple lag and lead terms as well as the generalized information structure replicate all the first and second moments of revisions in expectations at all horizons.

Literature on information frictions often express information processing in terms of variables $\mathbf{y}_t$, instead of the shocks $\mathbf{u}_t$. Properties (1)-(5) can be expressed in variables, as well, by just replacing $\mathbf{u}_t$ with $\mathbf{y}_t$. The advantage of expressing information structure with shocks is that the shocks are uncorrelated and have a unit variance, i.e. $E[\mathbf{u}_{t+k}\mathbf{u}_t] = 0$, and $E[\mathbf{u}_t\mathbf{u}_t'] = I$, for $k \neq 0$, so that deviations in conditional expectations thereof can be simply attributed to the information structure.

The different properties of the information structure can be quantified by decomposing structural shocks $\mathbf{u}_t$ into the following components by IDI:

$$\mathbf{u}_t = \sum_{k=-H}^{K} \Big(E[\mathbf{u}_t|\mathcal{F}_{t+k}] - E[\mathbf{u}_t|\mathcal{F}_{t+k-1}]\Big), \tag{12}$$

where each component reflects what the agent learned about shock $\mathbf{u}_t$, k periods after, respectively $-k$ periods before realization. The shock $\mathbf{u}_t$ is fully revealed at time $t + K$ to the agent, and there is no information available before time $t - H$. This decomposition characterizes the agent's cumulative learning about the shock, which ultimately is fully revealed at $t+H$. More specifically, the components can be interpreted as follows, conditional on today's information set:

- *News components*, $E[\mathbf{u}_{t+h}|\mathcal{F}_t]-E[\mathbf{u}_{t+h}|\mathcal{F}_{t-1}]$, $h \geq 1$, are changes in expectations about a future shock. These will be non-zero whenever agents receive new information about future values of the shock.
- *Surprise components*, $E[\mathbf{u}_t|\mathcal{F}_t]-E[\mathbf{u}_t|\mathcal{F}_{t-1}]$, are changes in expectations about a current shock. This is the only non-zero component in FIRE models, where all the information is learned on impact.
- *Revise components*, $E[\mathbf{u}_{t-k}|\mathcal{F}_t]-E[\mathbf{u}_{t-k}|\mathcal{F}_{t-1}]$, $k \geq 1$, are changes in expectations about a past shock. These revisions will occur whenever agents have imperfect information about current or future shocks and only gradually learn about their values over time.

The information structure can then be quantified by the variance-covariance matrix of the shock components:

$$\mathfrak{D} \equiv E\Big[\big(E[\mathbf{U}_t|\mathcal{F}_t] - E[\mathbf{U}_t|\mathcal{F}_{t-1}]\big)\big(E[\mathbf{U}_t|\mathcal{F}_t] - E[\mathbf{U}_t|\mathcal{F}_{t-1}]\big)'\Big], \tag{13}$$

where $\mathbf{U}_t \equiv [\mathbf{u}'_{t-K} \ \cdots \ \mathbf{u}'_{t+H}]'$. For illustrative purpose, let $K = 1$, and $H = 2$ so that there is no uncertainty about $t - 1$, and no information on shocks beyond $t + 2$. The covariance matrix $\mathfrak{D}$ can then be decomposed into $(K + 1 + H)^2 = 16$ matrices, each with dimension

$N \times N$ as follows:

$$\mathfrak{D} = \begin{pmatrix} \mathfrak{D}_{-1,-1} & \mathfrak{D}_{-1,0} & \mathfrak{D}_{-1,1} & \mathfrak{D}_{-1,2} \\ \mathfrak{D}_{-1,0} & \mathfrak{D}_{00} & \mathfrak{D}_{01} & \mathfrak{D}_{02} \\ \mathfrak{D}_{-1,1} & \mathfrak{D}_{01} & \mathfrak{D}_{11} & \mathfrak{D}_{12} \\ \mathfrak{D}_{-1,2} & \mathfrak{D}_{02} & \mathfrak{D}_{12} & \mathfrak{D}_{22} \end{pmatrix}.$$

This variance-covariance matrix captures how information arrives in the economy. For example, under FI, shocks $\mathbf{u}_t$ are fully and only revealed at the time of realization so that $E[\mathbf{u}_t|\mathcal{F}_t] - E[\mathbf{u}_t|\mathcal{F}_{t-1}] = I\mathbf{u}_t$, which implies $\mathfrak{D}_{00} = I$, and $\mathfrak{D}_{hk} = 0$ for all other h and k. Relaxing FI means other matrices than $\mathfrak{D}_{00}$ might be non-zero. More specifically, the above properties can be quantified as follows:

(1) *Imperfect information:* $\sum_{h=0}^{H} \mathfrak{D}_{hh} \neq I$.

(2) *News shocks about the future:* $\sum_{h=1}^{H} \mathfrak{D}_{hh} \neq 0$.

(3) *Uncertainty about timing among future shocks: $diag(\mathfrak{D}_{hk}) \neq 0$ for some $h > k \geq 0$, and by RTR, $\mathfrak{D}_{h,k} = 0$ for all $h \leq -1$ and for all k.*

(4) *Uncertainty about type: $\mathfrak{D}_{hh}$ is not diagonal for some h.*

(5) *Uncertainty about timing and type: $\mathfrak{D}_{hk} - diag(\mathfrak{D}_{hk}) \neq 0$ for some $h > k \geq 0$.*

If information has no real effects, $\mathfrak{D}$ can be altered arbitrarily without affecting $\mathbf{y}_t$, as long as the properties of the shocks remain satisfied: $E[\mathbf{u}_t\mathbf{u}'_{t+h}] = 0$, $\forall h \neq 0$, and $E[\mathbf{u}_t\mathbf{u}'_t] = I$, which imply $\sum_{h=-K}^{H} \mathfrak{D}_{h,h+s} = 0$, for all $s \neq 0$, and $\sum_{h=-K}^{H} \mathfrak{D}_{hh} = I$. One extreme is to change $\mathfrak{D}$ so that everything is only observed ex post, by setting all matrices equal to zero except $\mathfrak{D}_{-1-1} = I$. Another extreme is to put the identity at the bottom right so that everything is observed H periods in advance, $\mathfrak{D}_{HH} = I$. If $\mathbf{y}_t$ is different for these two cases, the timing of arrival of information affects economic dynamics, which means expectations must matter for economic decisions. If $\mathbf{y}_t$ remains the same, information is irrelevant meaning that forward-looking behaviour has no real impact. Section 6.3 calculates these two counterfactuals, as well as other manipulations of $\mathfrak{D}$.

4.2 Solution

Model (1) depends on variables and shocks conditional on two different information sets, the information set of the period when $\mathbf{y}_t$ is realized, $\mathcal{F}_t$, and the information set when $\mathbf{y}_t$ is fully revealed. The economy thus propagates differently for information that arrives before time t, than for information that arrives after time t:

$$\mathbf{y}_t = \underbrace{E[\mathbf{y}_t|\mathcal{F}_t]}_{\text{info before } t} + \underbrace{(\mathbf{y}_t - E[\mathbf{y}_t|\mathcal{F}_t])}_{\text{info after } t} \tag{14}$$

Let's solve for the part of the model that is based on information that arrives before time t. The solution of $E[\mathbf{y}_t|\mathcal{F}_t]$ can be expressed as in Binder and Pesaran (1997):

$$E[\mathbf{y}_t|\mathcal{F}_t] = \mathbf{A}E[\mathbf{y}_{t-1}|\mathcal{F}_t] + E[\mathbf{z}_t|\mathcal{F}_t], \tag{15}$$

$$E[\mathbf{z}_t|\mathcal{F}_t] = \mathbf{B}_1 E[\mathbf{z}_{t+1}|\mathcal{F}_t] + \mathbf{B}_0\mathbf{\Pi}E[\mathbf{u}_t|\mathcal{F}_t]. \tag{16}$$

Appendix B.2 derives $\mathbf{A} = (I - \mathbf{\Phi}\mathbf{A})^{-1}\mathbf{\Gamma}$, $\mathbf{B}_1 = (I - \mathbf{\Phi}\mathbf{A})^{-1}\mathbf{\Phi}$, and $\mathbf{B}_0 = (I - \mathbf{\Phi}\mathbf{A})^{-1}$. Hence, once $\mathbf{A}$ and $\mathbf{B}_1$ are identified, it is possible to solve for $\mathbf{\Gamma}$ and $\mathbf{\Phi}$ without parameter restrictions. This solution nests the solution of the previous section where FI.i is imposed. FI.i implies $E[\mathbf{z}_{t+1}|\mathcal{F}_t] = 0$ so that no variation goes through $\mathbf{B}_1$, and therefore only $\mathbf{A}$ shows up.

The second part of the solution, $\mathbf{y}_t - E[\mathbf{y}_t|\mathcal{F}_t]$, is based on information that arrives after time t. From the rational expectations model (1) it is apparent that fluctuations in $\mathbf{y}_t$ are either caused directly by fundamental changes through $\{\mathbf{y}_{t-1}, \mathbf{u}_t\}$, or by changes in beliefs through $E[\mathbf{y}_{t+1}|\mathcal{F}_t]$. Note that beliefs are fully observed when the period is realized. There are therefore no revisions in $E[\mathbf{y}_{t+1}|\mathcal{F}_t]$ in the future, only in $\{\mathbf{y}_{t-1}, \mathbf{u}_t\}$. Mathematically, this is just the law of iterated expectations, which holds under rational expectations: $E[E[\mathbf{y}_{t+1}|\mathcal{F}_t]|\mathcal{F}_{t+h}] = E[\mathbf{y}_{t+1}|\mathcal{F}_t]$, $\forall h \geq 0$. Future revisions in the representative agent's belief about $\mathbf{y}_t$ can thus only be caused directly by unobserved fluctuations in fundamentals

$\{\mathbf{y}_{t-1}, \mathbf{u}_t\}$, which agents learn about only after time t, and not by changes in beliefs, since the relevant beliefs which affected economic outcomes at time t are known in subsequent periods. Hence, the part of the model that depends on beliefs about the future cancels out:

$$(\mathbf{y}_t - E[\mathbf{y}_t|\mathcal{F}_t]) = \tilde{\mathbf{\Gamma}} (\mathbf{y}_{t-1} - E[\mathbf{y}_{t-1}|\mathcal{F}_t]) + \mathbf{\Pi} (\mathbf{u}_t - E[\mathbf{u}_t|\mathcal{F}_t]) \,, \tag{17}$$

where $\tilde{\mathbf{\Gamma}} \neq \mathbf{\Gamma}$ is possible if beliefs about the past matter, which is further discussed in appendix C.1. When agents revise their beliefs about past outcomes, it can only be because they learned since then something about past fundamentals.

Combining both parts of $\mathbf{y}_t$ according to systems (14), (15), (16), and (17):

$$\begin{aligned} \mathbf{y}_t =& \mathbf{A}\mathbf{y}_{t-1} + \sum_{h=1}^{H} (\mathbf{B}_1)^h \mathbf{B}_0 \mathbf{\Pi} E[\mathbf{u}_{t+h}|\mathcal{F}_t] + \mathbf{B}_0 \mathbf{\Pi} \mathbf{u}_t \\ &+ (\tilde{\mathbf{\Gamma}} - \mathbf{A}) (\mathbf{y}_{t-1} - E[\mathbf{y}_{t-1}|\mathcal{F}_t]) + (I - \mathbf{B}_0) \mathbf{\Pi} (\mathbf{u}_t - E[\mathbf{u}_t|\mathcal{F}_t]) \,. \end{aligned} \tag{18}$$

The first line illustrates how the economy depends on realizations $\{\mathbf{y}_{t-1}, \mathbf{u}_t\}$, but also on expectations about the future $E[\mathbf{u}_{t+h}|\mathcal{F}_t]$, $h \geq 1$. The second line adjusts for information that is gathered after $\mathbf{y}_t$ is realized. The impact of these unobserved fluctuations differ, as they don't experience any feedback from expectations.

This section solved the model without imposing assumptions on the information process. The next section shows how the parameters of the model can be identified by a researcher who has access to data on the expectations of the representative agent.

4.3 Identification of the Model

This section identifies $\mathbf{A}$ and $\mathbf{B}_1$, which solve $\mathbf{\Gamma}$, $\mathbf{\Phi}$ and $\mathbf{B}_0$. The identification strategy of section 2 no longer works to get composite parameter $\mathbf{A}$. The regression of $E[\mathbf{y}_t|\mathcal{F}_t]$ on $E[\mathbf{y}_{t-1}|\mathcal{F}_{t-1}]$ is biased, because of the confounding factor $\mathbf{z}_t$: $E[\mathbf{y}_t|\mathcal{F}_t]$ depends on $E[\mathbf{z}_t|\mathcal{F}_t]$ through $\mathbf{B}_0\mathbf{\Pi} \neq 0$, while $E[\mathbf{y}_{t-1}|\mathcal{F}_{t-1}]$ depends on $E[\mathbf{z}_t|\mathcal{F}_{t-1}]$ through $\mathbf{B}_1\mathbf{B}_0\mathbf{\Pi} \neq 0$.

A new identification strategy is introduced that makes use of fluctuation that is only observed in retrospect. Unobserved fluctuations have no impact on expectations and thus do not depend on the confounding factor. This unobserved variation is captured with *back-cast revisions*, which are defined as changes in expectations about the past, $E[\mathbf{y}_{t-k}|\mathcal{F}_t] - E[\mathbf{y}_{t-k}|\mathcal{F}_{t-1}]$, $k \geq 1$.

Formally, system (15) shows nowcast revisions depend on backcast revisions through $\mathbf{A}$:

$$E[\mathbf{y}_t|\mathcal{F}_t] - E[\mathbf{y}_t|\mathcal{F}_{t-1}] = \mathbf{A}\left(E[\mathbf{y}_{t-1}|\mathcal{F}_t] - E[\mathbf{y}_{t-1}|\mathcal{F}_{t-1}]\right) + \left(E[\mathbf{z}_t|\mathcal{F}_t] - E[\mathbf{z}_t|\mathcal{F}_{t-1}]\right), \qquad (19)$$

where backcast revisions don't depend on $\mathbf{z}_t$ according to system (17). Hence, there is no correlation between backcast revisions and the error term coming from the shocks directly, as they are uncorrelated over time, $E[\mathbf{u}_{t+h}\mathbf{u}_t] = 0$, $h \neq 0$, implies $E[\mathbf{z}_t\mathbf{u}_{t-k}] = 0$, $k \geq 1$. Section 4.1 shows however that *beliefs* about the shocks might still be correlated so that identification of $\mathbf{A}$ requires the RTR assumption. The revealed timing at realization (RTR) assumption makes regression (19) unbiased, because the error term, $E[\mathbf{z}_t|\mathcal{F}_t] - E[\mathbf{z}_t|\mathcal{F}_{t-1}]$, depends on today's and yesterday's beliefs about current and future shocks, while backcast revisions depend on today's and yesterday's beliefs about all past shocks. They are not correlated because there is no more uncertainty about timing between past shocks versus current and future shocks according to RTR. Formally, inspection of (19) shows that the regressor is uncorrelated with the error term under RTR:

$$\begin{aligned}
&E\Big[\left(E[\mathbf{y}_{t-1}|\mathcal{F}_t] - E[\mathbf{y}_{t-1}|\mathcal{F}_{t-1}]\right)\left(E[\mathbf{z}_t|\mathcal{F}_t] - E[\mathbf{z}_t|\mathcal{F}_{t-1}]\right)\Big] \\
&= E\left[\left(\sum_{s=1}^{\infty}(\tilde{\mathbf{\Gamma}})^s\left(E[\mathbf{u}_{t-s}|\mathcal{F}_t] - E[\mathbf{u}_{t-s}|\mathcal{F}_{t-1}]\right)\right)\left(\sum_{h=0}^{\infty}(\tilde{\mathbf{B}}_1)^h\tilde{\mathbf{B}}_0\tilde{\mathbf{\Pi}}\left(E[\mathbf{u}_{t+h}|\mathcal{F}_t] - E[\mathbf{u}_{t+h}|\mathcal{F}_{t-1}]\right)\right)\right] \\
&= \sum_{s=1}^{\infty}\sum_{h=0}^{\infty}(\tilde{\mathbf{\Gamma}})^s(\tilde{\mathbf{B}}_1)^h\tilde{\mathbf{B}}_0\tilde{\mathbf{\Pi}}\Big(Cov\left(E[\mathbf{u}_{t-s}|\mathcal{F}_t], E[\mathbf{u}_{t+h}|\mathcal{F}_t]\right)) - Cov\left(E[\mathbf{u}_{t-s}|\mathcal{F}_{t-1}], E[\mathbf{u}_{t+h}|\mathcal{F}_{t-1}]\right))\Big) \\
&= 0.
\end{aligned}$$

Once $\mathbf{A}$ is known, identification of $\mathbf{B}_1$ can be done in two steps. First, use $\mathbf{A}$ to extract $E[\mathbf{z}_{t+h}|\mathcal{F}_t] = E[\mathbf{y}_{t+h}|\mathcal{F}_t] - \mathbf{A}E[\mathbf{y}_{t+h-1}|\mathcal{F}_t]$ according to system (15). Second, regress $E[\mathbf{z}_t|\mathcal{F}_t]$ on $E[\mathbf{z}_{t+1}|\mathcal{F}_t]$ to get $\mathbf{B}_1$ of system (16). $\mathbf{A}$ and $\mathbf{B}_1$ are then sufficient to solve for $\mathbf{\Gamma}$ and $\mathbf{\Phi}$, and $\mathbf{B}_0 = (I - \mathbf{\Phi}\mathbf{A})^{-1}$ so that the parameters are identified. Regression (16) is unbiased because regressor and error terms are uncorrelated under RTR:

$$
\begin{aligned}
E\Big[E[\mathbf{z}_{t+1}|\mathcal{F}_t]E[\mathbf{u}_t|\mathcal{F}_t]\Big] &= E\left[\sum_{h=1}^{\infty}(\tilde{\mathbf{B}}_1)^h\tilde{\mathbf{B}}_0\tilde{\mathbf{\Pi}}E[\mathbf{u}_{t+h}|\mathcal{F}_t]E[\mathbf{u}_t|\mathcal{F}_t]\right] \\
&= \sum_{h=1}^{\infty}(\tilde{\mathbf{B}}_1)^h\tilde{\mathbf{B}}_0\tilde{\mathbf{\Pi}}Cov\left(E[\mathbf{u}_{t+h}|\mathcal{F}_t], E[\mathbf{u}_t|\mathcal{F}_t]\right) = 0.
\end{aligned}
$$

Next, I impose IDI in addition to RTR to identify the shocks of the model as well as the shock components. The variance-covariance matrix of the shock components then identifies the information structure.

IDI implies that structural shocks $\mathbf{u}_t$ can be decomposed into $K + H + 1$ news, surprise, and revise components according to system (12). Define *residual* components as linear combinations of these components: $\mathbf{\Pi}(E[\mathbf{u}_{t+s}|\mathcal{F}_t] - E[\mathbf{u}_{t+s}|\mathcal{F}_{t-1}])$, guided by the impact matrix $\mathbf{\Pi}$. Systems (16) and (17) show that revise, surprise, and news *residual* components can be extracted as follows:

$$
\begin{aligned}
E[\mathbf{y}_{t-k}|\mathcal{F}_t] - E[\mathbf{y}_{t-k}|\mathcal{F}_{t-1}] &= \tilde{\mathbf{\Gamma}}\left(E[\mathbf{y}_{t-k-1}|\mathcal{F}_t] - E[\mathbf{y}_{t-k-1}|\mathcal{F}_{t-1}]\right) \\
&\quad + \mathbf{\Pi}\left(E[\mathbf{u}_{t-k}|\mathcal{F}_t] - E[\mathbf{u}_{t-k}|\mathcal{F}_{t-1}]\right),\ 1 \le k \le K, \qquad (20) \\
E[\mathbf{z}_{t+h}|\mathcal{F}_t] - E[\mathbf{z}_{t+h}|\mathcal{F}_{t-1}] &= \mathbf{B}_1(E[\mathbf{z}_{t+h+1}|\mathcal{F}_t] - E[\mathbf{z}_{t+h+1}|\mathcal{F}_{t-1}]) \\
&\quad + \mathbf{B}_0\mathbf{\Pi}(E[\mathbf{u}_{t+h}|\mathcal{F}_t] - E[\mathbf{u}_{t+h}|\mathcal{F}_{t-1}]),\ 0 \le h \le H. \qquad (21)
\end{aligned}
$$

Parameters $\mathbf{B}_1$, $\mathbf{B}_0$, and $\mathbf{\Gamma}$, as well as unobservable $E[\mathbf{z}_{t+h}|\mathcal{F}_t]$ are identified in the previous section. If $\tilde{\mathbf{\Gamma}} \neq \mathbf{\Gamma}$, which is the case when beliefs about the past matter (see appendix C.1), system (20) can be used to identify $\tilde{\mathbf{\Gamma}}$ under RTR.

Residuals, $\mathbf{\Pi u}_t$, are identified by summing up all identified residual components:

$$\mathbf{\Pi u}_t = \sum_{k=-H}^{K} \mathbf{\Pi}\Big(E[\mathbf{u}_t|\mathcal{F}_{t+k}] - E[\mathbf{u}_t|\mathcal{F}_{t+k-1}]\Big). \tag{22}$$

Since $E[\mathbf{u}_t\mathbf{u}_t'] = I$, the covariance matrix is identified: $E[(\mathbf{\Pi u}_t)(\mathbf{\Pi u}_t)'] = \mathbf{\Pi\Pi'}$.

Identification of shocks $\mathbf{u}_t$ requires restrictions on the impact matrix $\mathbf{\Pi}$. Identified covariance matrix $\mathbf{\Pi\Pi'}$ provides $\frac{(N+1)N}{2}$ out of N^2 parameters so that $\frac{(N-1)N}{2}$ additional restrictions are necessary. Constraints can be derived from dynamic stochastic general equilibrium (DSGE) models, directly, or from restrictions on impulse response functions. Section 5 discusses the different strategies to get appropriate identifying restrictions. Identification of $\mathbf{\Pi\Pi'}$ is sufficient to quantify the importance of availability of information (see section 6.3), while identification of the impact matrix $\mathbf{\Pi}$ is necessary for impulse response analyses (see section 6.4).

The information structure is characterized by the $N(H+K+1)\times N(H+K+1)$ variance-covariance matrix $\mathfrak{D}$ according to system (13), which is derived using the identified shock components. Remember that there is uncertainty about type so that the shock components, $E[u_{i,t+h}|\mathcal{F}_t] - E[u_{i,t+h}|\mathcal{F}_{t-1}]$, might be correlated across i. In order to orthogonalize across i, one needs to identify information matrix $\mathbf{D}$, where $\mathbf{DD'} = \mathfrak{D}$. Identifying restrictions on $\mathbf{D}$ are, for example, whether agents confuse demand shocks with supply shocks, the other way around, or if they confuse demand shocks with supply shocks as often as they confuse supply shocks with demand shocks.

To summarize, section 4 relaxes FI by introducing a non-parametric description of how information diffuses over time, then it solves the model despite not having a parametric expression for the information process, and finally, section 4 uses data on expectations to identify parameters, shocks, and information structure of the rational expectations model.

5 Identifying Restrictions on Shocks

The previous section shows how one can identify rational expectations models under more general information structures than normally considered in standard macroeconomic models. The assumptions of section 4 are however not sufficient to identify impact matrix $\mathbf{\Pi}$, which is required to identify the shocks of the model. Identified variance-covariance matrix $\mathbf{\Pi\Pi}'$ provides some information so that only $\frac{(N-1)N}{2}$ additional restrictions are necessary to get a unique impact matrix $\mathbf{\Pi}$.

The next two sections provide strategies to extract meaningful shocks $\mathbf{u}_t$ from identified residuals $\mathbf{\Pi u}_t$. Section 5.1 uses DSGE models to find restrictions on $\mathbf{\Pi}$, while section 5.2 uses restrictions on impulse responses as it is common for structural VARs.

5.1 Restrictions from DSGE Models

The rational expectations model (1) can be written as follows:

$$\mathbf{\Pi}^{-1}\mathbf{y}_t = \mathbf{\Pi}^{-1}\mathbf{\Gamma}\mathbf{y}_{t-1} + \mathbf{\Pi}^{-1}\mathbf{\Phi}E[\mathbf{y}_{t+1}|\mathcal{F}_t] + \mathbf{u}_t, \tag{23}$$

which can be related to a system of equations derived by a DSGE model. Theory might provide zero or sign restrictions on $\mathbf{\Pi}^{-1}$, $\mathbf{\Pi}^{-1}\mathbf{\Gamma}$, or $\mathbf{\Pi}^{-1}\mathbf{\Phi}$. Note that only a few zero restrictions are necessary, for example, in a bivariate model only $\frac{(2-1)2}{2} = 1$ restriction, and in a model with three equations only $\frac{(3-1)3}{2} = 3$ restrictions.

For illustrative purpose, consider a standard New-Keynesian DSGE model with a Philips curve (24), a Taylor rule (25), and an IS-curve 26:

$$\pi_t = \omega E_t \pi_{t+1} + (1-\omega)\pi_{t-1} + \lambda y_t + \sigma_1 u_{1t}, \tag{24}$$

$$y_t = E_t y_{t+1} - \frac{1}{\sigma}(r_t - E_t \pi_{t+1}) + \sigma_2 u_{2t}, \tag{25}$$

$$i_t = \rho i_{t-1} + (1-\rho)(\gamma \pi_t + \eta y_t + \psi(y_t - y_{t-1})) + \sigma_3 u_{3t}. \tag{26}$$

This NK-model imposes thirteen zero restrictions on the rational expectations model:

$$
\begin{aligned}
\mathbf{\Pi}^{-1}\mathbf{y}_t &= \mathbf{\Pi}^{-1}\mathbf{\Gamma}\mathbf{y}_{t-1} + \mathbf{\Pi}^{-1}\mathbf{\Phi}E[\mathbf{y}_{t+1}|\mathcal{F}_t] + \mathbf{u}_t, \\
\begin{pmatrix} 1 & \kappa_{12} & 0 \\ 0 & 1 & \kappa_{23} \\ \kappa_{31} & \kappa_{32} & 1 \end{pmatrix} \begin{pmatrix} \pi_t \\ y_t \\ i_t \end{pmatrix} &= \begin{pmatrix} \rho_{11} & 0 & 0 \\ 0 & \rho_{22} & 0 \\ 0 & \rho_{32} & \rho_{33} \end{pmatrix} \begin{pmatrix} \pi_{t-1} \\ y_{t-1} \\ i_{t-1} \end{pmatrix} \\
&+ \begin{pmatrix} \psi_{11} & 0 & 0 \\ \psi_{21} & \psi_{22} & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} E_t[\pi_{t+1}] \\ E_t[y_{t+1}] \\ E_t[i_{t+1}] \end{pmatrix} + \begin{pmatrix} u_{1t} \\ u_{2t} \\ u_{3t} \end{pmatrix}.
\end{aligned} \tag{27}
$$

Identification of impact matrix $\mathbf{\Pi}$ only requires three restrictions out of the thirteen so that the NK model provides more than enough restrictions to identify the shocks. Hence, the challenge is not finding restrictions, but rather choosing the right restrictions that clearly separate the different equations.

5.2 Restrictions from Impulse Response Functions

An alternative strategy is to impose restrictions on the impulse response function as it is common in the structural VAR literature. Once the model is identified, solution 18 can be used to calculate IRFs for different $\mathbf{\Pi}$ until the desired short- or long-run, and zero- or sign-restrictions are satisfied. While shocks in structural VARs always produce the same impulse responses, the responses of the rational expectations model depend on how much information is available about the shock before, at, and after it is realized. This can be accommodated by imposing impulse response restrictions on the average response, which is the average response to shocks across time, where information sometimes arrive earlier and sometimes later (see definition in appendix A.3). Appendix A.3 shows that the average response is equal to the response to the *average shock*, which is defined as the shock which realization is learned with the same speed as the agent learns on average. Calculation of impulse response to the average shock does not require simulation, which makes identification straight-forward.

The model does not only identify shocks $\mathbf{u}_t$, but it also identifies the shocks' revise, surprise, and news components according to system (12). Instead of imposing restrictions on the response to the sum of all components, $\mathbf{u}_t$, one could also impose restrictions on the response to specific components, $E[\mathbf{u}_t|\mathcal{F}_{t-k}] - E[\mathbf{u}_t|\mathcal{F}_{t-k-1}]$, for a given k. For example, if u_{it} is a monetary policy shock, one could impose that when information about the shock comes as surprise, $E[u_{it}|\mathcal{F}_{t-k}] - E[u_{it}|\mathcal{F}_{t-k-1}] = u_{it}$ for $k = 0$ and zero otherwise, the impact on output is zero within the quarter. Another example is to set the impact of a TFP shock u_{jt} on TFP itself equal to zero before it is realized. Hence, the impact of a TFP news component, $E[u_{jt}|\mathcal{F}_{t-k}] - E[u_{it}|\mathcal{F}_{t-k-1}] = u_{it}$ for $k = 1$ and zero otherwise, can be restricted to have no impact on TFP before time t. These timing restrictions are more precise so that the impact matrix $\mathbf{\Pi}$ can be identified using weaker assumptions.

The examples of this section show how identifying restrictions can be found in many places. The application of section 6.4 will use a traditional approach following Blanchard and Quah (1989), by separating demand and supply shocks with a zero long-run restriction on the response of real GDP to an average demand shock.

6 Application

This section discusses data on expectations with a new metric of information gain in section 6.1, estimates models in section 6.2, and uses estimates for the following two applications. Section 6.3 quantifies the effect of availability of information with the moments of the data. I find that early arrival of information makes data more persistent and more pro-cyclical. Section 6.4 identifies impulse responses to demand and supply shocks and identifies how well these shocks are observed on average. I find that almost half of the variance of supply shocks are only observed after the shocks are realized, and the effect of information is stronger for supply than for demand shocks.

6.1 Forecaster Data

For applications, I use data on expectations from the Greenbooks, which are provided by the Federal Reserve Bank of Philadelphia. These back-, now, and forecasts are prepared for meetings of the Federal Open Market Committee (FOMC), and are available to the public with a delay of five years. I aggregate the forecasts to a quarterly series by taking average, if there is more than one meeting per quarter.

The Federal Reserve is likely the most informed economic agent in the economy, which is desirable for identification in this paper. This is because Fed's backcast revisions are likely caused by fluctuations that are also not observed by other agents so that there should be no feedback coming from expectations to this unobserved variation. An additional advantage of using Greenbooks is the wealth of forecasts across variables and horizons, which is useful for implementation.

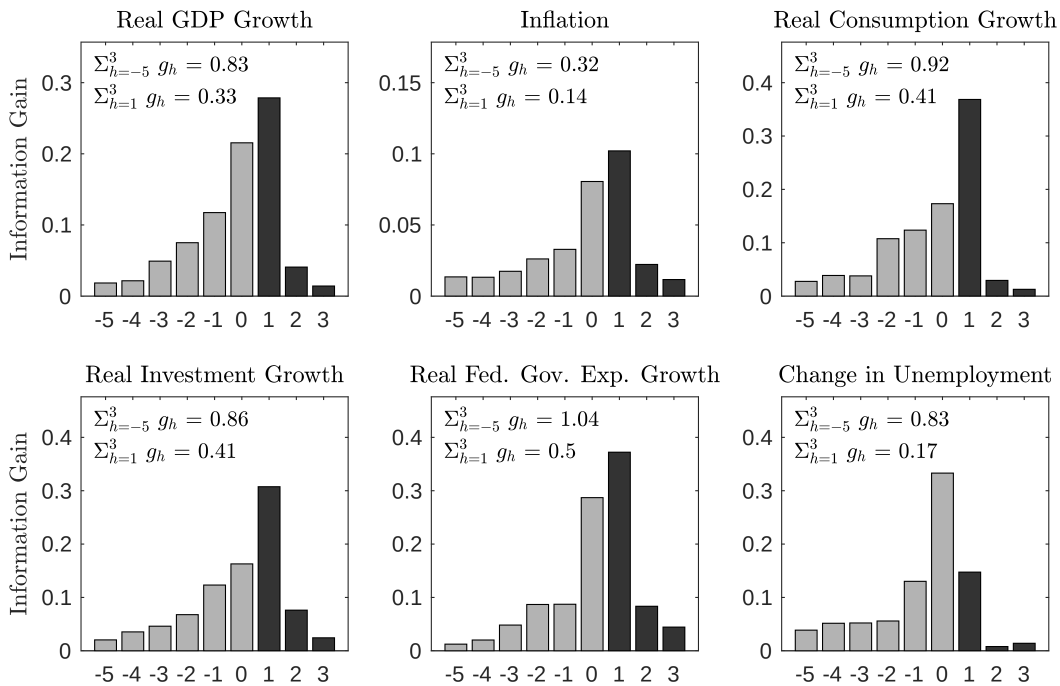
The key identification condition for my proposed methodology, namely that agents revise their views not just about the future but also about the past, can be quantified directly through the following metric of information gain g_h:

$$g_h \equiv \frac{Var\left(E[y_t|\mathcal{F}_{t+h}] - E[y_t|\mathcal{F}_{t+h-1}]\right)}{Var(y_t)} \in [0,1]. \tag{28}$$

This metric measures the share of the variance of a variable y_t that is learned h periods after its realization, or $-h$ periods before its realization.

Under FI, information gain g_h would be zero for all $h \geq 0$, as everything is observed when it is realized. For $h \leq -1$, the only reason why g_h would not be equal to zero under FI is that the system is persistent so that information about today propagates into the future and thus changes expectations about the future. Under FI, these changes would reflect one to one the changes in expectations about today, which is why Manski's (1993) reflection problem

Figure 2: Information Gain



Notes: This figure displays information gain g_h for six macro indicators. The metric is defined in equation (28) and it measures the share of variance that is learned h quarters after realization, respectively, -h quarters before realization, where h is plotted on the x-axis. The share of the variance that is learned between 5 periods before and 3 periods after realization is equal to the sum of the gains, $\Sigma^3_{h=-5} g_h$, which is displayed at the top left in each plot. Right below that is the share of the variance that is learned after realization, $\Sigma^3_{h=1} g_h$. Information gain is measured using quarterly data on expectations from the Greenbook, 1978Q3 - 2011Q4. Inflation is defined as GDP price deflator, consumption as personal consumption expenditures, investment as business fixed investment, government expenditures as federal government consumption and gross investment.

occurs. When relaxing FI.ii, there is information gain after realization, $g_h \neq 0$ for $h \geq 0$, so that people learn about a period even after they experience that period. Relaxing FI.i means that people learn about the future from news shocks directly, and not just indirectly from propagating what they learn about today into the future. Hence, this metric provides a simple quantitative metric for assessing when information about different macroeconomic variables arrives and is processed by agents.

Full revelation, $\lim_{s\to\infty} E_{t+s}[y_t] = y_t$, together with RE imply $\sum_{h=-\infty}^{\infty} g_h = 1$. The share of the variance that is only observed after realization is equal to $\sum_{h=1}^{\infty} g_h$. Figure 2 plots information gain observed in Greenbook forecasts. Strikingly, around one third of real GDP growth is only observed after realization. This contradicts the full information assumption. Since this is a lower bound for missing information, as the Fed is one of the best forecaster in the U.S. economy, this finding provides particularly strong support for relaxing FI.

6.2 Estimation

For estimation, consider a more general version of model (1):

$$\mathbf{y}_t = \mathbf{q}_t + \sum_{p=1}^{P} \mathbf{\Gamma}_p \mathbf{y}_{t-p} + \sum_{r=1}^{R} \mathbf{\Phi}_r E[\mathbf{y}_{t+r}|\mathcal{F}_t] + \mathbf{\Pi}\mathbf{u}_t, \tag{29}$$

where $\mathbf{q}_t$ are N trend components that are observed in advance, $E[\mathbf{q}_t|\mathcal{F}_{t-k}] = \mathbf{q}_t$, for all k, the same way as model parameters are observed in advance. These trend components take care of changes in the first moments of the data over time. This is necessary as for some variables, e.g. inflation, far horizon forecasts don't converge to a constant mean, but rather converge to a mean that changes over time. Model 29 also extends the number of lead and lag terms from one to P lag variables, $\{\mathbf{y}_{t-1}, ..., \mathbf{y}_{t-P}\}$, and R lead terms, $\{E[\mathbf{y}_{t+1}|\mathcal{F}_t], ..., E[\mathbf{y}_{t+R}|\mathcal{F}_t]\}$.

I will impose the same assumptions as discussed in section 4. Namely, I assume rational expectations (RE), but relax full information (FI), and I restrict the information structure so that there is no uncertainty about timing after shocks are realized (RTR), and information about shocks diffuse H periods before until K periods after realization (IDI).

Estimation is based on the variance-covariance matrix of the revisions observed in data on expectations, $E[\mathbf{y}_{t+h}|\mathcal{F}_t] - E[\mathbf{y}_{t+h}|\mathcal{F}_{t-1}]$, for all $h \in \{-K, ..., H\}$, where K and H are the same limits as assumed under IDI. Hence, there is no information gathered about shocks beyond horizon $H = 4$, which is reasonable according to figure 2, as information gains are close to zero for revisions further than three quarters before impact. I consider the one quarter backcasts as truth by setting $K = 1$ so that $E[\mathbf{y}_{t-1}|\mathcal{F}_t] = \mathbf{y}_{t-1}$. The reason is that backcast revisions about the further past are infrequent and sometimes capture changes in definitions rather than unobserved fluctuations. Revisions are uncorrelated across t under RE, by law of iterated expectations. The generalized model with $P = H + 1$, and $R = H$ is exactly identified so that the model can fully match the variance-covariance matrix.

Similar to the more restricted model in section 4, appendix B.3 shows that the model can be solved in terms of $\{\mathbf{A}_p\}_{p=1}^P$, $\{\mathbf{B}_r\}_{r=1}^R$, and $\mathbf{B}_0$, which are invertible functions of $\{\mathbf{\Gamma}_p\}_{p=1}^P$ and $\{\mathbf{\Phi}_r\}_{r=1}^R$. For illustrative purpose, let $K = 1$ and $H = 3$. The covariance matrix of revisions in expectations can be decomposed into the following elements:

$$\mathbf{R} \equiv E[\tilde{\mathbf{Y}}_t\tilde{\mathbf{Y}}_t'] = (\mathcal{A}\mathcal{B}_1\mathcal{B}_0\mathcal{C})\mathfrak{D}(\mathcal{A}\mathcal{B}_1\mathcal{B}_0\mathcal{C})', \quad \tilde{\mathbf{Y}}_t \equiv \begin{pmatrix} E[\mathbf{y}_{t-1}|\mathcal{F}_t] - E[\mathbf{y}_{t-1}|\mathcal{F}_{t-1}] \\ E[\mathbf{y}_{t}|\mathcal{F}_t] - E[\mathbf{y}_{t}|\mathcal{F}_{t-1}] \\ E[\mathbf{y}_{t+1}|\mathcal{F}_t] - E[\mathbf{y}_{t+1}|\mathcal{F}_{t-1}] \\ E[\mathbf{y}_{t+2}|\mathcal{F}_t] - E[\mathbf{y}_{t+2}|\mathcal{F}_{t-1}] \\ E[\mathbf{y}_{t+3}|\mathcal{F}_t] - E[\mathbf{y}_{t+3}|\mathcal{F}_{t-1}] \end{pmatrix}, \tag{30}$$

where appendix B.4 shows that $\mathcal{A}$ is function of $\{\mathbf{A}_p\}_{p=1}^P$, $\mathcal{B}_1$ of $\{\mathbf{B}_r\}_{r=1}^R$, $\mathcal{B}_0$ of $\mathbf{B}_0$, $\mathcal{C}$ of $\mathbf{\Pi}$, and $\mathfrak{D}$ is defined in (13). Properties of the shocks, $E[\mathbf{u}_t\mathbf{u}_{t+h}] = 0$, $\forall h \neq 0$, $E[\mathbf{u}_t\mathbf{u}_t'] = I$, together with RTR and IDI imply the following covariance matrix of shock components:

$$\mathfrak{D} \equiv \begin{pmatrix} \mathfrak{D}_{-1,-1} & 0 & 0 & 0 & 0 \\ 0 & \mathfrak{D}_{00} & \mathfrak{D}_{01} & \mathfrak{D}_{02} & 0 \\ 0 & \mathfrak{D}_{01} & \mathfrak{D}_{11} & \mathfrak{D}_{12} & -\mathfrak{D}_{02} \\ 0 & \mathfrak{D}_{02} & \mathfrak{D}_{12} & \mathfrak{D}_{22} & -\sum_{i=0}^{1} \mathfrak{D}_{i,i+1} \\ 0 & 0 & -\mathfrak{D}_{02} & -\sum_{i=0}^{1} \mathfrak{D}_{i,i+1} & I - \sum_{i=0}^{2} \mathfrak{D}_{ii} \end{pmatrix}. \tag{31}$$

There are enough restrictions in the variance-covariance matrix of the shock components to exactly identify all the model parameters from matrix $\mathbf{R}$, which can be done with a minimization algorithm, or more efficiently with matrix calculations (see appendix B.4). Inference is performed by bootstrapping the covariance matrix of revisions, $\mathbf{R}$, and re-estimating the model for each draw.

6.3 Counterfactual Moments

This section compares the effect of shocks on business cycle fluctuations when they are observed ahead of time to shocks that are only observed ex-post. If both types of shocks generate the same characteristics, the null that information and thus expectations matter can be rejected. Information enters the model exogenously, hence, according to Lucas's (1976) critique, this exercise does not reveal whether a policy that aims to provide more or less information is effective, because there might be a change in behaviour in response to such policy, which would then change the parameters of the model. There is however still a lot to learn about effects of expectations from quantifying the difference between early and late arrival of information.

Arrival of information is measured using the same metric of information gain (28) on shocks u_{it} so that g_h is the i^{th} diagonal element in $\mathfrak{D}_{hh}$. Counterfactuals are generated by manipulating the information structure, $\mathfrak{D}$, while still maintaining the properties of the shocks discussed in section 2.5. Note that the counterfactual exercise does not require identification of the impact matrix $\mathbf{\Pi}$ of section 5, because the moments do not depend on how the residuals are decomposed. The only specification needed is the set of variables included in $\mathbf{y}_t$.

I estimate counterfactuals in a model with real GDP growth, real consumption growth, real investment growth, and inflation. I chose these four variables as they should capture the state of the economy in both, real business cycle (RBC) models and New Keynesian (NK) models at business cycle frequencies. NK models generally include interest rates, but they are fully observed in real time so that the identification strategy does not apply. If the interest rate is a state variable, estimation without it is fine as long as the four variables and expectations thereof capture the state of the interest rate. RBC models generally include total factor productivity (TFP), but the Greenbook does not provide expectations thereof so that the same idea applies for TFP as for interest rates. Counterfactuals are estimated with model (29) using five lag and four lead terms, based on quarterly Greenbook forecasts from 1978Q3 to 2011Q4. The moments are calculated for detrended data, hence, observed changes in means are subtracted from the variables, $\mathbf{y}_t - \mathbf{q}_t$, before calculating the moments, where $\mathbf{q}_t$ is defined in model (29).

Table 1 shows the moments of the data, and six counterfactuals. The first column lists the moments without changing the arrival of information. The first six rows describe when information arrives by calculating how much of the variance of an average shock is learned at $t-4$, $t-3$, up until $t+1$, where t is the time when shocks are realized. About half the information is available before the shocks are realized $(0.07 + 0.04 + 0.07 + 0.12 + 0.25 = 0.54)$, whereas the other half is available after realization (0.46). The second column simulates a *delayed info* counterfactual, where information is only gathered after the shocks are realized so that all the variance of the shocks is learned ex-post, $\mathfrak{D}_{-1-1} = I$. The *full info* counterfactual of column three simulates an economy with the standard full information (FI) assumption so that surprise components are the only non-zero components, and have a variance of one, $\mathfrak{D}_{00} = I$. The *news* counterfactuals of columns four to seven list the

Table 1: Moments of Data when Changing Arrival of Information

Variance Observed	Data	Delayed Info	Full Info	News			
$t-4$	0.07	0	0	0	0	0	1.00
$t-3$	0.04	0	0	0	0	1.00	0
$t-2$	0.07	0	0	0	1.00	0	0
$t-1$	0.12	0	0	1.00	0	0	0
t	0.25	0	1.00	0	0	0	0
$t+1$	0.46	1.00	0	0	0	0	0
Real GDP Growth (y_t), detrended							
$\sigma(y_t)$	2.55	2.46	2.33	2.41	2.68	3.04	3.47
Auto(1)	0.43	0.32	0.31	0.46	0.51	0.55	0.61
Real Consumption Growth, detrended							
$\sigma(\cdot)$	2.73	2.90	2.29	2.43	2.69	2.76	3.34
Auto(1)	0.15	0.01	0.05	0.20	0.26	0.34	0.42
$\rho(\cdot, y_{t-1})$	0.34	0.30	0.31	0.21	0.28	0.40	0.46
$\rho(\cdot, y_t)$	0.61	0.51	0.64	0.65	0.67	0.80	0.84
$\rho(\cdot, y_{t+1})$	0.25	0.09	0.16	0.29	0.39	0.50	0.53
Real Investment Growth, detrended							
$\sigma(\cdot)$	8.68	8.06	7.79	9.10	8.80	10.20	11.17
Auto(1)	0.35	0.17	0.23	0.59	0.63	0.65	0.65
$\rho(\cdot, y_{t-1})$	0.08	0.04	0.05	0.09	0.11	0.14	0.17
$\rho(\cdot, y_t)$	0.16	0.13	0.16	0.16	0.18	0.19	0.20
$\rho(\cdot, y_{t+1})$	0.13	0.11	0.11	0.12	0.13	0.15	0.18
Inflation (GDP Deflator), detrended							
$\sigma(\cdot)$	1.44	1.24	1.06	1.12	1.40	1.42	2.67
Auto(1)	0.31	0.17	0.16	0.33	0.42	0.44	0.29
$\rho(\cdot, y_{t-1})$	-0.33	-0.13	-0.15	-0.04	-0.29	-0.14	-0.61
$\rho(\cdot, y_t)$	-0.30	-0.23	-0.32	-0.28	-0.18	-0.08	-0.35
$\rho(\cdot, y_{t+1})$	-0.19	-0.06	-0.05	-0.15	-0.06	-0.09	-0.17

Notes: This table shows the moments of the data for different assumptions on the arrival of information. The first column lists the moments of detrended data, $\mathbf{y}_t - \mathbf{q}_t$, where $\mathbf{q}_t$ is the trend of system (29). The remaining columns show counterfactual moments by shifting the arrival of information from one quarter after realization (column 2) up to four quarters before realization (column 7). The moments of interest are standard deviation, $\sigma(\cdot)$, autocorrelation, $Auto(1)$, and correlations to detrended output growth with lag $\rho(\cdot, y_{t-1})$, without lag, $\rho(\cdot, y_t)$, and with a lead, $\rho(\cdot, y_{t+1})$. Trend and counterfactual moments are estimated based on model (29) with five lag and four lead terms, using quarterly data on expectations from the Greenbook, 1978Q3-2011Q4. The first six rows of column 1 list the average share of variances that is learned at time $t-k$, where t is the period when shocks are realized: $\overline{Var}(E[\mathbf{u}_t|\mathcal{F}_{t-k}] - E[\mathbf{u}_t|\mathcal{F}_{t-k-1}])$. In the *delayed info* economy of column 2, the arrival of information is shifted to one period after realization so that information is only available ex-post. Column 3 displays the *full info* economy where all the information about shocks is learned at time of realization. The remaining four columns list the moments of *news* economies, where all the information arrives one quarter ahead of time (column 4), up to four quarters ahead of time (column 7). The counterfactual moments are generated by manipulating the variance-covariance matrix of the shock components, $\mathfrak{D}$, so that for example in column 7: $Var(E[\mathbf{u}_t|\mathcal{F}_{t-4}] - E[\mathbf{u}_t|\mathcal{F}_{t-5}]) = I$. Price level is measured as GDP price deflator, consumption as personal consumption expenditures, and investment as business fixed investment. Units are annualized percentage points.

moments when all the information is gathered one to four periods ahead of time so that all the weights are put on the news components, $\mathfrak{D}_{hh} = I$, for $h = 1$ up to $h = 4$.

Table 1 reveals that whether information is available has economically significant implications. For example, the ability to respond to economic shocks ahead of time seems to increase persistence, measured as autocorrelation. In particular, the autocorrelations of output, consumption, and investment growth increase when moving from the *delayed info* to the *news* counterfactuals. Moreover, available information about the future seems to increase the overall variance, which suggests that information amplifies the effect of economic shocks.

Table 1 shows that the relationships between output, consumption, and investment growth become stronger the more information is available; the variables are more pro-cyclical in the *news* counterfactuals compared to the *delayed info* counterfactual, i.e. $\rho(\cdot, y_t)$ increases with information. Expectations thus seem to contribute to the observed co-movement of macro series. According to table 1, investment growth is a leading variable, $\rho(\cdot, y_{t-1}) = 0.08 < 0.13 = \rho(\cdot, y_{t+1})$, while consumption growth lags output growth, $\rho(\cdot, y_{t-1}) = 0.34 > 0.25 = \rho(\cdot, y_{t+1})$. This difference seems to be driven by lack of information, only, because once information is available ahead of time, this leading and lagging behaviour no longer occurs: in the *news* counterfactual where information is available four quarters ahead of time, investment barely leads output growth anymore, $\rho(\cdot, y_{t-1}) = 0.17 < 0.18 = \rho(\cdot, y_{t+1})$, while consumption growth slightly leads output growth, $\rho(\cdot, y_{t-1}) = 0.46 < 0.53 = \rho(\cdot, y_{t+1})$, instead of responding with a delay.

The changes in the moments are more noisy for inflation. Inflation is countercyclical, but less so when information is available one to three quarters ahead of time. In particular, table 1 shows that the correlation of inflation and output growth is more negative in the actual economy of column 1, $\rho(\cdot, y_t) = -0.30$, than in the counterfactual economy of column 6, when information is available three quarters ahead of time, $\rho(\cdot, y_t) = -0.08$. The response of expectations to information about the near future thus seems to move inflation in the same direction as output growth, suggesting that this information causes a response on the demand side of the economy. However, if information about the future is available beyond horizon three, inflation becomes more countercyclical again, $\rho(\cdot, y_t) = -0.35$, suggesting that the supply side responds, too, as long as information is available far enough ahead of time.

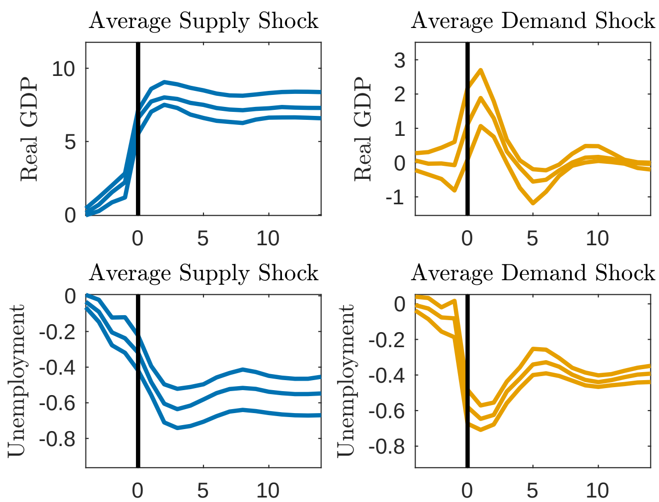
To summarize, three stylized facts show that information and therefore expectations have real effects on the economy. First, effects of information contribute to the variance of macro variables, second, they increase persistence of shocks, and third, they strengthen interdependences across variables. Hence, the exercise justifies the effort of including expectations in dynamic models, which is a common practice in the field of macroeconomics. Importantly, this conclusion is not drawn from a model where expectations need to matter by construction, in order to match persistence of the data. Instead, the model is general enough to replicate the persistence of the data without relying on effects of expectations. In other words, the model is not subject to Manski's reflection problem so that the importance of information is identified rather than assumed.

6.4 Impulse Responses

This section identifies economic shocks, how well they are observed, and their average impulse responses. In particular, the aim is to separate demand from supply shocks. Economic shocks are extracted from identified residuals, $\mathbf{\Pi u}_t$, but additional identifying restrictions are necessary to make the impact matrix unique, i.e. the variance-covariance matrix $\mathbf{\Pi\Pi}'$ together with identifying restrictions should provide a unique solution for $\mathbf{\Pi}$. Section 5 proposes different strategies to get restrictions from DSGE models or alternatively, from impulse responses directly. I will use impulse response functions to separate demand and supply shocks, based on the same idea as Blanchard and Quah (1989): supply shocks are the only shocks with long run impact on real GDP, while demand shocks have only a temporary effects on output. Similar to Blanchard and Quah (1989), I will separate demand from supply shocks in a bivariate system with real GDP growth and changes in unemployment rate.

Model (29) is estimated with five lag and four lead terms using quarterly Greenbook forecasts from 1969Q1-2011Q4. In contrast to Blanchard and Quah (1989), the model estimated in this paper differentiates between shock components observed before, at, or after the shock hits the economy. Depending on when information arrives, a shock has a stronger or a weaker effect on the economy. The zero restriction on the long run effect of a demand shock on real GDP is thus imposed on the *average* response according to (35).

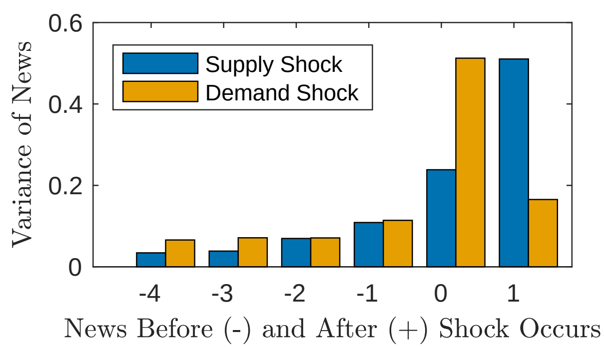
Figure 3: Impulse Response Functions



Notes: This figure displays responses of unemployment and real GDP to average supply and demand shocks, as obtained from system (29) together with a zero restriction on the long run impact of real GDP in response to an average demand shock. The model is estimated based on a backcast, a nowcast, and five forecasts, and it includes five lag and four lead terms. Data for estimation are quarterly Greenbook forecasts from 1969Q1-2011Q4. 68% confidence bands are estimated using bootstrap method. Units of real GDP are cumulated annualized percentage points, and unemployment is measured in percentage points..

Figure 3 plots the average impulse responses to demand and supply shocks. As imposed by the zero long run restriction, real GDP increases in response to an average demand shock, but then converges back to zero fairly quickly, while the average supply shock increases real GDP permanently. While Blanchard and Quah (1989) find that supply shocks lead to increased unemployment on impact, the impulse responses of figure 3 suggest that both demand and supply shocks reduce unemployment rates at all horizons, consistent with standard RBC models. Interestingly, the shocks cause significant responses even before they are realized. Information gathered ahead of time thus have real effects on the economy, by changing behaviour and decisions made in preparation of this future shock. This finding is consistent with the literature on news shocks started by Beaudry and Portier (2004, 2006) and extended by Barsky and Sims (2011), Beaudry, Dupaigne and Portier (2011), Beaudry and Portier (2014), and Barsky, Basu and Lee (2015) among others, who find that TFP shocks have real effects even before they are realized. These models use the assumption that identified TFP news shocks are indeed shocks that only realize in the future, rather than shocks that realize today. Hence, these models rely on assumptions to differentiate the direct effect of a shock, from an indirect effect of observing that shock (Manski's (1993) reflection problem). While relying on a different set of assumptions, the model estimated here can identify whether shocks are indeed realized or not so that this assumption is not necessary. Discovering significant responses before realization with a model where data alone identifies effects of expectations thus provides robust support in favour of the news shock literature.

Figure 4: Information Gain of Shocks



Notes: This figure displays information gain g_h for demand and supply shocks. The metric is defined in equation (28) and it measures the share of variance that is learned h quarters after realization, respectively, -h quarters before realization, where h is plotted on the x-axis. The metric is equal to the variance of the shock components, $diag(\mathfrak{D}_{hh}) = Var(E[\mathbf{u}_t|\mathcal{F}_{t-h}] - E[\mathbf{u}_t|\mathcal{F}_{t-h-1}])$, as obtained from system (29) together with a zero restriction on the long run impact of real GDP in response to an average demand shock. The model includes five lag and four lead terms, and is estimated using quarterly Greenbook forecasts on real GDP growth and changes in unemployment from 1969Q1-2011Q4.

The effect of expectations about the future seems to be stronger for supply shocks than for demand shocks, suggesting that either expectations matter more for supply shocks, or that there is more information available for supply shocks than for demand shocks. Figure 4 shows that the former is the case. The figure plots how much information on supply and demand shocks is learned before, at, and after impact. Demand shocks are better observed, while a large share of the variance of supply shocks is only observed ex-post. This provides evidence that the demand side is well understood, but there is little consensus on the production side. The finding that supply shocks cause a significant response ahead of time suggests that supply shocks come along with large adjustments ahead of time, as for example investments in new infrastructure.

7 Conclusion

Forward-looking behavior is one of the reasons why macro variables are correlated over time according to rational expectations (RE) models. I show that in the standard estimation framework, the extent to which this persistence is driven by forward-looking behavior is a result of the underlying model rather than a feature of the data. I relate this issue to Manski's (1993) reflection problem of social interaction models, where the researcher cannot distinguish whether expectations about a group cause individual behavior or whether they just reflect individual behavior without causing it. FIRE models suffer from the same reflection problem, as the researcher cannot distinguish whether expectations about the future cause current behavior or whether they just reflect current behavior without causing it.

I use data on expectations at multiple horizons to separately identify the forward- and backward-looking components of the RE model. I first generalize the RE model by relaxing the full information (FI) assumption, which allows me to match data on expectations at all horizons. Specifically, information about a shock is not only revealed at realization as it is the case under FI, but it also diffuses before and after the shock is realized. This generalization is based on the literature on news shocks (see Beaudry and Portier, 2004, 2006), where information arrives ahead of time, and the literature on information rigidities (see Mankiw and Reis, 2002; Woodford, 2003; Sims, 2003), where information may arrive after realization. In this generalized framework, expectations about the future are no longer proportional to current and past realizations, because the agent receives separate information about the future. Hence, the generalized model no longer suffers from the reflection problem.

Identification makes use of the timing assumption that expectations about the future can only depend on information that is available today. Revisions in expectations about the past are therefore purely backward-looking and identify how expectations reflect actions. Once this direct effect is identified, the difference between projections and observed forecasts is used to identify the feedback of expectations to actions. Applying the new approach to U.S. Greenbook forecasts, I find that persistence and co-movement of macro series depend significantly on how information diffuses over time. Since information can only cause macro movements when expectations matter, this finding supports the standard assumption that expectations are relevant for business cycle fluctuations. Moreover, based on Blanchard and Quah's (1989) decomposition, I find that supply shocks are subject to more information frictions than demand shocks. Finally, consistent with RBC models, both demand and supply shocks increase output and decrease unemployment.

References

Adam, Klaus, and Mario Padula. 2011. "Inflation dynamics and subjective expectations in the United States." *Economic Inquiry*, 49(1): 13–25.

Barsky, Robert B., and Eric R. Sims. 2011. "News shocks and business cycles." *Journal of Monetary Economics*, 58(3): 273–289.

Barsky, Robert B., and Eric R. Sims. 2012. "Information, Animal Spirits, and the Meaning of Innovations in Consumer Confidence." *American Economic Review*, 102(4): 1343–77.

Barsky, Robert B., Susanto Basu, and Keyoung Lee. 2015. "Whither News Shocks?" *NBER Macroeconomics Annual*, 29(1): 225–264.

Beaudry, Paul, and Franck Portier. 2004. "An exploration into Pigou's theory of cycles." *Journal of Monetary Economics*, 51(6): 1183 – 1216.

Beaudry, Paul, and Franck Portier. 2006. "Stock Prices, News, and Economic Fluctuations." *American Economic Review*, 96(4): 1293–1307.

Beaudry, Paul, and Franck Portier. 2014. "News-Driven Business Cycles: Insights and Challenges." *Journal of Economic Literature*, 52(4): 993–1074.

Beaudry, Paul, Martial Dupaigne, and Franck Portier. 2011. "Modeling news-driven international business cycles." *Review of Economic Dynamics*, 14(1): 72 – 91. Special issue: Sources of Business Cycles.

Binder, Michael, and M. Hashem Pesaran. 1997. "Multivariate Linear Rational Expectations Models: Characterization of the Nature of the Solutions and Their Fully Recursive Computation." *Econometric Theory*, 13(6): 877–888.

Blanchard, Olivier Jean, and Danny Quah. 1989. "The Dynamic Effects of Aggregate Demand and Supply Disturbances." *American Economic Review*, 79(4): 655–73.

Blanchard, Olivier J., Jean-Paul L'Huillier, and Guido Lorenzoni. 2013. "News, Noise, and Fluctuations: An Empirical Exploration." *American Economic Review*, 103(7): 3045–70.

Blundell, Richard, and Stephen Bond. 2000. "GMM Estimation with persistent panel data: an application to production functions." *Econometric Reviews*, 19(3): 321–340.

Bullard, James. 2018. "R-Star Wars: The Phantom Menace." Presentation at the 34th Annual National Association for Business Economics (NABE) Economic Policy Conference. Washington, D.C.

Christiano, Lawrence J., Cosmin Ilut, Roberto Motto, and Massimo Rostagno. 2010. "Monetary policy and stock market booms." *Proceedings - Economic Policy Symposium - Jackson Hole*, 85–145.

Clarida, Richard, Jordi Galí, and Mark Gertler. 2000. "Monetary Policy Rules And Macroeconomic Stability: Evidence And Some Theory." *The Quarterly Journal of Economics*, 115(1): 147–180.

Coibion, Olivier, and Yuriy Gorodnichenko. 2012. "What Can Survey Forecasts Tell Us about Information Rigidities?" *Journal of Political Economy*, 120(1): 116 – 159.

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

Davis, Joshua Mark. 2007. "News and the Term Structure in General Equilibrium." Unpublished.

Del Negro, Marco, and Frank Schorfheide. 2013. "Chapter 2 - DSGE Model-Based Forecasting." In *Handbook of Economic Forecasting.* Vol. 2, , ed. Graham Elliott and Allan Timmermann, 57 – 140. Elsevier.

Del Negro, Marco, and Stefano Eusepi. 2011. "Fitting observed inflation expectations." *Journal of Economic Dynamics and Control*, 35(12): 2105 – 2131.

Fernández-Villaverde, Jesús, Juan F. Rubio-Ramírez, Thomas J. Sargent, and Mark W. Watson. 2007. "ABCs (and Ds) of Understanding VARs." *The American Economic Review*, 97(3): 1021–1026.

Fuhrer, Jeff. 2017. "Expectations as a Source of Macroeconomic Persistence: Evidence from Survey Expectations in a Dynamic Macro Model." *Journal of Monetary Economics*, 86: 22 – 35.

Galí, Jordi, and Mark Gertler. 1999. "Inflation dynamics: A structural econometric analysis." *Journal of Monetary Economics*, 44(2): 195 – 222.

Hirose, Yasuo, and Takushi Kurozumi. 2012. "Identifying News Shocks with Forecast Data." Centre for Applied Macroeconomic Analysis, Crawford School of Public Policy, The Australian National University CAMA Working Papers 2012-01.

Jain, Monica. 2017. "Perceived Inflation Persistence." *Journal of Business & Economic Statistics*, 1–11.

Kydland, Finn E., and Edward C. Prescott. 1982. "Time to Build and Aggregate Fluctuations." *Econometrica*, 50(6): 1345–1370.

Laubach, Thomas, and John C. Williams. 2003. "Measuring the Natural Rate of Interest." *The Review of Economics and Statistics*, 85(4): 1063–1070.

Leeper, Eric M., Todd B. Walker, and Shu-Chun Susan Yang. 2013. "Fiscal Foresight and Information Flows." *Econometrica*, 81(3): 1115–1145.

Levinsohn, James, and Amil Petrin. 2003. "Estimating Production Functions Using Inputs to Control for Unobservables." *Review of Economic Studies*, 70(2): 317–341.

Lucas, Robert E. 1972. "Expectations and the neutrality of money." *Journal of Economic Theory*, 4(2): 103 – 124.

Lucas, Robert E. 1976. "Econometric policy evaluation: A critique." *Carnegie-Rochester Conference Series on Public Policy*, 1(1): 19–46.

Maćkowiak, Bartosz, and Mirko Wiederholt. 2009. "Optimal Sticky Prices under Rational Inattention." *The American Economic Review*, 99(3): 769–803.

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

Manski, Charles F. 1993. "Identification of Endogenous Social Effects: The Reflection Problem." *The Review of Economic Studies*, 60(3): 531–542.

Manski, Charles F. 2004. "Measuring Expectations." *Econometrica*, 72(5): 1329–1376.

Milani, Fabio, and Ashish Rajrhandari. 2012. "Observed Expectations, News Shocks, and the Business Cycle." University of California-Irvine, Department of Economics Working Papers 121305.

Muth, John F. 1961. "Rational Expectations and the Theory of Price Movements." *Econometrica*, 29(3): 315–335.

Nguyen, Thuy Lan, and Wataru Miyamoto. 2014. "News shocks and Business cycles: Evidence from forecast data." Society for Economic Dynamics 2014 Meeting Papers 259.

Olley, G. Steven, and Ariel Pakes. 1996. "The Dynamics of Productivity in the Telecommunications Equipment Industry." *Econometrica*, 64(6): 1263–1297.

Roberts, John M. 1995. "New Keynesian Economics and the Phillips Curve." *Journal of Money, Credit and Banking*, 27(4): 975–984.

Ryngaert, Jane. 2017. "What Do (and Don't) Forecasters Know about U.S. Inflation?" University of Texas at Austin.

Sims, Christopher A. 1980. "Macroeconomics and Reality." *Econometrica*, 48(1): 1–48.

Sims, Christopher A. 2003. "Implications of rational inattention." *Journal of Monetary Economics*, 50(3): 665 – 690.

Sims, Christopher A., and Tao Zha. 2006. "Were there regime switches in U.S. monetary policy?" *The American Economic Review*, 96(1): 54–81.

Sims, Eric R. 2012. "News, Non-Invertibility, and Structural VARs." In *Nathan Balke, Fabio Canova, Fabio Milani, Mark A. Wynne (ed.) Advances in Econometrics, Volume 28, DSGE Models in Macroeconomics: Estimation, Evaluation, and New Developments*. 81–135. Emerald Group Publishing Limited.

Watson, Mark W. 1986. "Vector autoregressions and cointegration." In *Handbook of Econometrics*. Vol. 4 of *Handbook of Econometrics*, , ed. R. F. Engle and D. McFadden, Chapter 47, 2843–2915. Elsevier.

Woodford, Michael. 2003. "Imperfect Common Knowledge and the Effects of Monetary Policy." In *Knowledge, Information, and Expectations in Modern Macroeconomics : In Honor of Edmund S. Phelps.* , ed. Philippe Aghion, Roman Frydman, Joseph Stiglitz and Michael Woodford, 25–58. Princeton, N.J.: Princeton University Press.

Appendix

A Proofs

A.1 Proof of Lemma 1

Proof. Guess $E[\mathbf{y}_{t+h+1}|\mathcal{F}_t] = \mathbf{A}E[\mathbf{y}_{t+h}|\mathcal{F}_t]$, and insert the guess in system (2), then by RE:

$$\begin{aligned} E[\mathbf{y}_{t+h}|\mathcal{F}_t] =& \boldsymbol{\Gamma} E[\mathbf{y}_{t+h-1}|\mathcal{F}_t] + \boldsymbol{\Phi}\mathbf{A}E[\mathbf{y}_{t+h}|\mathcal{F}_t] \\ =& (I - \boldsymbol{\Phi}\mathbf{A})^{-1}\boldsymbol{\Gamma} E[\mathbf{y}_{t+h-1}|\mathcal{F}_t]. \end{aligned}$$

Hence, the guess can be verified if $\mathbf{A} = (I - \boldsymbol{\Phi}\mathbf{A})^{-1}\boldsymbol{\Gamma}$. ∎

A.2 Proof of Proposition 1

Proof. (a) Under RE, and FI.i, lemma 1 shows that system (1) can be expressed in terms of expectations about today:

$$\mathbf{y}_t = \boldsymbol{\Gamma}\mathbf{y}_{t-1} + \boldsymbol{\Phi}\mathbf{A}E[\mathbf{y}_t|\mathcal{F}_t] + \boldsymbol{\Pi}\mathbf{u}_t. \tag{32}$$

Under FI, $E[\mathbf{y}_t|\mathcal{F}_t] = \mathbf{y}_t$, so that $\boldsymbol{\Gamma}$, $\boldsymbol{\Phi}\mathbf{A}$, and $\boldsymbol{\Pi}$ cannot be separately identified. This is the reflection problem of Manski (1993), who denotes coefficient $\beta \equiv \boldsymbol{\Phi}\mathbf{A}$ as *endogenous effect* and coefficient $\eta \equiv \boldsymbol{\Gamma}$ as *direct effect.* (b) Take expectations of system (32) and solve:

$$E[\mathbf{y}_t|\mathcal{F}_t] = (I - \boldsymbol{\Phi}\mathbf{A})^{-1}\boldsymbol{\Gamma}E[\mathbf{y}_{t-1}|\mathcal{F}_t] + (I - \boldsymbol{\Phi}\mathbf{A})^{-1}\boldsymbol{\Pi}E[\mathbf{u}_t|\mathcal{F}_t], \tag{33}$$

then shift information sets to get for any $k \geq 0$:

$$\begin{aligned} E[\mathbf{y}_{t+k}|\mathcal{F}_t] =& E[\mathbf{y}_{t+k}|\mathcal{F}_{t-1}] + (E[\mathbf{y}_{t+k}|\mathcal{F}_t] - E[\mathbf{y}_{t+k}|\mathcal{F}_{t-1}]) \\ =& (I - \boldsymbol{\Phi}\mathbf{A})^{-1}\boldsymbol{\Gamma}E[\mathbf{y}_{t+k-1}|\mathcal{F}_{t-1}] + (E[\mathbf{y}_{t+k}|\mathcal{F}_t] - E[\mathbf{y}_{t+k}|\mathcal{F}_{t-1}]). \end{aligned} \tag{34}$$

The regression of $E[\mathbf{y}_{t+k}|\mathcal{F}_t]$ on $E[\mathbf{y}_{t+k-1}|\mathcal{F}_{t-1}]$ is unbiased, as by law of iterated expectations:

$$E[E[\mathbf{y}_{t+k-1}|\mathcal{F}_{t-1}]\big(E[\mathbf{y}_{t+k}|\mathcal{F}_t] - E[\mathbf{y}_{t+k}|\mathcal{F}_{t-1}]\big)] = 0,$$

hence, composite parameter $\mathbf{A} = (I - \boldsymbol{\Phi}\mathbf{A})^{-1}\boldsymbol{\Gamma}$ is identified. ∎

A.3 Lemma 2 with Proof

Define average IRF, IRF_h, as the average difference between the variable at $t+h$ when the shock is set equal to its standard deviation, and the outcome when the shock is set equal to zero:

$$IRF_h \equiv E\left[E\left[\mathbf{y}_{t+h}|u_{it} = \sqrt{Var(u_{it})} = 1\right] - E[\mathbf{y}_{t+h}|u_{it} = 0]\right]. \tag{35}$$

Define average impulse or average shock as a shock which realization is learned with the same speed as the agent learns on average. The average shock $\bar{u}_{it}$ is learned as follows:

$$E[\bar{u}_{it}|\mathcal{F}_{t+k}] - E[\bar{u}_{it}|\mathcal{F}_{t+k-1}] = E\left[E[u_{it}|\mathcal{F}_{t+k}] - E[u_{it}|\mathcal{F}_{t+k-1}]\Big|u_{it} = 1\right] \equiv s_k,\ \forall k, t, \tag{35}$$

where $Var(E[u_{it}|\mathcal{F}_{t+k}] - E[u_{it}|\mathcal{F}_{t+k-1}]) \approx s_k$ is the linear approximation of the conditional expectation in (35), which is equal to s_k under the assumption of normality. In the model presented in section 4, the following holds:

Lemma 2 *Average response to a shock is equal to the response to the average shock.*

Proof. The second expression in (35) can be expressed as expectations conditional on an information set where no information about the shock is available, while the first expression can be expressed as conditional expectations where all of the shock is observed so that

$$\begin{aligned} IRF_h &= E\left[E\left[E[\mathbf{y}_{t+h}|\mathcal{F}_{t+K}] - E[\mathbf{y}_{t+h}|\mathcal{F}_{t-H-1}]\Big|u_{it} = 1\right]\right], \\ &= E\left[E\left[\sum_{k=-H}^{K}\Big(E[\mathbf{y}_{t+h}|\mathcal{F}_{t+k}] - E[\mathbf{y}_{t+h}|\mathcal{F}_{t+k-1}]\Big)\Bigg|u_{it} = 1\right]\right], \end{aligned}$$

where the second equation follows from the fact that the agent learns about u_{it} within the interval of H periods before and K periods after impact so that the effects of the shock are fully learned within that interval, as well. Let x and v be random variables, then $u_{it} = \sum_{m=-H}^{K}(E[u_{it}|\mathcal{F}_{t+m}] - E[u_{it}|\mathcal{F}_{t+m-1}])$, and $E[x|v=1] = E[E[x|v]|v=1]$ implies

$$IRF_h = E\left[E\left[\sum_{k=-H}^{K} E\left[E[\mathbf{y}_{t+h}|\mathcal{F}_{t+k}] - E[\mathbf{y}_{t+h}|\mathcal{F}_{t+k-1}]\Bigg|\sum_{m=-H}^{K}\Big(E[u_{it}|\mathcal{F}_{t+m}] - E[u_{it}|\mathcal{F}_{t+m-1}]\Big)\right]\right]\Bigg|u_{it} = 1\right]\right]. \tag{36}$$

Note that by law of iterated expectations, the changes in expectations about shocks and variables are uncorrelated so that for any z_1 and z_2 (e.g. $z_1 = z_2 = u_{it}$ or $z_1 = \mathbf{y}_{t+h}$ and $z_2 = u_{it}$), and for all $m > n$:

$$\begin{aligned} &E\left[\Big(E[z_1|\mathcal{F}_{t+m}] - E[z_1|\mathcal{F}_{t+m-1}]\Big)\Big(E[z_2|\mathcal{F}_{t+n}] - E[z_2|\mathcal{F}_{t+n-1}]\Big)\right] \\ =&E\left[E[z_1|\mathcal{F}_{t+n}]E[z_2|\mathcal{F}_{t+n}]\right] - E\left[E[z_1|\mathcal{F}_{t+n-1}]E[z_2|\mathcal{F}_{t+n-1}]\right] \\ &- E\left[E[z_1|\mathcal{F}_{t+n}]E[z_2|\mathcal{F}_{t+n}]\right] + E\left[E[z_1|\mathcal{F}_{t+n-1}]E[z_2|\mathcal{F}_{t+n-1}]\right] = 0. \end{aligned} \tag{37}$$

The conditional expectation inside (36) can be illustrated as $E\left[E[x|v_1+v_2]\big|v_1+v_2=1\right]$, where v_1, v_2, respectively the revisions in expectations about the shocks are uncorrelated according to (37). Given linearity, $x = av_1 + bv_2$, the previous expression is equal to

$E\left[E[x|v_1]+E[x|v_2]\middle|v_1+v_2\middle|v_1+v_2=1\right]=E\left[E[x|v_1]+E[x|v_2]\middle|v_1+v_2=1\right]$, hence,

$$IRF_h=E\left[E\left[\sum_{k=-H}^{K}\sum_{m=-H}^{K}E\left[E[\mathbf{y}_{t+h}|\mathcal{F}_{t+k}]-E[\mathbf{y}_{t+h}|\mathcal{F}_{t+k-1}]\middle|E[u_{it}|\mathcal{F}_{t+m}]-E[u_{it}|\mathcal{F}_{t+m-1}]\right]\middle|u_{it}=1\right]\right].$$

By law of iterated expectations, $E\left[E[\mathbf{y}_{t+h}|\mathcal{F}_{t+k}]-E[\mathbf{y}_{t+h}|\mathcal{F}_{t+k-1}]\middle|E[u_{it}|\mathcal{F}_{t+m}]-E[u_{it}|\mathcal{F}_{t+m-1}]\right]=0$ for all $m\neq k$ so that

$$IRF_h=E\left[E\left[\sum_{k=-H}^{K}E\left[E[\mathbf{y}_{t+h}|\mathcal{F}_{t+k}]-E[\mathbf{y}_{t+h}|\mathcal{F}_{t+k-1}]\middle|E[u_{it}|\mathcal{F}_{t+k}]-E[u_{it}|\mathcal{F}_{t+k-1}]\right]\middle|u_{it}=1\right]\right].$$

The above expression has the following form: $E\left[E\left[E[x_1|v_1]+E[x_2|v_2]\middle|v_1+v_2=1\right]\right]$, which is identical to $E\left[E\left[E[x_1|v_1]\middle|v_1+v_2=1\right]\right]+E\left[E\left[E[x_2|v_2]\middle|v_1+v_2=1\right]\right]$ so that

$$IRF_h=\sum_{k=-H}^{K}E\left[E\left[E\left[E[\mathbf{y}_{t+h}|\mathcal{F}_{t+k}]-E[\mathbf{y}_{t+h}|\mathcal{F}_{t+k-1}]\middle|E[u_{it}|\mathcal{F}_{t+k}]-E[u_{it}|\mathcal{F}_{t+k-1}]\right]\middle|u_{it}=1\right]\right].$$

An expression inside the above equation has the following form: $E[E[x_1|v_1]|v_1+v_2=1]$, which is equal to $E[x_1|E[v_1|v_1+v_2=1]]$, since $E[x_1|v_1]$ is linear in v_1, hence,

$$IRF_h=\sum_{k=-H}^{K}E\left[E[\mathbf{y}_{t+h}|\mathcal{F}_{t+k}]-E[\mathbf{y}_{t+h}|\mathcal{F}_{t+k-1}]\middle|E\left[E[u_{it}|\mathcal{F}_{t+k}]-E[u_{it}|\mathcal{F}_{t+k-1}]\middle|u_{it}=1\right]\right].$$

Note that $E\left[E[u_{it}|\mathcal{F}_{t+k}]-E[u_{it}|\mathcal{F}_{t+k-1}]\middle|u_{it}=1\right]$ is defined as s_k in (35) and it is constant across t:

$$IRF_h=\sum_{k=-H}^{K}E\left[E[\mathbf{y}_{t+h}|\mathcal{F}_{t+k}]-E[\mathbf{y}_{t+h}|\mathcal{F}_{t+k-1}]\middle|E[u_{it}|\mathcal{F}_{t+k}]-E[u_{it}|\mathcal{F}_{t+k-1}]=s_k\right]. \tag{38}$$

Taking the sum back inside expectations, and since $\sum_{k=-H}^{K}\left(E[\mathbf{y}_{t+h}|\mathcal{F}_{t+k}]-E[\mathbf{y}_{t+h}|\mathcal{F}_{t+k-1}]\right)$ is equal to $\left(\mathbf{y}_{t+h}-E[\mathbf{y}_{t+h}|\mathcal{F}_{t+H-1}]\right)$:

$$IRF_h=\sum_{k=-H}^{K}E\left[\mathbf{y}_{t+h}\middle|E[u_{it}|\mathcal{F}_{t+k}]-E[u_{it}|\mathcal{F}_{t+k-1}]=s_k\right]-E[\mathbf{y}_{t+h}].$$

Let w be a random variable that is linear in v_1 and v_2, then $E[w|v_1=r_1]+E[w|v_2=r_2]=E[w|v_1=r_1,v_2=r_2]$, respectively, in the big model:

$$\begin{aligned}IRF_h&=E\left[\mathbf{y}_{t+h}\middle|\begin{pmatrix}E[u_{it}|\mathcal{F}_{t-H}]-E[u_{it}|\mathcal{F}_{t-H-1}]\\ \vdots\\ E[u_{it}|\mathcal{F}_{t+K}]-E[u_{it}|\mathcal{F}_{t+K-1}]\end{pmatrix}=\begin{pmatrix}s_{-H}\\ \vdots\\ s_K\end{pmatrix}\right]-E\left[\mathbf{y}_{t+h}\middle|\begin{pmatrix}E[u_{it}|\mathcal{F}_{t-H}]-E[u_{it}|\mathcal{F}_{t-H-1}]\\ \vdots\\ E[u_{it}|\mathcal{F}_{t+K}]-E[u_{it}|\mathcal{F}_{t+K-1}]\end{pmatrix}=0\right],\\ &=E\left[E\left[\mathbf{y}_{t+h}|u_{it}=\bar{u}_{it}\right]-E[\mathbf{y}_{t+h}|u_{it}=0]\right],\end{aligned}$$

which is the response to the average shock, where average shock is defined in (35). ■

B Model Solutions

B.1 Micro-Foundation and Solution of Simple RE Model (3)

Consider the following law of motion for inflation π_t:

$$\pi_t = \kappa\pi_{t-1} + e_t + \epsilon_t. \tag{39}$$

The central bank can intervene and exert effort e_t to adjust inflation, while being subject to an *iid* error which is only observed with a delay, i.e. $\epsilon_t \overset{iid}{\sim} N(0, \sigma^2_\epsilon)$, $E[\epsilon_t|\mathcal{F}_t] = 0$ and $E[\epsilon_t|\mathcal{F}_{t+1}] = \epsilon_t$. The central bank chooses effort e_t to solve the following problem:

$$\max \sum_{s=0}^{\infty} \beta^s \Big[(a + \mu_{t+s})\pi_{t+s} - b\pi^2_{t+s} - ce^2_{t+s}\Big]. \tag{40}$$

Some level of inflation is beneficial to the economy, $a > 0$, as it keeps unemployment balanced, but too high inflation distorts prices and thus harms the economy, $b > 0$, and large interventions are costly, $c > 0$. Inflation is sometimes more beneficial than other times, $\mu_t \overset{iid}{\sim} N(0, \sigma^2_\mu)$, and the future is discounted by factor β.

Reformulate problem (40) in terms of expected inflation by replacing effort using (39), $e_{t+s} = E[\pi_{t+s}|\mathcal{F}_t] - \kappa E[\pi_{t+s-1}|\mathcal{F}_t]$, so that the central bank chooses $E[\pi_{t+s}|\mathcal{F}_t]$ to maximize the following expression:

$$\max \sum_{s=0}^{\infty} \beta^s \Big[(a + \mu_{t+s|t})E[\pi_{t+s}|\mathcal{F}_t] - bE[\pi_{t+s}|\mathcal{F}_t]^2 - c(E[\pi_{t+s}|\mathcal{F}_t] - \kappa E[\pi_{t+s-1}|\mathcal{F}_t])^2\Big].$$

Solve for first order condition:

$$a + \mu_t - 2bE[\pi_t|\mathcal{F}_t] - 2c(E[\pi_t|\mathcal{F}_t] - \kappa E[\pi_{t-1}|\mathcal{F}_t]) + \beta\kappa c(E[\pi_{t+1}|\mathcal{F}_t] - \kappa E[\pi_t|\mathcal{F}_t]) \overset{!}{=} 0,$$

and rearrange:

$$\begin{aligned}
[2(b+c) + \beta\kappa^2 c]E[\pi_t|\mathcal{F}_t] &= a + 2\kappa c E[\pi_{t-1}|\mathcal{F}_t] + \beta\kappa c E[\pi_{t+1}|\mathcal{F}_t] + \mu_t, \\
E[\pi_t|\mathcal{F}_t] &= \frac{a}{2(b+c) + \beta\kappa^2 c} + \frac{2\kappa c}{2(b+c) + \beta\kappa^2 c} E[\pi_{t-1}|\mathcal{F}_t] \\
&\quad + \frac{\beta\kappa c}{2(b+c) + \beta\kappa^2 c} E[\pi_{t+1}|\mathcal{F}_t] + \frac{1}{2(b+c) + \beta\kappa^2 c}\mu_t, \\
E[\pi_t|\mathcal{F}_t] &= \alpha + \gamma E[\pi_{t-1}|\mathcal{F}_t] + \phi E[\pi_{t+1}|\mathcal{F}_t] + \mu_t.
\end{aligned}$$

From (39) we know that $\pi_t - E[\pi_t|\mathcal{F}_t] = \epsilon_t$, hence,

$$\pi_t = \alpha + \gamma\pi_{t-1} + \phi E[\pi_{t+1}|\mathcal{F}_t] + \sigma u_t,$$

where $\sigma u_t \equiv \mu_t + \epsilon_t$, and demeaning the system gets rid of α.

B.2 Solution of Generalized RE Model

The rational expectations model (1) can be expressed similar to Binder and Pesaran (1997):

$$E[\mathbf{y}_t|\mathcal{F}_t] = \mathbf{A}E[\mathbf{y}_{t-1}|\mathcal{F}_t] + E[\mathbf{z}_t|\mathcal{F}_t] \tag{15}$$

$$E[\mathbf{z}_t|\mathcal{F}_t] = \mathbf{B}_1 E[\mathbf{z}_{t+1}|\mathcal{F}_t] + \mathbf{B}_0\mathbf{\Pi}E[\mathbf{u}_t|\mathcal{F}_t], \tag{16}$$

where $E[\mathbf{z}_t|\mathcal{F}_t]$ is a forward-looking place-holder. Now plug system (15) into current expectations of system (1):

$$\begin{aligned}\Big(\mathbf{A}E[\mathbf{y}_{t-1}|\mathcal{F}_t] + E[\mathbf{z}_t|\mathcal{F}_t]\Big) =& \mathbf{\Gamma}E[\mathbf{y}_{t-1}|\mathcal{F}_t] + \mathbf{\Phi}\Big(\mathbf{A}E[\mathbf{y}_t|\mathcal{F}_t] + E[\mathbf{z}_{t+1}|\mathcal{F}_t]\Big) + \mathbf{\Pi}E[\mathbf{u}_t|\mathcal{F}_t].\\ =& \mathbf{\Gamma}E[\mathbf{y}_{t-1}|\mathcal{F}_t] + \mathbf{\Phi}\Big(\mathbf{A}^2E[\mathbf{y}_{t-1}|\mathcal{F}_t] + \mathbf{A}E[\mathbf{z}_t|\mathcal{F}_t] + E[\mathbf{z}_{t+1}|\mathcal{F}_t]\Big)\\ &+ \mathbf{\Pi}E[\mathbf{u}_t|\mathcal{F}_t]\end{aligned}$$

Rearrange:

$$\begin{aligned}(I - \mathbf{\Phi}\mathbf{A})E[\mathbf{z}_t|\mathcal{F}_t] =& (\mathbf{\Gamma} - \mathbf{A} + \mathbf{\Phi}\mathbf{A}^2)E[\mathbf{y}_{t-1}|\mathcal{F}_t] + \mathbf{\Phi}E[\mathbf{z}_{t+1}|\mathcal{F}_t] + \mathbf{\Pi}E[\mathbf{u}_t|\mathcal{F}_t]\\ E[\mathbf{z}_t|\mathcal{F}_t] =& \left[(I - \mathbf{\Phi}\mathbf{A})^{-1}\mathbf{\Gamma} - \mathbf{A}\right] E[\mathbf{y}_{t-1}|\mathcal{F}_t] + (I - \mathbf{\Phi}\mathbf{A})^{-1}\mathbf{\Phi}E[\mathbf{z}_{t+1}|\mathcal{F}_t]\\ &+ (I - \mathbf{\Phi}\mathbf{A})^{-1}\mathbf{\Pi}E[\mathbf{u}_t|\mathcal{F}_t],\end{aligned}$$

so that together with system (16) one gets $[(I - \mathbf{\Phi}\mathbf{A})^{-1}\mathbf{\Gamma} - \mathbf{A}] = 0$, $\mathbf{B}_1 = (I - \mathbf{\Phi}\mathbf{A})^{-1}\mathbf{\Phi}$, and $\mathbf{B}_0 = (I - \mathbf{\Phi}\mathbf{A})^{-1}$.

B.3 Solution with Multiple Lags, Leads, and Trend

The rational expectations model (29) can be expressed similar to Binder and Pesaran (1997):

$$E[\mathbf{y}_t|\mathcal{F}_t] = \mathbf{B}_0\mathbf{q}_t + \sum_{p=1}^{P}\mathbf{A}_pE[\mathbf{y}_{t-p}|\mathcal{F}_t] + E[\mathbf{z}_t|\mathcal{F}_t], \tag{41}$$

$$E[\mathbf{z}_t|\mathcal{F}_t] = \sum_{r=1}^{R}\mathbf{B}_rE[\mathbf{z}_{t+r}|\mathcal{F}_t] + \mathbf{B}_0\mathbf{\Pi}E[\mathbf{u}_t|\mathcal{F}_t]. \tag{42}$$

Now plug equation (41) into (29) repetitively until $E[\mathbf{z}_t|\mathcal{F}_t]$ is no longer function of current and future $E[\mathbf{y}_{t+r}|\mathcal{F}_t]$:

$$\begin{aligned}E[\mathbf{z}_t|\mathcal{F}_t] =& \sum_{p=1}^{\max\{P,R\}}\left(\mathbb{1}_{p\le P}(\mathbf{\Gamma}_p - \mathbf{A}_p) + \mathbb{1}_{p\le Q}\sum_{r=1}^{R}\mathbf{\Phi}_r\mathbf{M}_p^{1+r}\right)E[\mathbf{y}_{t-p}|\mathcal{F}_t]\\ &+ \sum_{r=1}^{R}\sum_{s=0}^{r}\mathbf{\Phi}_r\mathbf{M}_1^sE[\mathbf{z}_{t+r-s}|\mathcal{F}_t] + \mathbf{\Pi}E[\mathbf{u}_t|\mathcal{F}_t],\end{aligned} \tag{43}$$

where $\mathbf{M}_p^{1+r}$ is function of $\{\mathbf{A}_1, ..., \mathbf{A}_P\}$. In particular, $\mathbf{M}_i^k$ is the i^{th} top $N \times N$ submatrix from the left of $(\bar{\mathbf{M}})^k$, where $\bar{\mathbf{M}}$ is the companion matrix:

$$\bar{\mathbf{M}} \equiv \begin{bmatrix} \mathbf{A}_1 & \mathbf{A}_2 & \cdots & \mathbf{A}_{P-1} & \mathbf{A}_P \\ 1_N & 0_N & \cdots & 0_N & 0_N \\ \vdots & & \ddots & & \\ 0_N & 0_N & \cdots & 1_N & 0_N \end{bmatrix}.$$

The coefficients in front of past $E[\mathbf{y}_{t-p}|\mathcal{F}_t]$ need to be zero, as $E[\mathbf{z}_t|\mathcal{F}_t]$ is only forward-looking so that the following conditions need to hold:

$$\begin{aligned} \mathbf{\Gamma}_p - \mathbf{A}_p + \sum_{r=1}^{R} \mathbf{\Phi}_r \mathbf{M}_p^{1+r} &= 0, \quad \forall p \leq \min\{P, R\} \\ \mathbf{\Gamma}_p - \mathbf{A}_p &= 0, \quad \forall p > R \\ \sum_{r=1}^{R} \mathbf{\Phi}_r \mathbf{M}_p^{1+r} &= 0, \quad \forall p > P. \end{aligned}$$

Under these conditions, system (43) can be written as follows:

$$\begin{aligned} E[\mathbf{z}_t|\mathcal{F}_t] =& \sum_{r=1}^{R} \sum_{s=0}^{r} \mathbf{\Phi}_r \mathbf{M}_1^s E[\mathbf{z}_{t+r-s}|\mathcal{F}_t] + \mathbf{\Pi} E[\mathbf{u}_t|\mathcal{F}_t], \\ =& \sum_{r=1}^{R} \left(1 - \sum_{s=1}^{R} \mathbf{\Phi}_s \mathbf{M}_1^s\right)^{-1} \sum_{s=l}^{R} \mathbf{\Phi}_s \mathbf{M}_1^{s-r} E[\mathbf{z}_{t+r}|\mathcal{F}_t] + \left(1 - \sum_{s=1}^{R} \mathbf{\Phi}_s \mathbf{M}_1^s\right)^{-1} \mathbf{\Pi} E[\mathbf{u}_t|\mathcal{F}_t] \\ \equiv& \sum_{r=1}^{R} \mathbf{B}_r E[\mathbf{z}_{t+r}|\mathcal{F}_t] + \mathbf{B}_0 \mathbf{\Pi} E[\mathbf{u}_t|\mathcal{F}_t], \end{aligned}$$

and therefore,

$$\mathbf{B}_r = \sum_{s=l}^{R} \mathbf{\Phi}_s \mathbf{M}_1^{s-r}, \quad \mathbf{B}_0 = \left(1 - \sum_{s=1}^{R} \mathbf{\Phi}_s \mathbf{M}_1^s\right)^{-1}.$$

B.4 Decomposition of Revisions Variance-Covariance Matrix

Define $\mathbf{R}_{ij}$ as the $\{ij\}^{th}$ $N \times N$ matrix in $\mathbf{R}$ of system 30. Inspection of 18 reveals:

$$\begin{pmatrix} \mathbf{M}_1^1 \\ \mathbf{M}_1^2 \\ \mathbf{M}_1^3 \\ \mathbf{M}_1^4 \end{pmatrix} = \begin{pmatrix} \mathbf{R}_{21} \\ \mathbf{R}_{31} \\ \mathbf{R}_{41} \\ \mathbf{R}_{51} \end{pmatrix} \mathbf{R}_{11}^{-1}, \quad \mathcal{A} \equiv \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ -\mathbf{A}_1 & 1 & 0 & 0 & 0 \\ -\mathbf{A}_2 & -\mathbf{A}_1 & 1 & 0 & 0 \\ -\mathbf{A}_3 & -\mathbf{A}_2 & -\mathbf{A}_1 & 1 & 0 \\ -\mathbf{A}_4 & -\mathbf{A}_3 & -\mathbf{A}_2 & -\mathbf{A}_1 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ \mathbf{M}_1^1 & 1 & 0 & 0 & 0 \\ \mathbf{M}_1^2 & \mathbf{M}_1^1 & 1 & 0 & 0 \\ \mathbf{M}_1^3 & \mathbf{M}_1^2 & \mathbf{M}_1^1 & 1 & 0 \\ \mathbf{M}_1^4 & \mathbf{M}_1^3 & \mathbf{M}_1^2 & \mathbf{M}_1^1 & 1 \end{pmatrix}$$

Define $\bar{\mathbf{R}} \equiv \mathcal{A}^{-1}\mathbf{R}\mathcal{A}'^{-1}$. Inspection of (21) reveals

$$\begin{pmatrix}\mathbf{B}_1' \\ \mathbf{B}_2' \\ \mathbf{B}_3'\end{pmatrix} = \begin{pmatrix}\sum_{i=3}^{5}\bar{\mathbf{R}}_{ii} & \sum_{i=3}^{4}\bar{\mathbf{R}}_{i,i+1} & \bar{\mathbf{R}}_{35} \\ \sum_{i=3}^{4}\bar{\mathbf{R}}_{i,i+1} & \sum_{i=4}^{5}\bar{\mathbf{R}}_{ii} & \bar{\mathbf{R}}_{45} \\ \bar{\mathbf{R}}_{35} & \bar{\mathbf{R}}_{45} & \bar{\mathbf{R}}_{55}\end{pmatrix}^{-1} \begin{pmatrix}\sum_{i=2}^{4}\bar{\mathbf{R}}_{i,i+1} \\ \sum_{i=2}^{3}\bar{\mathbf{R}}_{i,i+2} \\ \bar{\mathbf{R}}_{25}\end{pmatrix}.$$

Matrices $\mathcal{B}_1$, $\mathcal{B}_0$, and $\mathcal{C}$ are defined as follows:

$$\mathcal{B}_1 \equiv \begin{pmatrix}1 & 0 & 0 & 0 & 0 \\ 0 & 1 & -\mathbf{B}_1 & -\mathbf{B}_2 & -\mathbf{B}_3 \\ 0 & 0 & 1 & -\mathbf{B}_1 & -\mathbf{B}_2 \\ 0 & 0 & 0 & 1 & -\mathbf{B}_1 \\ 0 & 0 & 0 & 0 & 1\end{pmatrix}^{-1}, \ \mathcal{B}_0 \equiv \begin{pmatrix}1 & 0 & 0 & 0 & 0 \\ 0 & \mathbf{B}_0 & 0 & 0 & 0 \\ 0 & 0 & \mathbf{B}_0 & 0 & 0 \\ 0 & 0 & 0 & \mathbf{B}_0 & 0 \\ 0 & 0 & 0 & 0 & \mathbf{B}_0\end{pmatrix}, \ \mathcal{C} \equiv \begin{pmatrix}\mathbf{\Pi} & 0 & 0 & 0 & 0 \\ 0 & \mathbf{\Pi} & 0 & 0 & 0 \\ 0 & 0 & \mathbf{\Pi} & 0 & 0 \\ 0 & 0 & 0 & \mathbf{\Pi} & 0 \\ 0 & 0 & 0 & 0 & \mathbf{\Pi}\end{pmatrix}.$$

Matrix $\mathbf{B}_0$ requires mapping of $\{\mathbf{A}_p\}_{p=1}^{P}$ and $\{\mathbf{B}_r\}_{r=1}^{R}$ into $\{\mathbf{\Gamma}_p\}_{p=1}^{P}$ and $\{\mathbf{\Phi}_r\}_{r=1}^{R}$, which is derived in appendix B.3. Define $\bar{\bar{\mathbf{R}}} \equiv (\mathcal{B}_\mathbf{1}\mathcal{B}_\mathbf{0})^{-1}\bar{\mathbf{R}}(\mathcal{B}_\mathbf{0}'\mathcal{B}_\mathbf{1}')^{-1}$ so that $\mathbf{\Pi}\mathbf{\Pi}' = \sum_{i=1}^{5}\bar{\bar{\mathbf{R}}}_{ii}$, where additional identifying restrictions are imposed to get $\mathbf{\Pi}$, as discussed in section 5. Once $\mathbf{\Pi}$ is identified, the covariance matrix of shock components is identified, as well: $\mathfrak{D} = \mathcal{C}^{-1}\bar{\bar{\mathbf{R}}}\mathcal{C}'^{-1}$.

C Model Extensions

C.1 RE Model where All Beliefs Matter

Consider the following extension of RE model (1):

$$\begin{aligned}\mathbf{y}_t =& \mathbf{\Gamma} E[\mathbf{y}_{t-1}|\mathcal{F}_t] + \mathbf{\Phi} E[\mathbf{y}_{t+1}|\mathcal{F}_t] + \mathbf{\Pi} E[\mathbf{v}_t|\mathcal{F}_t] \\ &+ \tilde{\mathbf{\Gamma}}(\mathbf{y}_{t-1} - E[\mathbf{y}_{t-1}|\mathcal{F}_t]) + \tilde{\mathbf{\Pi}}(\mathbf{v}_t - E[\mathbf{v}_t|\mathcal{F}_t]).\end{aligned} \tag{44}$$

Define $\mathbf{u}_t = E[\mathbf{v}_t|\mathcal{F}_t] + \mathbf{\Pi}^{-1}\tilde{\mathbf{\Pi}}(\mathbf{v}_t - E[\mathbf{v}_t|\mathcal{F}_t])$:

$$\begin{aligned}\mathbf{y}_t =& \mathbf{\Gamma} E[\mathbf{y}_{t-1}|\mathcal{F}_t] + \mathbf{\Phi} E[\mathbf{y}_{t+1}|\mathcal{F}_t] + \mathbf{\Pi}\mathbf{u}_t \\ &+ \tilde{\mathbf{\Gamma}}(\mathbf{y}_{t-1} - E[\mathbf{y}_{t-1}|\mathcal{F}_t]).\end{aligned} \tag{45}$$

By RTR assumption, $E[\mathbf{u}_t\mathbf{u}_{t+h}'] = 0$ for all $h \neq 0$. Redefine $\mathbf{\Pi}$ so that $E[\mathbf{u}_t\mathbf{u}_t'] = I$.

Taking expectations cancels out prediction errors so that the system resembles standard first order conditions:

$$E[\mathbf{y}_t|\mathcal{F}_t] = \mathbf{\Gamma} E[\mathbf{y}_{t-1}|\mathcal{F}_t] + \mathbf{\Phi} E[\mathbf{y}_{t+1}|\mathcal{F}_t] + \mathbf{\Pi} E[\mathbf{u}_t|\mathcal{F}_t].$$

Backcast revisions on the other hand only depend on revisions in fundamentals, rather than beliefs:

$$\begin{aligned}E[\mathbf{y}_{t-1}|\mathcal{F}_t] - E[\mathbf{y}_{t-1}|\mathcal{F}_{t-1}] =& \tilde{\mathbf{\Gamma}}\left(E[\mathbf{y}_{t-2}|\mathcal{F}_t] - E[\mathbf{y}_{t-2}|\mathcal{F}_{t-1}]\right) \\ &+ \mathbf{\Pi}\left(E[\mathbf{u}_{t-1}|\mathcal{F}_t] - E[\mathbf{u}_{t-1}|\mathcal{F}_{t-1}]\right).\end{aligned} \tag{17}$$

C.2 Additional Shock Terms

Relaxing FI might change the structure of the model, depending on how the shocks are micro-founded. A larger set of models is nested in a structure that depends on expected future shocks separately. Consider the following extension of RE model (29):

$$\mathbf{y}_t = \mathbf{q}_t + \sum_{p=1}^{P} \mathbf{\Gamma}_p \mathbf{y}_{t-p} + \sum_{r=1}^{R} \mathbf{\Phi}_r E[\mathbf{y}_{t+r}|\mathcal{F}_t] + \sum_{r=1}^{R} \mathbf{\Pi}_r E[\mathbf{u}_{t+r}|\mathcal{F}_t] + \mathbf{\Pi}_0 \mathbf{u}_t. \tag{46}$$

Repeated substitution of expected future shocks lead to an infinite forward-looking model that no longer depends on future shocks separately:

$$\mathbf{y}_t = \check{\mathbf{q}}_t + \sum_{p=1}^{P} \check{\mathbf{\Gamma}}_p \mathbf{y}_{t-p} + \sum_{r=1}^{\tilde{R}} \check{\mathbf{\Phi}}_r E[\mathbf{y}_{t+r}|\mathcal{F}_t] + \check{\mathbf{\Pi}} \mathbf{u}_t, \tag{47}$$

where $\tilde{R}$ is infinity, but it can be truncated if system is stable. The structure of (47) is the same as in system (29) so that parameters $\check{\mathbf{\Gamma}}_p$ and $\check{\mathbf{\Phi}}_r$, as well as residuals $\check{\mathbf{\Pi}}\mathbf{u}_t$ with all the components are identified. The identified parameters cannot be mapped to the original model parameters without additional assumptions. The identified parameters together with impulse response restrictions on $\check{\mathbf{\Pi}}$ are enough to calculate variance decompositions and impulse response functions to the shocks of the model.