Monetary News, Surprises and the Macroeconomy.

Francesco Gallio *
†

September 2015

Abstract

In this paper we investigate the effects of anticipated and unanticipated monetary policy shocks. We bring forward empirical evidence that news are a relevant channel of the monetary transmission mechanism accounting in between 25 and 50% of the overall policy effects. Consistently with the theory, a monetary tightening generates humped-shaped responses of GDP, consumption and investment and a fall in prices. Interestingly, aggregate variables adjust even before the announced policy shift actually happens. Accordingly, we testify anticipated feedback effects on the interest rate, via the policy rule. Our results are robust to alternative proxies for market expectations, to different Cholesky ordering and to alternative identification strategies, relying on sing rather than zero restrictions.

JEL classification: C32, E52, E58.

Keywords: Monetary Policy, News, Surprises, Non-fundamentalness, SVAR, Informational Sufficiency.

---

*Universitat Autònoma de Barcelona and Barcelona Graduate School of Economics. E-mail: francesco.gallio@uab.cat
†I am thankful to the MOVE foundation for the support given through the research contract. I am deeply indebted to Luca Gambetti for the time devoted to this work. The author also acknowledges Albert Marcet, Franck Portier, Patrick Fève, Ryan Chahrou, Alain Guay, Angela Fiedler, Nezih Guner, Francesc Obiols and Edgardo Lara Córdoba all the colleagues from the UAB and TSE departments for their precious advice.
1 Introduction.

The effects of monetary policy have long been object of interest in the economic profession. At present day a large amount of studies substantiated its transmission mechanism either using general equilibrium models as experimental laboratories or structural VARs as statistical tools. Christiano et al. (1999) survey the empirical literature, showing that, notwithstanding the variety of identification schemes proposed by different authors, there is a considerable agreement about the qualitative effects of exogenous monetary policy shocks. That is, regardless of the fact that identification relies on the recursiveness assumption (e.g. Sims (1992); Christiano et al. (2005)), on sign restrictions (e.g. Uhlig (2005)) or on narrative methods (e.g. Romer and Romer (2004)), the responses of various economic aggregates is broadly consistent. In a nut shell, after an exogenous monetary tightening, short term interest rates rise, aggregate output, profit and various monetary aggregates fall, and the price level responds sluggishly. Also, there is a widespread consensus that monetary policy shocks account for a modest percentage of the volatility of aggregate output and inflation.

The bulk of the aforementioned literature was focused on exogenous shocks which take the private sector by surprise. There exists some early contributions on the role of monetary anticipation (e.g. Mishkin (1982); Cochrane (1998)) and on the fact that central banks can steer market perceptions signalling through their communicates (Amato et al., 2002). However, it was only recently that the issue captured greater attention. With the zero lower bound impeding further adjustments of the short term rates, monetary authorities were robbed their traditional policy tool and had to resort to unconventional policy instruments, igniting the debate on forward guidance.

The Federal Reserve, through the press releases of the Federal Open Market Committee (FOMC), made extensive use of forward guidance, in order to move long-term yields and stimulate aggregate expenditure. Gürkaynak et al. (2005) and more recently Campbell et al. (2012) present empirical evidence that the central bank effectively telegraphs its intentions, thus affecting market expectations (measured as in Söderström (2001) and Kuttner (2001) with federal funds futures) and consequently asset prices. However, quantitative measures of the effects of such announcements on output and other macroeconomic aggregates are scarce. This is not too surprising, given that the challenge of identifying monetary policy shock is harshened in the case of shocks that are anticipated.

The idea that foresight is a source of sizeable economic fluctuations have been widely explored in fiscal policy (e.g. Yang (2005); Leeper et al. (2013)) and in real business cycle (e.g. Beaudry and Portier (2006); Schmitt-grohe and Uribe (2012); Blanchard et al. (2013); Barsky and Sims (2009); Forni et al. (2014)). Conversely, few are the theoretical
contributions in the field of monetary economic (Lorenzoni (2007); Milani and Rajbhandari (2012)) and even fewer the empirical estimates. Our work is aimed at filling this gap in the literature by disentangling the effects of anticipated (news) and unanticipated (surprise) monetary shocks.

We start presenting a small scale new-Keynesian model, introducing the extra feature of monetary foresight. The mechanism of transmission of the policy shocks is standard: due to price rigidities, the central bank has some leverage on real variables and can stimulate the economy by affecting market expectations of the nominal interest rate. The novelty resides in that agents form their beliefs both on the basis of anticipated and unanticipated information, therefore, the central bank has an extra channel to signal its intentions to the private sector.

Intuitively, when future decisions matter, the information contained in present and past data is not sufficient to capture anticipated policy reactions. This translates in the fact that the model is non-invertible and standard structural VARs are unfit to correctly recover the underlying economic shocks. To dodge this problem we propose two alternative approaches. The first, based on dynamic identification, echoes Forni et al. (2013b) and is a useful theoretical exercise to spell out the bias implied by classical VAR techniques. The second resides in closing the informational wedge in between agents and econometrician, by augmenting the VAR with market expectations. Model simulations allow to test the performance of the proposed strategies and, in line with Féve and Jidoud (2012), we find that non-invertibility is especially severe when the roots of the structural MA are close to zero.

Interestingly, while non-fundamentalness impedes the exact identification of both shocks together, it has little to say on the possibility of recovering one shock alone. Indeed, we find that even in the non-invertible representation the effects of unanticipated policy are correctly pinned down in the VAR. That is, disregarding news does not imply any bias in the estimates of the classical surprise shock, which allows for direct comparison with previous literature. However, it does imply neglecting a possibly relevant part of the transmission mechanism with consequent underestimation of the overall policy effects.

The main contribution of this work resides exactly in presenting empirical evidence that this channel is also relevant. Our estimates suggest that anticipation accounts for in between 25 and 50% of the effects of monetary policy on output. Not only news have sizeable effects on the real economy but also they significantly affect prices and market expectations. More in detail, a monetary policy tightening, both anticipated and unanticipated, provokes a contraction in aggregate production and a fall in prices, which is consistent with the theory. Noticeably, agents have foresight on policy actions, thus aggregate variables adjust even before the actual realization of future expected shocks. Lastly, as regards the interest rates,
we found impulse responses unprecedented in the literature that testify the indirect feedback of announcements through the Taylor rule. That is, following a news shock, rates fall in the short run, as a reflection of the economic contraction, and raise at longer horizons, when the announced policy shift happens.

In the remainder of the paper we proceed as follows. Section 2 presents the model; Section 3 discusses its econometric implications; Section 4 proposes two strategies to solve the issue of non-invertibility; Section 5 presents the general econometric model; Section 6 performs simulation exercises; Section 7 presents the empirical evidence; Section 8 concludes.

2 The Model.

In this section we present a simple model that explores the relevant monetary mechanisms while keeping analytical tractability. We want to capture the idea that policy shocks have an anticipated or news component - which embeds the notion of forward guidance - and an unanticipated or surprise component - reflecting the “classical” exogenous disturbance traditionally analysed in the literature.

2.1 A simple model of News and Surprises.

We consider an economy with a plain supply side, where we have no capital and output is completely demand driven \( y_t = c_t \). Given the actual level of productivity, labor inputs adjust to match the quantity \( y_t \). As customary in the literature, we summarize the behavior of the central bank with a simple Taylor rule, assuming that monetary authorities fix the short term nominal interest rate according to:

\[
i_t = i^* + \phi_\pi \pi_t + \phi_y y_t + v_t
\]  

where \( i^* = -\log \beta \), the discount factor. \( \phi_\pi \) and \( \phi_y \) are positive parameters reflecting the strength of the central bank reaction to inflation and output gap respectively, while \( v_t \) captures exogenous deviations from the systematic policy rule. To introduce news in the model we use a variant of Campbell et al. (2012) and Milani and Treadwell (2012), assuming that the aforementioned shock is the sum of two components:

\[
v_t = \rho(L) \epsilon_t + \rho(L) \eta_{t-q}.
\]  

where the i.i.d. disturbance \( \epsilon_t \) represents an unanticipated change in policy while \( \eta_{t-q} \) corresponds to its anticipated part. In plain words, in every period private agents face a surprise shift in the target rate of \( \epsilon_t \) and receive news about future monetary deviations, which will only materialize \( q \) periods into the future. As customary, \( \rho(L) \equiv \sum_{j=0}^{\infty} \rho^j L^j \) with
\( \rho < 1, \) stands for the persistence of the policy shock. For the sake of parsimony we set \( \rho = 0, \) which simplifies the above equation to:

\[
v_t = \varepsilon_t + \eta_{t-q}.
\] (3)

The mechanism of transmission of such disturbances is standard: due to nominal rigidities, a change in \( i_t \) provokes shifts in current and expected future real rates, at least over some horizons. This gives the central bank the leverage to manipulate aggregate spending and, as a result, aggregate output. More in specific, present deviations of output from its natural level match the discounted sum of expected monetary disturbances, according to:

\[
\tilde{y}_t = -\Omega \sum_{k=0}^{\infty} \Omega^k E_t\{v_{t+k}\}
\] (4)

where \( \Omega = \frac{1}{1+\phi_y} \) is a positive parameter smaller than one. Following Blanchard et al. (2013), we show in the appendix that this model can be derived as the limit case of a standard New Keynesian model with Calvo pricing, when \( \theta - \) the probability of non re-optimizing prices - goes to one (or equivalently the frequency of price adjustment goes to zero).

It could be argued that the above assumptions are over restrictive. However, some lengthy algebra (also in the appendix) shows that it is possible to extend the model, allowing for different calibrations of \( \theta \) and for positive persistence of monetary policy shocks. The relevant mechanisms are widely unchanged across specifications and thus we privilege the basic setup, which is a cleaner explanatory tool.

Furthermore, in line with Gali (2008), we posit that agents have perfect knowledge of current and past realizations of the shocks, that is, their information set is \( I_t = (\varepsilon_{t-j}, \eta_{t-j})_{j=0}^{\infty}. \) This assumption slightly differs from Christiano et al. (2005), where agents observe monetary policy up to \( t-1. \) Either specification could be employed to derive the results presented in this paper and, although we favour the former for the theoretical section, we will use both of them in the empirical application.

In our setting the expectation of the policy residual is:

\[
E_t\{v_{t+j}\} = \begin{cases} 
\varepsilon_t + \eta_{t-q} & \text{for } j = 0 \\
\eta_{t+q-j} & \text{for } 0 < j \leq q \\
0 & \text{for } j > q.
\end{cases}
\] (5)

which, plugged into (4) simplifies the infinite summation to a finite number of addends:

\[
\tilde{y}_t = -\Omega (\varepsilon_t + \eta_{t-q} + \Omega \eta_{t-q+1} + \ldots + \Omega^q \eta_q). \]

1Such assumption introduces a one lag delay in the response of the real economy to monetary disturbances and justifies the recursive identification scheme used in their VAR.
Notice that by construction $\eta_t$ does not affect $v_t$ on impact, and in fact it does not move the policy indicator till $t + q$ periods in the future. However, being that output adjusts to expected future policy shocks, news on forthcoming realizations of $v_t$ move $\tilde{y}_t$ before the shock is actually realized.

Given $q$ periods of anticipation, equation (3) and (6) join in the MA representation:

$$
\begin{pmatrix}
v_t \\
\tilde{y}_t
\end{pmatrix} = \begin{pmatrix}
1 & L^2 \\
-\Omega & -\Omega \omega_q(L)
\end{pmatrix}
M(L) \begin{pmatrix}
\varepsilon_t \\
\eta_t
\end{pmatrix} \quad (7)
$$

where

$$\omega_q(L) \equiv \sum_{k=0}^q \Omega^k L^{q-k} \quad (8)$$

is a polynomial in the lag operator gathering all information received up to $q$ periods into the past$^2$.

Two properties $\omega_q(L)$ are worth mentioning. First, the $q$ complex roots of $\omega_q(L)$ lie inside the open unit disc, more specifically at $\Omega$ distance from the origin$^3$. Second, some easy algebra shows that:

$$\omega_q(L) = L^q + \Omega \omega_{q-1}(L).$$

This equation has a simple interpretation: at time $t$ the effect of anticipated information on output is given by the old piece of news received at $t - q$ plus a composition of more recent news ($\omega_{q-1}(L)$) discounted by $\Omega$. This formulation features the same characteristic of reverse discounting found by Leeper et al. (2013) regarding fiscal foresight: latest news are discounted more heavily than past one, that is, the effect of $\eta_{t-k}$ on output is smaller than the one of $\eta_{t-s}$ for all $k < s$. This stems naturally from the model because recent information affects interest rates further away in the future while past announcements relate to more imminent policy changes.

Hereby, we present the case of $q = 2$, which will be used as a working example throughout the discussion:

$$
\begin{pmatrix}
v_t \\
\tilde{y}_t
\end{pmatrix} = \begin{pmatrix}
1 & L^2 \\
-\Omega & -\Omega(L^2 + \Omega L + \Omega^2)
\end{pmatrix}
\begin{pmatrix}
\varepsilon_t \\
\eta_t
\end{pmatrix}. \quad (9)
$$

$^2$Examples for different anticipation horizons are:

- $\omega_0(L) = 1$.
- $\omega_1(L) = L + \Omega$.
- $\omega_2(L) = L^2 + \Omega L + \Omega^2$.

$^3$As an example $\omega_2(L)$ has two roots: $r_1 = -\frac{\Omega(1 - \sqrt{3}i)}{2}$ and $r_2 = -\frac{\Omega(1 + \sqrt{3}i)}{2}$ with modulus $|r_1| = |r_2| = \Omega$. 

6
In this circumstance the effects of \( \eta_t \) on the policy residual materialize in \( t+2 \). Output gap

\[
\tilde{y}_t = -\Omega \left( \varepsilon_t + \eta_{t-2} + \Omega \eta_{t-1} + \Omega^2 \eta_t \right)
\]

depends on three increasingly discounted elements: the current policy shift \( v_t = \varepsilon_t + \eta_{t-2} \), the one expected for next period \( \Omega E_t \{ v_{t+1} \} = \Omega \eta_{t-1} \) and the one expected for two periods ahead \( \Omega^2 E_t \{ v_{t+2} \} = \Omega^2 \eta_t \). Notice that the exponent on \( \Omega \) goes in parallel with the expectation horizon, while recent information is more heavily discounted. In short, policy shocks \( v_t \) are discounted the usual way, while policy news are discounted in reverse order.

3 Failure of Classical SVARs.

3.1 Non-Fundamentalness.

Standard SVAR techniques assume that the structural shocks are a linear combination of the reduced form residuals. However, if the system is not fundamental, this assumption fails and the method is no longer valid. Introducing anticipation in the model has severe econometric implications in this sense. Consider the determinant of \( M(z) \) in (7):

\[
\begin{align*}
Det(M(z)) &= -\Omega(\omega_q(z) - z^q) \\
&= -\Omega(z^q + \Omega \omega_{q-1}(z) - z^q) \\
&= -\Omega^2 \omega_{q-1}(z)
\end{align*}
\]

where the second equation derives from the aforementioned properties of \( \omega_q(z) \). Clearly, the determinant shares the same roots with \( \omega_{q-1}(z) \) which, as discussed above, are all inside the unit circle. That is, the system is non-fundamental and non-invertible\(^4\). The only notable exception is \( q = 1 \), for which the determinant is identically equal to 1 and the system admits a VAR representation.

In other words, for anticipation horizons \( q \geq 2 \), the VAR based on \( \tilde{y}_t \) and \( v_t \) does not contain enough information to correctly identify the two structural shocks. This does not stem from the difficulty of proxying for \( v_t \) and \( \tilde{y}_t \) - usually not directly observable in the data. It is rather the consequence of agents' decision making which incorporates in every period a full set of news. Agents keep track of \( \eta_t, \ldots, \eta_{t-q+1} \) but since at time \( t \) their effect on \( v_t \) has not materialized yet, knowledge of the policy indicator cannot be fully revealing of the underlying shocks. As a matter of fact, \( v_t \) only contains \( \eta_{t-q} \) and not \( \eta_t \).

\(^4\)Recall that non-fundamentalness stems whenever the determinant of the MA representation vanishes for values of \( z \) within the open unit disc. Non-invertibility arises in that case too, but it also appears for roots equal to unity. Therefore, non-fundamentalness implies non-invertibility but not all the way round.
Also, the combination of \( v_t \) and \( \tilde{y}_t \) is informationally deficient, even if the latter depends on present and past news. The example with \( q = 2 \) can clarify this point. From (9) we have:

\[
(L + \Omega)\eta_t = \frac{\tilde{y}_t + \Omega v_t}{-\Omega^2}
\]

which contains the non invertible polynomial \((L + \Omega)\). Clearly, it is not possible to derive an expression for \( \eta_t \) in terms of past realization of the data. However, multiplying both sides times \( F = L^{-1} \) and rearranging we obtain:

\[
\eta_t = F(1 + \Omega F)^{-1} \frac{\tilde{y}_t + \Omega v_t}{-\Omega^2}.
\]

That is, the news shock does not reside in the past of the data, but rather in its future. This is why the information set of the econometrician \( I^e_t = (\tilde{y}_{t-j}, v_{t-j})_{j=0}^\infty \) - who observes data outcomes - is strictly smaller than the one of the agents \( I_t = (\epsilon_{t-j}, \eta_{t-j})_{j=0}^\infty \) - who have perfect knowledge of both shocks.

Looking on the bright side, it is important to remark that non-fundamentalness implies VAR inadequacy to recover the two structural shocks at the same time. Nevertheless, it has nothing to say on the possibility of identifying only the unanticipated component. Consider, for instance, the simplified model:

\[
\begin{align*}
v_t &= \epsilon_t + \eta_{t-1} \\
z_t &= \eta_{t-1}
\end{align*}
\]

where \( z_t \) is fully news driven and \( v_t \), as usual, contains news and surprises. Once more, the econometrician faces a VAR failure, being that

\[
\begin{pmatrix}
v_t \\
z_t
\end{pmatrix} =
\begin{pmatrix}
1 & L \\
0 & L
\end{pmatrix}
\begin{pmatrix}
\epsilon_t \\
\eta_t
\end{pmatrix}
\]

(11)

features a root at 0 and thus it is not invertible. As discussed above, this can be imputed to the fact that news \( \eta_t = Fz_t \) reside in the future of \( z_t \) and cannot be recovered with present and past data. However, \( \epsilon_t \) can be easily expressed as a liner combination of the observables: \( \epsilon_t = v_t - z_t \). That is, data up to time \( t \) is fully revealing of the surprise shock. Therefore, the variables in (11) are informationally deficient for \( \eta_t \) and \( \epsilon_t \) together but informationally sufficient for \( \epsilon_t \) alone. As we will see in the simulations, and later in the empirical section, this will also be the case for monetary policy. Informational deficiency constrains the econometrician, as far as news are concerned, but it does not alter the effectiveness of VARs in recovering monetary surprises from the data.
3.2 What can SVARs do? Blaschke Matrix.

Sims (2012) and Beaudry and Portier (2014) showed that non-invertibility is not to be thought as an “either/or” proposition. Even in a model with foresight and roots in the unit circle, the wedge between VAR innovations and economic shocks might be small and SVAR techniques might still be reliable. Therefore, it is instructive to spell out clearly what the econometrician obtains when trying to estimate our non-fundamental system.

In order to do so, we rely on the use of Blaschke Matrices as presented in Lippi and Reichlin (1994). An \( n \times n \) matrix \( B(z) \) is a Blaschke matrix (BM henceforth) if it has no poles of modulus smaller or equal to unity and if:

\[
B(z)B^*(z^{-1}) = I
\]

where \( B^*(z) \) denotes the complex conjugate of \( B(z) \). A special case of BM can be computed using the Blaschke product as follows: let \( r_1, r_2, \ldots, r_n \) be a sequence of complex roots smaller than one in modulus\(^5\). Their Blaschke product\(^6\) is defined by:

\[
b_n(L) = \prod_{j=1}^{n} \frac{L - r_j}{1 - \bar{r}_j L}.
\]

where \( \bar{r}_j \) is the complex conjugate of \( r_j \). \( b_n(L) \) can be used to build the following BM:

\[
B(L) = \begin{pmatrix}
I & 0 \\
0 & b_n(L)
\end{pmatrix}
\]

where \( I \) denotes the \((n-1)\) dimensional unit matrix.

To understand how to employ BM, we can work out our example for \( q = 2 \) which will be later generalized. Recall that model (9):

\[
\begin{pmatrix}
v_t \\
\tilde{y}_t
\end{pmatrix} = \begin{pmatrix}
1 & L^2 \\
-\Omega & -\Omega(L^2 + \Omega L + \Omega^2)
\end{pmatrix} \begin{pmatrix}
\varepsilon_t \\
\eta_t
\end{pmatrix}
\]

has a root exactly at \( L = -\Omega \). The Blaschke factor corresponding to this single roots is:

\[
b_1(L) = \frac{L + \Omega}{1 + \Omega L}.
\]

\(^5\)In the case of \( \omega_{q-1}(L) \) all the roots have modulus \( \Omega \), meaning \( n = q - 1 \). However, for the more general case \( M_{21}(L) \) might have a different number of roots in the unit disc, as well as some roots falling outside.

\(^6\)The Blaschke product delivers an analytical function in the open unit disc constrained to have zeros at a finite (or infinite) sequence of prescribed complex numbers \( r_1, r_2, \ldots \).
Next, we can make use of the associated BM to split the system in two blocks:

\[
\begin{pmatrix}
    v_t \\
    \tilde{y}_t
\end{pmatrix}
= 
\begin{pmatrix}
    1 & \frac{\varepsilon_t^2}{b_1(L)} \\
    -\Omega & -\Omega b_1(L)
\end{pmatrix}
\begin{pmatrix}
    \varepsilon_t \\
    x_t
\end{pmatrix},
\]

(15)

\[
\begin{pmatrix}
    \varepsilon_t \\
    x_t
\end{pmatrix}
= 
\begin{pmatrix}
    1 & 0 \\
    0 & b_1(L)
\end{pmatrix}
\begin{pmatrix}
    \varepsilon_t \\
    \eta_t
\end{pmatrix}.
\]

(16)

The former is a fundamental representation. Indeed, in (15) all the roots smaller than one are eliminated dividing over \(b_1(L)\) and the determinant is left with \(-\Omega^2(1+\Omega L)\), which vanishes well outside the unit circle. Therefore, structural VARs apply and the first subsystem is what the econometrician can actually estimate. The latter block, in (16) relates the fundamental residuals obtained in the first step to the true economic shocks.

A first point to be made - which circles back to the discussion on non-invertibility - is that the RHS of (15) contains exactly \(\varepsilon_t\). Being that the system is fundamental, the surprise shock (together with \(x_t\)) resides in the space spanned by VAR innovations and can be correctly identified from the data. In other words, data is informationally sufficient for \(\varepsilon_t\) even if the original system (9) is overall non-invertible.

Wholly antithetical conclusions apply to \(\eta_t\). If the naive econometrician were to rely solely on (15), she would draw misleading inference on the actual monetary surprise. Indeed, what the VAR delivers is \(x_t\), a linear combination of reduced form residuals. A glimpse at the second sub-system (16) shows that

\[x_t = \frac{L + \Omega}{1 + \Omega L} \eta_t,\]

i.e., a non-invertible convolution of present and past values of \(\eta_t\) rather than the structural shock itself. Consequently, \(\eta_t = b_1(F)x_t\) has its expansion in the future of the fundamental shock \(x_t\) which, as previously discussed, is the very reason of data informational deficiency.

It is clear that the first step alone delivers a biased estimate of the effects of monetary news. However, it is not equally evident how severe such bias is and whether, on an empirical ground, the performance of structural VARs is hopelessly compromised. What we can state with confidence is that as \(\Omega\) decreases:

\[
\lim_{\Omega \to 0} x_t = L \eta_t
\]

the fundamental shock matches a lagged value of the news, and \(\eta_t\) is fully revealed only with one period delay. Conversely, as \(\Omega\) approaches the unit circle the wedge between the fundamental and the structural shocks disappears:

\[
\lim_{\Omega \to 1} x_t = \eta_t.
\]

\(^7^\text{Notice that } b(L)^{-1} = b(L^{-1}) = b(F).\)
This makes clear that non-invertibility is not an “either/or” problem. Indeed, how likely it is to reasonably approximate the news shocks depends on the size of the root in the unit circle. The biggest is $\Omega$ the more reliable are the estimates from a structural VAR.

To grasp an intuition of why this is the case, let’s explicit $\eta_t = b_1(F)x_t$:

$$
\eta_t = \Omega x_t + (1 - \Omega^2)(1 + \Omega F)^{-1}Fx_t.
$$

The above expression conveniently separates $x_t$ in two parts. The coefficient associated to $x_t$ is $\Omega$, while the sum of the coefficients associated its leads is $(1 - \Omega)$. This has a simple interpretation: the news shock is a combination of $x_t$ and its future, weighted by $\Omega$. High values of $\Omega$ convey more relevance to the present, thus pushing the news towards the space spanned by the data. Conversely, low $\Omega$ give more weight to future values, exacerbating the problem on non-fundamentalness.

4 Solving the non-invertibility.

In this section we present two alternatives to circumvent the problem of non-invertibility. In line with Forni et al. (2013b,a), the first is based on Blaschke matrices and designs a dynamic identification scheme. This is an interesting exercise, being that it allows to spell out clearly the bias arising from non-fundamental VARs. Therefore, we will employ it in model simulations, in order to grasp a better insight of the results presented above.

As regards the empirical application we will privilege the second approach, which is based on complementing the VAR with private sector expectations. As we discussed, non-invertibility is essentially a problem of information misalignment. Intuitively, a possible solution is to expand the scope of the VAR with variables that capture agents’ expectations about future policy. Also in this case, simulations will allow to test the performance of the proposed strategy.

4.1 Blaschke Matrix: general model.

In this section we remove the simplifying assumptions of the toy model - $\theta = 1$ and $\rho = 0$ - and we use its extended formulation, which allows for generic parameters calibration and for higher persistence of the policy shock, as in (2). Some lengthy algebra (in the appendix) shows that the moving average representation of $v_t$ and $\tilde{y}_t$ in terms of the monetary policy

$^{8}$set $F = 1$
shocks is given by:

\[
\begin{pmatrix}
v_t \\
\tilde{y}_t
\end{pmatrix} = \begin{pmatrix}
\rho(L) & \rho(L)L^q \\
\psi_1(L) & \gamma(L)
\end{pmatrix} \begin{pmatrix}
\varepsilon_t \\
\eta_t
\end{pmatrix}
\]

where \( \gamma(L) = \phi_1^{q-1}(L) + \psi_1(L)L^q \). Both \( \phi_1^{q-1}(L) \) and \( \psi_1(L) \) are polynomials in the lag operator, whose exact definition is available in the appendix. Notice the parallel with the simpler case presented above: the effect of the anticipated shock on output is a composition of the oldest information \( \psi_1(L)L^q \) and a collection of \( q - 1 \) more recent news \( \phi_1^{q-1}(L) \). Hence, \( \phi_1^{q-1}(L) \) is the generalized version of \( \Omega_{q-1}(L) \), and similarly to this latter, it is the one generating roots in the unit circle\(^9\). Therefore, also under a general calibration the model is non-invertible. In fact, the determinant \( \det(D(z)) = \phi_1^{q-1}(z)\rho(z) \) shares the same roots with \( \phi_1^{q-1}(z) \).

This calls for the use of \( b(L) \), that can be built using all the small roots of \( \phi_1^{q-1}(L) \). The corresponding matrix \( B(L) \) allows to decompose the model in its fundamental and structural couplet:

\[
\begin{pmatrix}
v_t \\
\tilde{y}_t
\end{pmatrix} = \begin{pmatrix}
\rho(L) & \frac{\rho(L)L^q}{b(L)} \\
\psi_1(L) & \frac{\phi_1^{q-1}(L) + \psi_1(L)L^q}{b(L)}
\end{pmatrix} \begin{pmatrix}
\varepsilon_t \\
x_t
\end{pmatrix}
\]

and

\[
\begin{pmatrix}
\varepsilon_t \\
x_t
\end{pmatrix} = \begin{pmatrix}
1 & 0 \\
0 & b(L)
\end{pmatrix} \begin{pmatrix}
\varepsilon_t \\
\eta_t
\end{pmatrix}.
\]

The first block has no roots problems, being that \( b(L) \) takes care of twisting inside out all the small roots. Hence it is possible to obtain a consistent estimate \( \varepsilon_t \) and \( x_t \) from the data. Next, with \( B(L) \) we can recover the true structural shocks from the first step. This differs from classical techniques in that the shocks are not obtained as a static combination of VARs residuals. On the contrary, they are pinned down by a dynamic convolution of fundamental innovations as in (21). As a consequence, a consistent estimate of \( B(L) \) is needed to correct for the first step innovations. Thus, the identification procedure follows these steps:

1. Run a VAR on the data and impose the restriction that the \( x_t \) does not affect \( v_t \) contemporaneously. This arises naturally from the fact that \( x_t \) contains a combination of news shocks whose effect on interest is delayed.

\(^9\)Figure 3 exemplifies it with a numerical study.
2. In order to find the roots falling in the unit circle, exploit the estimates of $\hat{G}_{11}(L)$ and $\hat{G}_{12}(L)$:

(a) Compute:

$$r(L) = \frac{L^q}{b(L)} = \hat{G}_{11}(L)^{-1}\hat{G}_{12}(L).$$

Notice that $r(L)$ can be expressed as $b(F)F^{-q}$. This means that it shares the same roots of $b(F)$, and as a consequence of $b(L)$.

(b) Find the the roots $\hat{r}_j$ of $r(L)$, smaller than one in modulus.

(c) Use $\hat{r}_j$ to build $\hat{b}(L)$ and $\hat{B}(L)$ as in (12) and (13)

3. The structural representation of the system in terms of the news and surprise shock is given by:

$$\begin{pmatrix} v_t \\ \tilde{y}_t \end{pmatrix} = \hat{G}(L)\hat{B}(L)\begin{pmatrix} \eta_t \\ \varepsilon_t \end{pmatrix}.$$

4.2 Adding information to the VAR.

As discussed earlier, the non-fundamentalness of the system in (19) reflects a mismatch between agents and econometrician’s information. When agents anticipate future expected policy developments the naif econometrician can only derive a convolution of the true structural shocks. In this case BMs are a valid tool to identify the underlying shocks dynamically.

Other than BMs, a different way out is to expand the information set of the econometrician by adding agents’ expectations on future policy movements. The intuition behind this approach is the following: when market participants act in advance of policy, the econometrician is unable to disentangle which movements in current output are attributable to surprise shocks and which are due to changes in expectation. This opens the wedge between agents and econometrician’s information sets and translates into non-invertibility. However, if the econometrician knew markets beliefs about future rates she would have sufficient information to properly disentangle anticipations and surprises.

But what is sufficient information in this setting? The answer to this question is intuitive: to solve the informational misalignment the econometrician must have access to market expectations over an horizon that is at least as wide as the central bank announcement. The intuition behind it is simple. Consider the baseline example with $q = 2$ and $\rho = 0$. Notice that $E_t\{v_{t+1}\} = \eta_{t-1}$ only contains past news, while $E_t\{v_{t+2}\} = \eta_t$ contains the current news. In short, agents use $\eta_t$ to change their projections for $t + q$, thus any forecast below this threshold does not carry the information needed to unveil the structural shock.

Simple algebra can formalize this point in the more general setting. Maintaining the assumption that the policy shock is a compound of news and surprises as in (2), we can
compute projections over any arbitrary horizon \( s \) using Weiner-Kolmogorov formula (see appendix).

\[
E_t\{v_{t+s}\} = \left[ \frac{\rho(L)\varepsilon_t + \rho(L)L^q\eta_t}{L^s} \right] + \begin{cases} 
\rho^s\rho(L)\varepsilon_t + \rho(L)L^{q-s}\eta_t & \text{for } s < q \\
\rho^s\rho(L)\varepsilon_t + \rho^{q-s}\rho(L)\eta_t & \text{for } s \geq q 
\end{cases}
\]

Therefore, when we condition the VAR on expected policy shocks we need to consider the two cases \( s < q \) and \( s \geq q \) separately.

\[
\begin{pmatrix}
\varepsilon_t \\
\eta_t
\end{pmatrix}
= \begin{pmatrix}
\rho(L) & \rho(L)L^q \\
\rho^s\rho(L) & \rho(L)L^{q-s}
\end{pmatrix}
\begin{pmatrix}
\varepsilon_t \\
\eta_t
\end{pmatrix}
\]

(\text{case}1)

\[
\begin{pmatrix}
\varepsilon_t \\
\eta_t
\end{pmatrix}
= \begin{pmatrix}
\rho(L) & \rho(L)L^q \\
\rho^s\rho(L) & \rho^{q-s}\rho(L)
\end{pmatrix}
\begin{pmatrix}
\varepsilon_t \\
\eta_t
\end{pmatrix}
\]

(\text{case}2)

Notice that when computing the complex zeros of the determinant of the above representations we have:

\[
\text{Det}(\mathcal{R}^{(1)}(z)) = \rho(z)^2z^{q-s} - \rho^s\rho(z)^2z^q
\]

which vanishes for \( z = 0 \) inside the circle. That is, when the market projections available do not match the anticipation horizon \( s < q \), the information set of the econometrician is still too poor to solve the non-invertibility. Conversely:

\[
\text{Det}(\mathcal{R}^{(2)}(z)) = \rho^{q-s}\rho(z)^2 - \rho^s\rho(z)^2z^q
\]

has roots of modulus \( z = \rho^{-1} \), outside the unit circle. In this case the econometrician has access to market beliefs for horizons comparable or exceeding the announcement horizons \( s \geq q \) and thus can confidently use classical SVAR techniques to derive the underlying shocks. Namely, the identification restrictions needed is \( \mathcal{R}^{(2)}_{12}(0) = 0 \) reflecting that, by construction, news do not move the policy rate up to \( q \) periods into the future.

5 Multivariate extension.

Let us now consider the extension of the bivariate model developed above, both in the Blaschke and in the additional information setting.

5.1 Blaschke Matrix.

Let \( z_t \) be a \( n - 2 \) vector of times series, to be included in the analysis. Stacking the vector below \( \tilde{y}_t \) and \( v_t \) we obtain:
where $d(L)$, $f(L)$, $h(L)$, $j(L)$, $k(L)$ are conformable matrices and vector of rational functions of $L$. One restriction in this case is no longer sufficient to completely identify the model, thus we impose full Cholesky triangularization on impact, which is enough to grant that $D_{12}^*(0)$, that is, the news has delayed effects on the policy indicator.

The structural representation can be derived from the fundamental through BMs as follows:

\[
\begin{pmatrix}
v_t \\
y_t \\
z_t
\end{pmatrix} =
\begin{pmatrix}
\rho(L) & \rho(L) L^q & d(L) \\
\psi_1(L) & \gamma(L) & f(L) \\
h(L) & j(L) & k(L)
\end{pmatrix}
\begin{pmatrix}
\varepsilon_t \\
\eta_t \\
w_t
\end{pmatrix},
\]

(23)

where $I_{n-2}$ is an identity matrix and $\mathbf{0}$ is a $n-2$ dimensional column vector of zeros. In this case the determinant of the first sub-system cannot be derived explicitly, since it depends on a whole set of matrices and vectors in the lag polynomial. Therefore, for empirical applications it would be necessary to test rather than assuming the fundamentalness of the first step, for instance using the procedure of Forni and Gambetti (2014). Once the first sub-system is estimated, identification of the roots needed to build $b(L)$ is achieved using the entries of $\hat{F}_{11}$ and $\hat{F}_{12}$ as explained in the bivariate case. Finally, matrix multiplication of the two moving averages above returns the original representation in $D^*(L)$.

5.2 Additional information.

In case additional information (equal or exceeding the foresight horizon $q$) is introduced in the VAR, we can derive the multivariate extension of the model by stacking the $n-2$ dimensional vector $z_t$ either on the top or on the bottom of the policy instrument. The first option corresponds to the case in which $z_t$ does not react to monetary policy within the same period, and closely relates to Christiano et al. (2005). We will use it as our baseline since it allows clear comparison with previous literature. The latter case brings the model
closer the specification of Gali (2008), in which the reaction of the macroeconomy to policy shock is simultaneous. We also present this specification, which will be used as robustness check.

5.2.1 Policy indicator on the bottom.

In the former case the moving average representation reads:

\[
\begin{pmatrix}
    z_t \\
    v_t \\
    E_t \{v_{t+s}\}
\end{pmatrix}
= 
\begin{pmatrix}
    \alpha(L) & \zeta(L) & \var{\theta}(L) \\
    \nu(L) & \rho(L) & \rho(L)L^2 \\
    \xi(L) & \rho^s(\rho(L)) & \rho^{s\rightarrow\rho}\rho(L)
\end{pmatrix}
\begin{pmatrix}
    w_t \\
    \varepsilon_t \\
    \eta_t
\end{pmatrix}
\]

(24)

where \(\alpha(L), \zeta(L), \var{\theta}(L), \nu(L), \xi(L)\) are conformable matrices and vectors in the lag operator. Again, identification requires \(\zeta(0), \var{\theta}(0),\) and the upper triangular part of \(\alpha(0)\) to equal zero.

In this setting the second last column of \(\mathbf{H}(L)\) represents the effects of the classical unanticipated monetary policy shock, in which a recursive scheme is applied: all variables pre-determined for the central bank do not react instantaneously to policy shocks. This is the case both for the anticipated and the unanticipated component which can be distinguished thanks to their effect on market expectations. On the one hand \(v_t\) is not moved on impact by news since, by construction, they only affect future policy rates. On the other hand expectations are moved by news instantaneously, given that \(\eta_t\) contains fruitful information to produce policy projections.

Finally, in case the VAR included also non pre-determined variables (e.g. Christiano et al. (2005)), it is possible to split the vector \(z_t\) in two blocks: \(z_{1t}\) of length \(k - 1\) with the pre-determined variables and \(z_{2t}\) of length \(n - k - 1\) containing the other variables. In this case the ordering of the VAR would be \((z_{1t} \ v_t \ E_t \{v_{t+s}\} \ z_{2t})'\) and, after imposing Cholesky restrictions, the effect of surprise and news shock are to be found in column \(k\) and \(k + 1\) of \(\mathbf{H}(L)\) respectively.

5.2.2 Policy indicator on top.

If we want to remove the delayed response assumption, it is sufficient to revert the ordering of the variables, stacking the policy block on top or \(z_t\). This is enough to introduce a contemporaneous reaction of real variables to monetary policy shocks. In this case the
The multivariate representation of the system turns into:

\[
\begin{pmatrix}
  v_t \\
  E_t\{v_{t+1}\} \\
  z_t
\end{pmatrix} =
\begin{pmatrix}
  \rho(L) & \rho(L)L^q & \nu(L) \\
  \rho^2 \rho(L) & \rho^{2-q} \rho(L) & \xi(L) \\
  \zeta(L) & \vartheta(L) & \alpha(L)
\end{pmatrix}
\begin{pmatrix}
  \varepsilon_t \\
  \eta_t \\
  \omega_t
\end{pmatrix},
\]

(25)

where \(\alpha(L), \zeta(L), \vartheta(L), \nu(L), \xi(L)\) are conformable matrices and vectors in the lag operator. The effects of unanticipated and anticipated monetary policy shocks are now to be found in the first and second column of \(\hat{H}\) respectively. The interesting restriction is \(\hat{H}_{12}(0) = 0\), mirroring the delayed response of the policy indicator to news. The difference with the previous case is that the Cholesky scheme does not restrict \(\zeta(L)\) and \(\vartheta(L)\) to be zero on impact. Put it differently, in this case the whole set of variables in \(z_t\) can react instantaneously both to policy decisions and announcements.

As a robustness check, and to take a more agnostic stance on the timing of the policy effects, we will use both specifications in the empirical section. Notice that, both ordering are legitimate as long as the policy measure proxies for \(v_t\), which is orthogonal to the variables included in the Taylor rule. Conversely, when the policy indicator used is the interest rate, more caution is needed. In fact, letting macro aggregates vary with news within the same quarter implies a prompt adjustment of the short term rate, according to the policy rule. If news have an indirect, though immediate, effect on the policy instrument, the restriction \(\hat{H}_{12}(0) = 0\) is a misspecification. As a consequence, the comparison with previous literature is circumscribed to the case of the federal funds rate ordered last.

6 Simulations.

In this section we run some simulation exercises in order to test the proposed methodology. We draw a sequence of Gaussian white noise errors for \(\eta_t\) and \(\varepsilon_t\), and given the MA representations \(D(z)\), as in (19), we construct a simulated sample of large size (1000 observations) for \(v_t\) and \(\hat{y}_t\). Furthermore, expectations for \(v_t\) at different horizons are obtained from \(R^{(1)}(L)\) and \(R^{(2)}(L)\). The model is calibrated to match quarterly data and, in the baseline, parameters are set as in Gali (2008)\(^{10}\). Simulations are repeated for 1000 times to obtain 95% confidence bands around the median estimates of the impulse responses.

Figure 4 displays results of the simulation in the bivariate case. The top panel compares the performance of classical structural VAR technique (blue line) against Blashcke-corrected impulse responses (green line). True impulse responses, that is, the d.g.p., are reported in

\[^{10}\theta = 2/3; \sigma = 1; \psi = 1; \beta = 0.99; \phi_t = 1.5; \phi_y = 0.5/4; \rho = 0.5.\]
red. As it is clear from the picture, the naive econometrician, who relies only on timing restrictions, would get a poor match of the effects of $\eta_t$. On the contrary, the estimates of $\varepsilon_t$ are relatively accurate\textsuperscript{11}. This supports our claim that data are informationally insufficient for the whole system, but still sufficiently informative for one shock alone. Furthermore, correcting the responses with the BM delivers more precise estimates of $\eta_t$, closing the wedge between the non-fundamental model and its empirical counterpart.

The bias of classical VAR is especially severe when the zeros of the theoretical representation are small. As discussed above, when the roots approach the unit circle non-fundamentalness becomes less stringent and the performance of structural VAR more reliable. The bottom panel of Figure 4 provides with an example: higher persistence of the policy shocks implies minimal root closer to unity. As a result, structural VAR techniques and Blaschke-corrected impulse responses behave almost equivalently. As noted by Sims (2012), this testifies that non-fundamentalness can be empirically problematic at quite different extents.

Results for the bivariate case complemented with market expectations are reported in Figure 5. The top panel shows $R^{(1)}(L)$, the case in which the econometrician has access to expectations falling short of monetary announcements ($s < q$). In this circumstance non-fundamentalness is stringent and the VAR fails to pin down the timing of the anticipated shock. In our specific example, the econometrician would infer that news act with a three periods delay - which corresponds to $s$ - rather than five - which is the actual announcement horizon $q$.

Notice that, conversely to the Blaschke case, we cannot reduce the empirical bias playing around with parameters calibration. This happens because the smallest root of the MA is at zero anyway, an extreme case for which non-fundamentalness is unsurmountable. Nevertheless, the responses associated to $\varepsilon_t$ are spot on, confirming once more that non-invertible systems might still contain enough information to recover one of the two structural shocks.

The bottom panel of figure 5 shows the case corresponding to $R^{(2)}(L)$. As expected, with $s \geq q$ the informational asymmetry is solved and VAR estimates match closely the true d.g.p. Conversely to the previous case, there is no confusion between $s$ and $q$. The structural VAR correctly estimates the delayed response of the news shock, thus both $\eta_t$ and $\varepsilon_t$ are recovered with precision. The excellent performance in model simulations coupled with the ease of use in different specifications justifies the choice of the latter approach for the empirical analysis that follows.

\textsuperscript{11}Notice that no Blaschke correction applies to the surprise shock, therefore the blue and green line perfectly overlap in the picture and only one of them is visible.
7 Empirical evidence.

In this section we apply the additional information method to recover and compare the effects of both anticipated and unanticipated policy disturbances. The main result is that monetary news explain a sizeable portion of the forecast error variance of GDP, consumption and inflation. Qualitatively the effects of news and surprise shocks are similar and quantitatively their relative relevance is comparable. This allows us to conclude that news have a salient role in the overall transmission of monetary policy.

7.1 Data.

The first step of our empirical analysis is to choose the series to proxy for \( v_t \) and \( E_t \{ v_{t+h} \} \). Recall that the former is the residual from the Taylor rule, that captures both current surprises and past news, while the latter reflects market expectations on future policy developments. For both we propose alternative proxies to perform a series of robustness checks.

To start with market expectations, one option is to use federal funds future contracts on the spirit of Kuttner (2001), Gürkaynak (2005) and Campbell et al. (2012). Federal funds futures (FFF henceforth) are traded by the Chicago Central Board of Trade (CBOT) since late 1988 and are a bet on future interest rates (see Söderström (2001) and Kuttner (2001) for further details). Briefly, the value at expiration of a contract for \( m + h \), i.e. at horizon \( h \) from the current month \( m \), is \( 100 - \bar{r} \), where \( \bar{r} \) is the average effective federal funds rate over the expiry month. The settlement price at time \( m \) is \( 100 - r_e \) and the seller (buyer) commits to compensate the other party in case the implied rate \( r_e \) turns out to exceed (fall behind) \( \bar{r} \) at \( m + h \). Being that such value is fully disclosed only at the end of the delivery month, \( r_e \) reflects market expectations about \( \bar{r} \).

To obtain a times series of private sector beliefs the literature suggests to use high frequency identification, looking at the meeting dates of the Federal Open Market Committee (FOMC). More in specific, the strategy is to compute the difference between FFF prices right after the meeting and their quotation the previous day. The underlying assumption is intuitive: the actions and communicates of the FOMC are promptly internalized by the private sector, whose beliefs are directly reflected on FFF prices. Being that in tight windows around the meeting the macroeconomic conditions are fairly stable, any shift in prices can be attributed to the outcomes of the meeting itself. That is, such movements capture market beliefs directly related to policy actions and announcements.

We withdraw the data from Barakchian and Crowe (2010), who report the intra-day difference in future rates for each contract from 1 to 6 months ahead, over the time span 1989 I – 2008 II. This provides with six different measures of market expectations related to
increasingly longer forecast horizons. Furthermore, through factor analysis, these authors condensate all the individual series into a single summary measure. We will employ this factor measure as our baseline, labelling it $\Delta fff$, and include the individual series alone as a robustness check.

Other than $\Delta fff$, we will also use alternative indicators of market expectations. One is the 6-months ahead FFF contract in levels, without the intra-day differencing. The others are SPF forecasts of the 3-Month Treasury Bill rate. Once more, we employ each forecast horizon as a different measure and as an extra robustness check.

Also as regards $v_t$ we exploit the data presented in Barakchian and Crowe (2010). Following the same logic as before, we take the intra-day difference of the federal funds rate around FOMC meetings and denote it $\Delta ffr$. In such a tight window the macro variables considered for systematic policy do not move significantly, thus any change in the funds rates reflects exogenous policy shifts\textsuperscript{12}. Notice that we do not discard the cases in which the central bank did not change the target rate being that zero intra-day difference might either mirror no news and surprise or a surprise movement exactly offsetting the effect of past news.

Both the series of $v_t$ and its expectations are passed to quarterly frequency with simple monthly average. The remaining quarterly data, i.e. GDP, consumption, GDP deflator, real investment, wages, productivity, federal fund rates, profits and M2, is obtained from FRED for the period 1989 I – 2014 IV.

### 7.2 VAR analysis.

In this section we present the impulse response functions and variance decomposition of three different VAR specifications. First, we study the bivariate case with the measures of $v_t$ and its expectation, which builds a bridge in between the theoretical simulations and the empirical application. Second, we run the baseline VAR with four variables, namely $\Delta fff$, $\Delta ffr$, GDP and GDP deflator. Lastly, we report the result of a larger scale VAR with 10 variables, using the specification of Christiano et al. (2005). This allows to compare our results with previous literature.

\textsuperscript{12}The idea is that the day before of the meeting the interest rate is given by $i_{d-1}^t = f(\Omega_{d-1}^t)$, where $\Omega_{d-1}^t$ is the information set of the central bank about the state of the macro economy the day prior to the meeting. At meeting dates the bank adjusts the rates according to $i_d^t = f(\Omega_d^t) + v_t$, where $v_t$ contains both the anticipated and unanticipated component. Under the assumption that the information set is fairly stable at high frequencies, $\Delta ffr = i_d^t - i_{d-1}^t = v_t$, is a proxy for the monetary policy stance.
7.2.1 Toy VAR.

The bivariate case is the simpler parallel of the simulation exercises, where $\Delta ffr$ and $\Delta fff$ stand for the policy residual and its expectation. We include four lags in the analysis and we identify the system with a zero Cholesky restriction. Figure 6 displays the corresponding impulse response functions. The baseline case is reported in black and refers to the factor measure of market expectations. In green we have the responses obtained replacing $\Delta fff$ with the individual series of each future contract separately.

$\Delta ffr$ represents the exogenous policy shift, and contains both anticipated and unanticipated shocks. Notice how, regardless of the measure of expectations used, the policy indicator shows a delayed response to news, whose higher effect materializes after four quarters. Conversely, its response to a surprise shock peaks on impact and goes to zero approximately after a year.

Recall that in our setting the intra-day difference in FFFs mirrors updates in expectations and it can be affected both by the current surprise and the last news received. Indeed, the bottom right panel of Figure 6 follows closely $\eta_t$, with unitary and significant response only on impact. This is in line with Barakchian and Crowe (2010) with the caveat that they consider such measure as being pure news, neglecting the possible effects of monetary surprises contained in it. It is true that the contribution of the surprise shock to $\Delta fff$ is limited when compared to news. However, expectations react significantly also to $\varepsilon_t$, at least within the first year. Our approach has therefore the advantage of providing with a refinement of the news shock series, cleaning it out from the effects of unanticipated monetary disturbances.

Lastly, it is interesting to underline the similarity between the simulations in Figure 5 and the 2-VAR presented here. This not only builds a bridge in between the theory and the empirical application but also reassures us in the proxy choice. Then, starting from this building block, we can expand the analysis to a VAR with a larger set of variables.

7.2.2 Baseline VAR.

The variables included in the baseline specification are GDP, GDP deflator, $\Delta ffr$ and $\Delta fff$. We include four lags in the estimation and rely on the assumption that the macroeconomy does not move with monetary policy within the same quarter. As a robustness check we will later remove this assumption. Also, we keep the theoretical restriction that news have a delayed effect on the policy indicator. This implies that the two shocks can be recovered as the last two Cholesky shocks with GDP and inflation ordered first. The relevant impulse response functions are displayed in Figure 7. As in the bivariate case, we report in black the
factor measure of $\Delta fff$ and in green all the measures from each individual future contract. Furthermore, the red line is the impulse response obtained removing $\Delta fff$ and identifying the unanticipated shock alone as the last Cholesky shock of the tri-variate system (with the policy instrument ordered last).

Consistent with the theory, a monetary policy contraction implies negative responses of output and prices. Noticeably, GDP reacts in a humped-shaped fashion to both shocks, with higher effect at around 4-5 quarters. On the contrary, movements in inflation are more sluggish and prices fall on longer horizons. As regard interest rates and expectations, the picture is similar to the bivariate case presented above, with the policy indicator responding with a delay to monetary news.

A first point to be made is that the effects of the surprise shock are consistent across all the specifications of $\Delta fff$. That is, notwithstanding the specific choice of the future contract, the responses of GDP and inflation to $\epsilon_t$ are widely unchanged. On the contrary, the effects of news are qualitatively similar but display more heterogeneity. This supports two of our main theoretical findings. The first one is that expectations on a wider horizon are a more reliable tool to solve the non-invertibility. In fact, FFF with longer maturity are richer in news and have increasingly higher effects on GDP. Analogous conclusions apply to inflation: notice how news produce an initial price puzzle which is absent for the surprise shock. However, as long as the foresight horizon of the FFF increases, the price puzzle disappears and the effects of both $\epsilon_t$ and $\eta_t$ became comparable both in shape and magnitude.

The second finding relates to data sufficiency. In the theory we argued that a subset of the structural shocks might be correctly recovered even if the whole system is flawed by non-fundamentalness. Indeed, under all specifications (with and without market expectations) the VAR performs equivalently in identifying $\epsilon_t$. This means that the extra information contained in $\Delta fff$, which is vital to pin down the news shock, is redundant to recover the surprise shock alone.

This latter result implies that previous empirical literature, focused only on unanticipated monetary policy shock, should not face large bias in recovering $\epsilon_t$, as long as the identification scheme is correct. However, Figure 7 makes it clear that both surprises and news have sizeable repercussions on GDP and prices. Therefore, not accounting for news leads to neglecting a significant portion of the overall monetary transmission mechanism. In order to quantify the relative relevance of both shocks Figure 8 reports their forecast error variance decomposition.

As noted by previous literature, monetary policy has a secondary role in explaining fluctuations of aggregate macroeconomic variables. However, if we consider news and surprises together, they account for 15% of the total output variance over a 25-periods horizon, equally
split among the two disturbances. That is, news account for half of the overall transmission mechanism on the real economy, which moderates the claim of limited scope of monetary policy. Similar conclusions apply to inflation. Both shocks have comparable effects, even if on a 25 periods span they account only for 10% of the total variance of prices.

A different picture emerges as regards financial markets. A high portion of the interest rate variance is explained by surprises especially at short horizons. This can be read as the cautious behaviour of monetary authorities, who prefer to adjust the short term rate according to specific contingencies rather than announcing an explicit path for the policy rate. This latter option, described by Campbell et al. (2012) as odyssey forward guidance, could be a powerful tool for steering market expectation. However, it has the drawback of limiting the future range of actions of central banks, exacerbating the trade-off between present commitment and subsequent credibility. Notwithstanding the fact that forward guidance is not completely explicit, monetary news explain a 10% of the interest rates movements in the medium run, which is still a sizeable portion. That is, announcements are a significant instrument in the monetary policy tool kit.

As regards expectations, the picture is completely reversed. Not surprisingly, the bulk of their fluctuations is attributable to news, which account for 80% of the total variance in the short run. Also at longer horizons this percentage does not fall below 60%, revealing that the private sector is much attentive to announcements that might contain information on future policy developments. This seems to corroborate the view of Gürkaynak et al. (2005) that the central bank communicates effectively his intentions, thus affecting private sector expectations.

### 7.2.3 Extended VAR.

In this section we extend our methodology to the specification of Christiano et al. (2005), which represents a popular benchmark in the literature. To do so, we use the same set of variables, namely: GDP, consumption, GDP deflator, real investment, wages, productivity, federal funds rate, profits and M2. On top of that we add a measure of expectations, which is fundamental to capture news effects. Notice that in this case the policy indicator is the interest rate itself instead of the Taylor residual. Therefore, to proxy for market expectations we will use the 6-months ahead federal funds future in levels (FF6), that corresponds more closely to the policy instrument at hand.

Echoing Christiano et al. (2005), identification relies on the assumption that the variables in the central bank information set are not affected by monetary policy within the same quarter. Furthermore, our theory suggests that news do not move the policy rate contemporaneously. Therefore the structural shocks can be retrieved with Cholesky restrictions
ordering the federal funds rate and the FF6 after the predetermined variables (namely in
the seventh and eighth position).

Figure 9 shows the impulse response functions of both shocks. Results are similar to
the previous case: after a contractionary policy shock GDP, consumption, productivity and
investment fall in a humped shaped manner. Inflation falls too, showing some initial price
puzzle only for the surprise component. What stands out in this specification is the behaviour
of the interest rate. We can appreciate how in response to a news shock the interest rate
falls, at least initially, and later reverts its trend, switching sign in the medium run. The
economic intuition behind this fact is as follows. The effects of news on the policy rate are
delayed, and likely more than one period. However, the private sector anticipates future
contractions, which explains why real variables and prices begin to fall from the beginning.
This happens before the actual materialization of the anticipated shock and feeds back into
the Taylor rule, pushing the rates down in the short run. Finally, when the period of inaction
of monetary news is over, the positive shock materializes, generating upwards pressure on
the policy rate and provoking its sign reversal.

Other than establishing evidence of feedbacks in the policy rate, we can exploit the larger
VAR to test for informational sufficiency in the data. To this end, we remove the federal
funds future and we re-estimate the effects of the unanticipated component alone. For ease
of comparison the corresponding impulse responses are added in red in Figure 9. Once
more, results are almost identical across specifications, supporting the hypothesis that data
is informationally sufficient to properly recover the surprise shock.

Lastly, Figure 10 reports the variance decomposition of the two shocks. Noticeably,
surprises explain a larger share of the fluctuations of GDP, consumption and investment.
However, news are responsible for a sizeable portion of the real movements, accounting in
between 25 and 30% of the overall policy effects. The same holds true for inflation, whose
variance explained by \( \eta_t \) adjusts to a third of the one attributable to \( \varepsilon_t \). Finally, as in the
previous section, the bulk of the movements in the interest rate are due to surprises while
news have a more prominent role in explaining expectations, especially in the short run.

To sum up, also in the extended setting we find that news and surprises have similar
effects and that anticipation offers a significant contribution to the overall monetary trans-
mission mechanism.

7.3 Robustness analysis.

A part from the seven measures of federal funds futures, which are already a preliminary
robustness check, in this section we report three additional specifications. First, we remove
the assumption of delayed response of the real economy, by reverting the order of the vari-
ables as in (25). Then, we propose a different identification strategy, based on sign rather than zero restrictions. Finally, we replace the measure of market expectations with SPF data.

7.3.1 Policy indicator ordered first.

As discussed in the theoretical section, removing the assumption of delayed response of the macro economy is as simple as reverting the order of the variables and looking at the first two Cholesky shocks. On the one hand, Figure 11 displays the resulting impulse responses, both for the factor and the individual measures of $\Delta fff$. To facilitate the comparison, the baseline VAR (with the policy indicator ordered last) is also reported in the graph, using red lines. On the other hand, Figure 14 contains the corresponding variance decomposition.

As in the baseline case, the specific choice of the federal funds future is irrelevant to pin down the surprise shock while it provokes higher dispersion in the responses of the anticipated component. As previously discussed, this phenomenon relates to the informational content of the data. The exercise we are performing in this section consists merely in rearranging the variables and, by its own nature, it cannot alter the relevant information carried by the VAR. Thus, it comes as no surprise that we find once more informational sufficiency for the surprise component alone.

Furthermore, from the graph it is evident how impulse responses are consistent regardless of the variable ordering, especially in what concerns the news shock. The most compelling difference with the baseline case resides in the stronger response of the macroeconomy to the surprise component. Output jumps significantly on impact, while maintaining its humped-shape pattern. Also inflation is pushed down with more intensity. However, the contemporaneous movements in prices and the impact response of GDP to news are not significant, showing that the zero restrictions imposed in the baseline are not too unreasonable after all.

As a consequence of the higher responses of output and inflation to the surprise shock, their variance decomposition associates more weight to $\epsilon_t$. However, this does not subtract relevance to news which maintain the same percentages as in the baseline, and still account for a 25% of the overall policy effects. Moreover, also in this alternative specification, news keep their primacy as explanatory factor of expectations, accounting for more that 60% of their variance.

7.3.2 Sign restrictions.

In this section we propose a different approach based on a mix of sing and zero restrictions, that is, a partial identification strategy. Accordingly, we only recover the effects of the news shocks and compare the results with the Cholesky scheme employed so far (always reported
in red in the pictures). We repeat this exercise for two different specifications: one with $\Delta ffr$ and the other using the federal funds rate.

As regards the former case we use a mixed strategy. Our theory suggests that news have delayed effects on $\Delta ffr$ which points at a zero contemporaneous restriction. Also, we posit a positive response of the policy indicator and a negative response of output after five horizons. In this fashion we leave the contemporaneous reaction of GDP unconstrained, while we impose the sign restriction only in a later moment, when the news shock is more likely to have materialized.

Figure 13 and Figure 14 show the corresponding impulse responses and variance decomposition. Following a news shock, the policy indicator takes four quarters to peak, while output, which is not constrained on impact, falls immediately. At first glance it is clear that the responses obtained with sign restrictions mimic closely the ones of the baseline model (in red). Indeed, the red lines are always well inside the 68% confidence bands. As regards variance decomposition, this methodology attributes high relevance to news in explaining GDP, inflation and market expectations. Again we find that anticipation plays role in monetary policy that should not be neglected.

Regarding the second example, with the federal funds rate as the policy instrument, we need to modify the identification restrictions. In fact, if output and inflation react to news on impact, any significant change in those variables would automatically translate into a change in the interest rates, according to the policy rule. That is, a zero contemporaneous restriction does not apply to the federal funds rate. Therefore, we use only sign restrictions to identify news, imposing negative effects on both output and inflation after five periods, and positive effects on expectations at short horizons. This allows us to take an agnostic stance on the reaction of the interest rate, which we leave unrestricted.

Results are presented in Figure 15 and Figure 16. The impulse responses of GDP, inflation and FF6 are in line with the theory and the variance decomposition is sensibly comparable to the baseline case. What really catches the eye is the response of the policy instrument. After an initial significant drop, the interest rate reverts its path and switches sign from negative to positive. The same behaviour shows up in the Cholesky case (red line in the graph) and it was already found in the extended VAR. This stands as an extra evidence that, through the policy rule, news shocks might feedback on the interest rate, even before their actual materialization.

7.3.3 SFP data.

As a last exercise we use a different measure of market expectations, namely SFP forecasts of the 3-Month Treasury Bill rate. Being that we cannot compute intra-day differences of
SPF data we keep it in levels. We choose the one and two quarters ahead projections, which relate more closely to the foresight horizons of the federal funds future at hand. Moreover, for comparison, we report the results obtained with the levels of the FF6.

SPF forecast is a rougher indicator of monetary news, being that it might be affected by other types of information and macro movements within the quarter. However, Figure 17 shows that it delivers impulse responses widely in line with those obtained with the FF6 in level. That is, both measures carry a similar informational content, with the difference that the latter is available at higher frequency and allows to better isolate the effects of policy announcements around meeting dates.

As regards financial variables, the relevant features are maintained: expectations are highly responsive to news especially on impact, while the policy indicator reacts with a certain delay. Results are slightly less satisfactory for the macro variables, given the initial price puzzle associated to news and the scarce reaction of output to surprises. However, the rest of the responses are in line with the theory and broadly consistent across specifications. Once more, the variance decomposition reported in figure Figure 18 attributes to monetary news a sizeable role both for the macroeconomic and the financial variables.

8 Conclusions.

In this paper we disentangled and compared the effects of monetary news and surprises on the macroeconomy. We have presented a small scale new Keynesian model with monetary foresight and we have showed that in this framework the structural MA representation of the economic variables is non-fundamental. This is in turn implies a failure of classical structural VARs.

We proposed two alternative solutions to the issue. The former, based on Blaschke matrix, is more a theoretical exercise that allows us to spell out clearly the bias implied by standard techniques. The latter resides in expanding the information set of the econometrician with market expectations on future policy developments. Through simulations we tested the performance of the proposed strategies and we concluded that non-invertibility is especially detrimental when the minimal root of the structural MA is close to zero. Also, we found that non-fundamentalness does not restrict the possibility of correctly recovering the unanticipated component alone, meaning that data is informationally sufficient for one of the two disturbances.

In the empirical section we have estimated a VAR complemented with market expectations and identified with Cholesky restrictions. This adds to previous literature in the sense that impulse responses and forecast error variance decomposition are estimated from the
data rather than obtained from DGSE simulations. Also, we showed that previous studies correctly identified the surprise component while neglecting foresight as a transmission channel of monetary policy. Therefore, the main contribution of the paper resides in presenting evidence that this channel is also relevant, and accounts in between 25 and 50% of the overall policy effects. Indeed, we found that news have sizeable effects on the real economy, prices and market expectations.

The impulse responses obtained are broadly consistent with the theory and in line with previous literature. Distinctly, we found a peculiar response of the short rate to news, testifying the indirect effect of announcement through the Taylor rule. Finally, we drew the conclusion that news are the main factor explaining market expectations, which is exactly the channel through which anticipation operates. Our results are robust to alternative proxies for market expectations, to different Cholesky ordering and to alternative identification strategies, relying on sign rather than zero restrictions.

This work establishes the empirical relevance of news in monetary policy and proposes solutions to the issue of non-invertibility, which is usually the main excuse to favour DSGE models over VARs. It goes without saying that it paves the way to a number of possible extensions. An example could be applying dynamic identification strategy, as in the Blaschke matrix case, to recover the structural disturbances. Furthermore, we have seen how news act through the expectation channel. It might be interesting to introduce a noise component and understand how miscommunication can affect market beliefs and cause aggregate effects, even when no underlying policy movement is implied. Another, interesting question is whether the role of news has changed over time and became more relevant in recent years with binding zero lower bound. Threshold VARs and consequent non-linearities could be an adequate approach to this problem. These questions and related issues, which are interesting to grasp a deeper understanding on the role of monetary news, are left for future research.
Appendix.

A Derivation of the model from new Keynesian theory.

Consider a standard New Keynesian model, as in Gali (2008). Preferences are given by:

\[ E \sum_{t=0}^{\infty} \beta^t \left( \log C_t - \frac{1}{1+\phi} N_t^{1+\phi} \right) \]

where \( N_t \) represents labor and \( C_t = \int_0^1 C_t(i) di \) is an aggregate consumption good. Each component \( C_t(i) \) is offered by a monopolistic firm which faces prices à la Calvo. In each period, the probability of keeping prices unchanged is \( \theta \). Output is demand driven, \( Y_t(i) = C_t(i) \), and firms hire labor at the flexible market wage \( W_t \) to produce with the linear technology:

\[ Y_t(i) = A_t N_t(i). \]

The period budget constraint for the agents takes the form:

\[ \int_0^1 P_t(i) C_t(i) di + Q_t B_t = B_{t+1} + W_t N_T + T_t \]

where \( B_t \) are nominal one-period bonds purchased at price \( Q_t \), \( P_t(i) \) is the price of good \( C_t(i) \) and \( T_t \) is a lump sum transfer that includes dividends from the firms.

Standard steps for the consumers and firm maximization problem and log-linearization around the zero inflation steady state lead to the New Keynesian Phillips curve and to the dynamic IS:

\[ \pi_t = \beta E_t \{ \pi_{t+1} \} + \kappa \bar{y}_t \]  

(26)

\[ \bar{y}_t = -(i_t - E_t \{ \pi_{t+1} \} - r^n_t) + E_t \{ \bar{y}_{t+1} \} \]  

(27)

where \( \kappa = (1+\phi)(1-\theta)(1-\beta\theta)/\theta \). Output gap \( \bar{y}_t \equiv y_t - y^n_t \) is defined as the deviation of output from its natural counterpart and \( r^n_t \) is the natural interest rate which we assume is unaffected by monetary shocks and is given by \( r^n_t = -\log \beta = i^\star \).

We also assume that central bank sets the nominal interest rate \( i_t = -\log Q_t \) according to the policy rule:

\[ i_t = i^\star + \phi_x \pi_t + \phi_y \bar{y}_t + v_t \]  

(28)

where \( \phi_x > 1, \phi_y > 0 \) and \( v_t \) is the exogenous disturbance that represents monetary policy shocks.
From (26),(27) and (28) it is possible to get to the following system of differential equations:

\[
\begin{pmatrix}
\tilde{y}_t \\
\pi_t
\end{pmatrix}
= A
\begin{pmatrix}
E_t \{\tilde{y}_{t+1}\} \\
E_t \{\pi_{t+1}\}
\end{pmatrix}
- B v_t
\tag{29}
\]

where:

\[
A \equiv \Omega \begin{pmatrix}
1 & 1 - \beta \phi \pi \\
\kappa & \kappa + \beta (1 + \phi y)
\end{pmatrix}
\]

\[
B \equiv \Omega \begin{pmatrix}
1 \\
\kappa
\end{pmatrix}
\]

\[
\Omega = \frac{1}{1 + \phi y + \kappa \phi \pi}
\]

The system has a unique solution as long as the eigenvalue of the matrix \(A\) lie inside the unit circle \(^{13}\). Forwards substitutions of (29) yield:

\[
\begin{pmatrix}
\tilde{y}_t \\
\pi_t
\end{pmatrix}
= -\sum_{k=0}^{\infty} A^k B E_t \{v_{t+k}\}.
\tag{30}
\]

Now on the spirit of Blanchard et al. (2013) we take a limit case of the model with \(\kappa \to 0\). In such case \(B = (\Omega \ 0)'\) and the matrix \(A\) is upper triangular. Its \(k^{th}\) power is again a upper triangular matrix of the form:

\[
A^k = \begin{pmatrix}
a_{11}^k & a_{12}^{(k-1)} a_{12} + a_{12}^{(k-1)} a_{22} \\
0 & a_{22}^k
\end{pmatrix}
\]

where \(a_{ij}^{(k)} = [A^k]_{i,j}\). Notice that the product \(A^k B\) always delivers,

\[
A^k B = \begin{pmatrix}
\Omega^{k+1} \\
0
\end{pmatrix}
\]

which substituted into (30) implies:

\[
\tilde{y}_t = -\Omega \sum_{k=0}^{\infty} \Omega^k E_t \{v_{t+k}\}.
\tag{31}
\]

Given that \(\kappa \to 0\) as \(\theta \to 1\) this completes the argument.

\(^{13}\)It can be shown that a necessary and sufficient condition for uniqueness is given by \(\kappa (\phi \pi - 1) + (1 - \beta) \phi y > 0\).
B Properties of $\omega_q(L)$.

The identity:

$$\omega_q(L) = L^q + \Omega \omega_{q-1}(L)$$

is easily proved.

$$\omega_q(L) \equiv \sum_{k=0}^{q} \Omega^k L^{q-k}$$

$$= L^q + \sum_{k=1}^{q} \Omega^k L^{q-k}$$

$$= L^q + \sum_{k=0}^{q-1} \Omega^{k+1} L^{q-1-k}$$

$$= L^q + \Omega \sum_{k=0}^{q-1} \Omega^k L^{q-1-k}$$

$$= L^q + \Omega \omega_{q-1}(L).$$

Moreover, notice that at $q = 1$ the only root is $r_1 = -\Omega$. For $q = 2$ we find two roots $r_{1,2} = \frac{-\Omega \pm \sqrt{3}\Omega i}{2}$ which both have modulus $|r_1| = |r_2| = \Omega$. For $q = 3$ we find three roots, $r_1 = -\Omega$, $r_{2,3} = \pm\Omega i$ whose modulus is again equal to $\Omega$.

Even if we do not provide a formal proof of it, the following picture exemplifies how this pattern is repeated for higher order of anticipation: the $q$ complex roots of $\omega_q(L)$ always fall within the open unit disc, and they all lie on the circle of radius $\Omega$.

![Figure 1: Roots of $\omega_q(L)$ for different anticipation periods.](image)

Whenever $\phi_y > 0$ or $\kappa > 0$ we have $\Omega < 0$ thus the roots always fall in the open unit disc. Only in the degenerate case of both $\phi_y = 0$ and $\kappa = 0$, $\Omega = 1$ and the roots of $\omega_q(L)$ are pushed on the unit circle.

---

14 Whenever $\phi_y > 0$ or $\kappa > 0$ we have $\Omega < 0$ thus the roots always fall in the open unit disc. Only in the degenerate case of both $\phi_y = 0$ and $\kappa = 0$, $\Omega = 1$ and the roots of $\omega_q(L)$ are pushed on the unit circle.
C Comparison with the literature.

Given that foresight is closely related to non-fundamentalness in VARs, previous literature privileged alternative methods to study the effects of monetary news shocks, namely, DSGE models (Laséen and Svensson, 2011; Milani and Treadwell, 2012; Del Negro et al., 2012) or univariate regressions (Gürkaynak et al., 2005; Campbell et al., 2012). For the sake of comparison, we can show how their models translate to our framework and understand the implied VAR failures.

These authors assume that in every period the central bank delivers a vector of information
\[ \eta_t = (\eta_t, 0, \eta_t, 1, \ldots, \eta_t, q)^t \]
where each element \( \eta_{t,s} \) represent news delivered at time \( t \) and affecting the interest rate at \( t + s \). Equation (3) becomes:
\[ v_t = \sum_{s=0}^{q} \eta_{t-s,s} \]  
where \( \eta_{t,0} \) is the monetary surprise \( \varepsilon_t \). In addition there are \( q \) different shocks affecting the economy at horizons 1, 2, \ldots, \( q \), respectively.

The expectation at time \( t \) of \( v_{t+j} \) is given by a summation of past news with different impact horizons:\(^{15}\):
\[ E_t\{v_{t+j}\} = \begin{cases} \sum_{s=0}^{q-j} \eta_{t-s,s+j} & \text{for } 0 < j \leq q \\ 0 & \text{for } j > q. \end{cases} \]  
(33)

Now we can plug (33) in (4) to obtain:
\(^{15}\)To better understand the formulation in (33) consider the values of \( j = 0 \) and \( j = q \). On the one hand, time \( t \) expectation of \( v_t \) reads:
\[ E_t\{v_t\} = \eta_{t,0} + \eta_{t,1} + \ldots + \eta_{t,q}. \]

This equation states that today’s changes in the policy rate are given by a summation of all the available information: the contemporaneous \( \eta_{t,0} \); the news delivered one period before and acting one period later; \( \ldots \); and so on down to the oldest new \( \eta_{t-q,q} \) received \( q \) period before and materializing with \( q \) periods of delay.

On the other the expected value at time \( t + q \):
\[ E_t\{v_{t+q}\} = \eta_{t,q} \]
can only be based on information received today for \( q \) periods into the future. This is because all other shocks affecting \( v_{t+q} \) will be revealed in the future and their present expectation equals zero.

All intermediary cases for \( 0 < j < q \) follow the same logic: only available past information, whose impact horizon is \( t + j \), can be used to generate the expectation. Notice that the further in time we project \( v_{t+s} \) the less news available we have. This explains the decreasing number of addends in \( \sum_{s=0}^{q-j} \eta_{t-s,s} \).
\[
\begin{align*}
\tilde{y}_t &= -\Omega \left( \sum_{s=0}^q \eta_{t-s,s} + \Omega \sum_{s=0}^{q-1} \eta_{t-s,s+1} + \ldots + \Omega^q \sum_{s=0}^0 \eta_{t-s,s+q} \right) \\
&= -\Omega \left( \eta_{t,0} + \sum_{s=0}^1 \Omega^s \eta_{t-1+s,1} + \ldots + \sum_{s=0}^q \Omega^s \eta_{t-q+s,q} \right) \\
&= -\Omega \left( \varepsilon_t + \omega_1(L) \eta_{t,1} + \ldots + \omega_q(L) \eta_{t,q} \right). \tag{34}
\end{align*}
\]

where the first equivalence is obtained expanding the summations and grouping anticipated shocks with the same horizon of impact. The last equation is given by the definition of \( \omega_q(L) \) in (8) and by the fact that \( \eta_{t,0} \) is \( \varepsilon_t \) in our setting.

It is now possible to rewrite (32) and (34) in matrix notation as:

\[
\begin{pmatrix}
\varepsilon_t \\
\eta_{t,1} \\
\vdots \\
\eta_{t,q}
\end{pmatrix}
= 
\begin{pmatrix}
1 & L & \cdots & L^q \\
-\Omega & -\Omega \omega_1(L) & \cdots & -\Omega \omega_q(L)
\end{pmatrix}
\begin{pmatrix}
\varepsilon_t \\
\eta_{t,1} \\
\vdots \\
\eta_{t,q}
\end{pmatrix}. \tag{35}
\]

where the rectangular matrix in (35) is \( 2 \times (q + 1) \) and the vector of shocks contains \( q + 1 \) fundamental innovations.

Once more, we rely on a simple example to clarify the general formulation. With \( q = 2 \), the policy residual is given by:

\[
v_t = \varepsilon_t + \eta_{t-1,1} + \eta_{t-2,2} \tag{36}
\]

and the implied MA representation is:

\[
\begin{pmatrix}
\varepsilon_t \\
\eta_{t,1} \\
\eta_{t,2}
\end{pmatrix}
= 
\begin{pmatrix}
1 & L & L^2 \\
-\Omega & -\Omega \omega_1(L) & -\Omega \omega_2(L)
\end{pmatrix}
\begin{pmatrix}
\varepsilon_t \\
\eta_{t,1} \\
\eta_{t,2}
\end{pmatrix}. \tag{37}
\]

Clearly (37) contains a rectangular matrix with more than two columns. VARs techniques are not viable in this setting because the number of shocks exceeds the number of dynamically independent time series. Therefore, to map the system into our framework it is necessary to simplify it to two structural shocks. This can be done assuming either:

\[
\begin{align*}
v_t &= \varepsilon_t + \eta_{t-1} & (v1) \\
v_t &= \varepsilon_t + \eta_{t-2} & (v2) \\
v_t &= \eta_{t-1} + \eta_{t-2} & (v3) \\
v_t &= \varepsilon_t + \eta_{t-1} + \eta_{t-2} & (v4)
\end{align*}
\]

The cases (v1) and (v2) correspond to the baseline model with \( q = 1 \) and \( q = 2 \), which have already been discussed. Assuming (v3) - i.e., no monetary surprises - is equivalent to
removing the first column from the MA matrix in (37). In this case the system would feature a root for \( L = 0 \), showing that a purely news-based model is also non fundamental. This evidence seems to support the claim of Milani and Treadwell (2012) that SVAR techniques are doomed to fail when news shocks are involved.

However, (v4) provides with an example that foresight does not necessarily imply non fundamentalness. In this latter case, the two news shocks acting at one and two horizons have been summarized in one single shock with a double delayed effect (possibly the closest mapping between \( \hat{\eta}_t \) and \( \eta_t \)). Under this specification the expectations of \( v_t \) is:

\[
E_t\{v_{t+j}\} = \begin{cases} 
\varepsilon_t + \eta_{t-1} + \eta_{t-2} & \text{for } j = 0 \\
\eta_t + \eta_{t-1} & \text{for } j = 1 \\
\eta_t & \text{for } j = 2 \\
0 & \text{for } j \geq 3.
\end{cases}
\]

and the equation for \( \tilde{y}_t \) turns into:

\[
\tilde{y}_t = -\Omega(\varepsilon_t + \Omega(1 + \Omega)\eta_t + (1 + \Omega)\eta_{t-1} + \eta_{t-2})
\]

leading to a MA representation of the kind:

\[
\begin{pmatrix}
v_t \\
\tilde{y}_t
\end{pmatrix} = \begin{pmatrix}
1 & L + L^2 \\
-\Omega & -\Omega^2(1 + \Omega) - \Omega(1 + \Omega)L - \Omega L^2
\end{pmatrix}
\begin{pmatrix}
\varepsilon_t \\
\eta_t
\end{pmatrix}
\]

The determinant of the system vanishes for \( L = -(1 + \Omega) \) which is outside the unit circle. Contrary to general wisdom, no problem of fundamentalness is found in this case. This apparently surprising result can be explained looking at the impulse responses of \( \tilde{y}_t \).

As noted by Beaudry and Portier (2014), news rich processes whose effects are delayed but not monotonically increasing up to \( q \) might have an invertible representation. Indeed, as exemplified in Figure 2, under (v4) the highest effect of the news shock is delayed only once, even if there are two periods of foresight. Conversely, assuming (v2) the impulse response of output peaks exactly after 2 horizons and generates a non-invertible representation.
Figure 2: Comparison of the the response of \( v_t \) and \( \tilde{y}_t \) to the news shock. Left hand columns refer to \( (v^2) \) and right hand columns refers to \( (v^4) \).

D The general model.

Results are presented in the paper under specific parameter values and the assumption of i.i.d. policy shocks. In order to extend the model we start from a general calibration of (30) and we allow for richer dynamics of the policy shock, as in (2):

\[
v_t = \rho(L) \eta_{t-q} + \rho(L) \varepsilon_t.
\]

where \( v^a_t \) and \( v^u_t \) are the anticipated and unanticipated component respectively. As customary, \( \rho(L) = \sum_{k=0}^{\infty} \rho^k L^k \) with \( \rho < 1 \) and \( L \) is the lag operator.

Expectations at \( t \) of \( v^a_t \) and \( v^u_t \) are obtained with linear projections using the Weiner-Kolmogorov formula. Exploiting the properties of the annihilator operator \([\cdot]_+\), which sets to zero all negative exponents of the lag polynomial, we easily obtain projections for the unanticipated component:

\[
E_t\{v^u_{t+s} \mid I_t\} = \left[ \frac{\rho(L)}{L^s} \right]_+ \varepsilon_t = \left[ \sum_{i=0}^{\infty} \rho^{i+s} L^i \right]_+ \varepsilon_t = \rho^s \rho(L) \varepsilon_t
\] (41)

As regards the anticipated component we need to be more cautious, since relevant information is received \( q \) periods beforehand:

\[
E_t\{v^a_{t+s} \mid I_t\} = \left[ \frac{\rho(L)}{L^s} \right]_+ \eta_{t-q} = \left[ \sum_{i=0}^{\infty} \rho^{i} L^{i+q-s} \right]_+ \eta_t
\]
and we need to consider horizons preceding and exceeding $q$ separately:

$$E_t\{v_{t+s} | I_t\} = \begin{cases} \sum_{i=0}^{\infty} \rho^i L^{q+s} \eta_t = \rho(L)L^{q-s}\eta_t & \text{for } s \leq q \\ \sum_{i=0}^{\infty} \rho^i L^{q+s} \eta_t = \rho^s q \eta_t & \text{for } s > q. \end{cases} \quad (42)$$

Equations (41) and (42) imply:

$$E_t\{v_{t+s}\} = \begin{cases} \rho(L)L^{q-s}\eta_t + \rho^s q \eta_t \varepsilon_t & \text{for } s \leq q \\ \rho^s q \eta_t + \rho^s q \eta_t \varepsilon_t & \text{for } s > q. \end{cases} \quad (43)$$

which, as expected, features no contemporaneous effect of $\eta_t$.

Substituting (43) in (30):

$$\begin{aligned} \begin{pmatrix} \tilde{y}_t \\ \varepsilon_t \end{pmatrix} &= -\sum_{s=0}^{\infty} A^s B E_t\{v_{t+s}\} \\ &= -\sum_{s=0}^{q-1} A^s B [\rho(L)L^{q-s}\eta_t + \rho^s q \eta_t \varepsilon_t] - \sum_{s=q}^{\infty} A^s B [\rho^s q \eta_t + \rho^s q \eta_t \varepsilon_t] \\ &= -\sum_{s=0}^{\infty} A^s B [\rho^s q \eta_t \varepsilon_t] - \sum_{s=q}^{\infty} A^s B [\rho^{s-q} \rho(L)\eta_t] - \sum_{s=q+1}^{\infty} A^s B [\rho^{s-q} \rho(L)\eta_t]. \end{aligned} \quad (44)$$

yields convolute infinite summations. In order to simplify this expression we proceed by sections as follows:

**a:** Given that by assumption the matrix $A$ has its eigenvalues inside the unit circle and that $\rho < 1$, the infinite summation in (a) converges to:

$$-\sum_{s=0}^{\infty} A^s B [\rho^s q \eta_t \varepsilon_t] = -(I - A\rho)^{-1} B \rho(L)\varepsilon_t$$

where:

$$\psi(L) = \begin{pmatrix} \psi_1(L) \\ \psi_2(L) \end{pmatrix} = -(I - A\rho)^{-1} B \sum_{s=0}^{\infty} \rho^s L^s$$

is a $2 \times 1$ vector in the lag operator.

**b:** This part of the summation contains $q$ MA processes:

$$\sum_{s=0}^{q} A^s B [\rho(L)L^{q-s}\eta_t] = A^q B \rho(L)\eta_t + A^{q-1} B \rho(L)\eta_{t-1} + \cdots + A^0 B \rho(L)\eta_{t-q}.$$ 

Expanding all its addends we find one $\eta_t$, two $\eta_{t-1}$, and so on till $\eta_{t-q}$ which appears $q+1$ times. All elements lagged more than $q$ will also appear $q+1$ times. Grouping together coefficients relative to the same lag we can rearrange the equation as:

36
\[ \sum_{s=0}^{q} A^s B \left[ \rho(L)L^{q-s}\eta_t \right] = \sum_{s=0}^{q-1} D_s BL^s \eta_t + \sum_{s=q}^{\infty} \rho^{s-q} D_q BL^s \eta_t. \]  

(b1)

where:

\[ D_s \equiv \sum_{m=0}^{s} A^q \rho^{s-m}. \]  

(46)

Notice that this matrix of coefficient varies with \( s \) within the first \( q - 1 \) lags, while from \( q \) onward (b1) always features:

\[ D_q = \sum_{m=0}^{q} A^m \rho^m \]

multiplied by increasing powers of \( \rho \). This fact allows us to divide the infinite summation at \( q - 1 \) (instead of \( q \)), which mirrors the misalignment between anticipation and surprise and will be useful in further steps.

(b2): We proceed as in (a). Rearranging the counter of the summation and solving the (convergent) sum we have:

\[
\sum_{s=q+1}^{\infty} A^s B \left[ \rho^{s-q} \right. \rho(L) \eta_t] = A^{q+1} \rho \sum_{s=0}^{\infty} (Ap)^s B \rho(L) \eta_t \\
= A^{q+1} \rho (I - Ap)^{-1} B \rho(L) \eta_t \]  

(b2)

where:

\[ C_q \equiv A^{q+1} \rho (I - Ap)^{-1} \]  

(47)

is a matrix of coefficients decreasing in the foresight horizon\(^\text{16}\). Adding (b2) and (b1) yields:

\( \text{Both (a) and (b2) deliver a re-scaling of the MA coefficients of } \rho(L). \text{ Few words on the intuition behind such result are in order. Notice that – as in (41) – linear projections for } t + s \text{ of an AR}(1) \text{ process contain the } s^{th} \text{ power of } \rho. \text{ This implies that further away projections are more heavily discounted by } \rho. \text{ In the multivariate case the powers of } A, \text{ which approach zero as the expectation horizon increases, act in a similar fashion. Therefore there are two sources of discounting: } \rho \text{ – coming from the AR nature of the policy shocks – and } A \text{ – coming from the recursive substitutions of the forward looking system (30).}

For the unanticipated shock agents can only create expectations though linear forecast at time \( t \), therefore the discounting due to forward substitutions and due to linear projections go in parallel (notice that we always have } A^s \rho^s). \text{ Conversely, regarding } \eta_t, \text{ agents have (anticipated) information for the first } q \text{ periods, out of which they can generate their expectations. After that horizon, that is the (b2) part of the summation, they are again constrained to use linear projections with the same philosophy employed for (a). This is enough to introduce a misalignment in discounting: since projections only begin in period } q + 1, \text{ results of (b2) are comparable to (a) but are additionally discounted by } A^{q+1} \text{ (which makes the difference between } C_q \text{ and } (I - Ap)^{-1}). \text{ Thus both (a) and (b2) are a re-scaling of different magnitude of the same process, and what is more relevant to understand the difference between } \eta_t \text{ and } \epsilon_t \text{ are the first } q \text{ periods – captured by the (b1) part of the summation.}

\( \text{37} \)
\[-[(b_2) + (b_1)] = - \sum_{s=0}^{\infty} C_q B \rho^s L^s \eta_t - \sum_{s=0}^{q-1} D_s B L^s \eta_t - \sum_{s=q}^{\infty} \rho^{s-q} D_q B L^s \eta_t \]
\[
= - \sum_{s=0}^{q-1} (C_q B \rho^s + D_s) B L^s \eta_t - \sum_{s=q}^{\infty} (C_q \rho^s + \rho^{s-q} D_q) B L^s \eta_t \\
= \phi^{q-1}(L) \eta_t - \sum_{s=q}^{\infty} (I - Ap)^{-1} \rho^{s-q} B L^s \eta_t \\
= \phi^{q-1}(L) \eta_t - \sum_{s=0}^{\infty} (I - Ap)^{-1} \rho^s B L^s \eta_t \\
= \phi^{q-1}(L) \eta_t + \psi(L) L^q \eta_t. \\
\]

where:
\[
\phi^{q-1}(L) = \begin{pmatrix} \phi_1^{q-1}(L) \\ \phi_2^{q-1}(L) \end{pmatrix} = - \sum_{s=0}^{q-1} (C_q \rho^s + D_s) B L^s \\
\]

is a 2 \times 1 vector in the lag operator and the third equality of (b) makes use of:
\[
C_q \rho^s + \rho^{s-q} D_q = A^{q+1} \rho (I - Ap)^{-1} \rho^s + \rho^{s-q} \sum_{m=0}^{q} A^m \rho^m \\
= A^{q+1} \rho^s + \rho^{s-q} \sum_{k=0}^{\infty} A^k \rho^k + \sum_{m=0}^{q} A^m \rho^m \\
= \rho^{s-q} \left( A^{q+1} \rho^s + \sum_{k=0}^{\infty} A^k \rho^k + \sum_{m=0}^{q} A^m \rho^m \right) \\
= \rho^{s-q} \left( \sum_{k=q+1}^{\infty} A^k \rho^k + \sum_{m=0}^{q} A^m \rho^m \right) \\
= (I - Ap)^{-1} \rho^{s-q}. \\
\]

Plugging (a) and (b) in (44) allows us to rewrite the system (30) as a MA representation in output gap and inflation as follows:
\[
\begin{pmatrix} y_t \\ \pi_t \end{pmatrix} = \begin{pmatrix} \psi_1(L) & \phi_1^{q-1}(L) + \psi_1(L) L^q \\ \psi_2(L) & \phi_2^{q-1}(L) + \psi_2(L) L^q \end{pmatrix} \begin{pmatrix} \varepsilon_t \\ \eta_t \end{pmatrix}. \\
\]

When we condition the VAR on monetary policy instead of inflation, we obtain:
\[
\begin{pmatrix} v_t \\ \hat{y}_t \end{pmatrix} = \begin{pmatrix} \rho(L) & \rho(L) L^q \\ \psi_1(L) & \gamma(L) \end{pmatrix} \begin{pmatrix} \varepsilon_t \\ \eta_t \end{pmatrix}, \\
\]

where \( \psi_1(L) \) and \( \phi_1^{q-1}(L) \) are the upper elements of the vectors \( \psi(L) \) and \( \phi^{q-1}(L) \) and \( \gamma(L) \) simplifies the notation:
\[
\gamma(L) \equiv \phi_1^{q-1}(L) + \psi_1(L) L^q. \\
\]
D.1 Roots of the MA determinant.

The study of the determinant $D(z)$ is analytically non trivial, and calls for numerical methods. We proceed as follows: we first calibrate the parameters$^{17}$. Then we build the relevant matrices to obtain the coefficients of $\psi(L)$ and $\phi(L)$. In this case we need to deal with a MA($\infty$), which for ease of computation we approximate with a MA of finite (but substantially high) order$^{18}$. Finally we compute $\text{Det}(D(z))$ and verify whether the characteristic polynomial has roots falling within the unit circle.

Results for $D(z)$ are shown in Figure 3 and are broadly in line with the findings of the limit-case model. Under the baseline calibration the systems features roots in the open unit disc at all values of $q$. Therefore, classical VARs need to be corrected either with Blaschke matrix or with additional information.

![Figure 3: Roots of the determinant of the MA representation with $\tilde{y}$ and $v$. The MA matrix is truncated at 100 elements. Specification for different foresight horizons $q$ are reported together in the same graph. The dashed line represent the unit circle in the complex plane.](image)

D.2 Deriving the toy model from the general case.

From the general setting, it is simple to derive the model in its limit-case calibration, with $\theta = 1$ and $\rho = 0$. Trivially:

$$v_t = \varepsilon_t + L^q \eta_t$$

and

$$\psi_1(L) = -1 - (1 - \Omega \rho) \phi(L) = -\Omega.$$  

$^{17}$The Baseline calibration is as in Gali (2008) p.52: $\theta = 2/3; \sigma = 1; \psi = 1; \beta = 0.99; \phi_y = 1.5; \phi_y = 0.5/4; \rho = 0.5$.

$^{18}$Namely we used a MA(100), whose last coefficients are already sufficiently close to zero to be negligible.
The operator $\phi_1^{q-1}(L)$ turns to:

$$
\phi_1^{q-1}(L) = -\sum_{s=0}^{q-1} \sum_{m=0}^{s} \Omega^{q+1-m} \rho^{s-m} L^s \\
= -\sum_{s=0}^{q-1} \Omega^{q+1-s} L^s \\
= -\Omega^2 \sum_{s=0}^{q-1} \Omega^{q-1-s} L^s \\
= -\Omega^2 \sum_{s=0}^{q-1} \Omega^{s} L^{q-1-s} \\
= -\Omega^2 \omega_{q-1}(L).
$$

which plugged into $\gamma(L)$ simplifies it to:

$$
\gamma(L) = -\Omega L^q - \Omega^2 \omega_{q-1}(L) \\
= -\Omega(L^q + \Omega \omega_{q-1}(L)) \\
= -\Omega \omega_q(L).
$$

Now, replacing the above expressions in (50), we obtain:

$$
\begin{pmatrix}
\hat{v}_t \\
\hat{y}_t
\end{pmatrix} = 
\begin{pmatrix}
1 & L^q \\
-\Omega & -\Omega \omega_q(L)
\end{pmatrix}
\begin{pmatrix}
\hat{v}_t \\
\hat{y}_t
\end{pmatrix}_M(L)
$$

which should make clear the parallel with the toy model discussed in the paper.
Figure 4: VAR and Blaschke methods compared. Red line is d.g.p, the blue line is the result of a standard SVAR, the green line is the correction of the SVAR by a Blaschke matrix. Top panel obtained with baseline calibration as in Gali, while the bottom panel features higher persistence of the policy shocks. Confidence bands are obtained with 1000 simulations from the d.g.p.
Figure 5: Adding expectation in the VAR. The red line is the d.g.p., \( R^{(1)}(L) \) in the top panel and \( R^{(2)}(L) \) in the bottom panel. The blue line is the result of standard SVAR. Confidence bands are obtained with 1000 simulations from the d.g.p.
E.2 VAR analysis.

E.2.1 Toy VAR.

Figure 6: Toy VAR with $\Delta ffr$ and $\Delta fff$ from Barakchian and Crowe (2010). Sample from 89Q1 to 08Q2. Individual measures for $h$-ahead contracts are reported in green. The factor measure that summarizes all of them is black.
E.2.2 Baseline VAR.

Figure 7: Baseline VAR with GDP, GDPDEF, $\Delta ffr$ and $\Delta fff$. Sample from 89Q1 to 08Q2 on Barakchian and Crowe (2010) data. Green lines are the measure derived from $h$-ahead future contract for $h = 1, \ldots, 6$. Black line is the factor measures summarizing all the individual contracts. 95% and 68% bands computed with bootstrap methods refer to the latter.
Figure 8: Variance decomposition of baseline VAR with GDP, GDPDEF, Δffr and Δfff (factor measure for the market expectations). One standard deviation bands in gray.
E.2.3 Extended VAR.
Figure 9: Christiano et al. (2005) VAR with timing restriction. The variables (in order) are GDP, CONS, GDPDEF, Real Investment, Wages, Productivity, FFR, log(FF6), Profits and M2 on the sample 89Q1-14Q4. All variables are in log levels, but M2 which is in growth rates. 95% and 68% bands computed with bootstrap methods. The red line compares the results obtained from the baseline case with no news specification.
Figure 10
Figure 10: Variance decomposition of Christiano et al. (2005) VAR with timing restriction. The variables (in order) are GDP, CONS, GDPDEF, Real Investment, Wages, Productivity, FFR, log(FF6), Profits and M2 on the sample 89Q1-14Q4. All variables are in log levels, but M2 which is in growth rates. One standard deviation bands in gray.
E.3 Robustness analysis.

E.3.1 Policy indicator ordered first.

Figure 11: VAR with $\Delta ffr$, $\Delta fff$, ordered first. Sample from 89Q1 to 08Q2. Green lines are the measure derived from $h$-ahead future contract for $h = 1, \ldots, 6$. Black line is the factor measures summarizing all the individual contracts. 95% and 68% bands computed with bootstrap methods refer to the latter. The red line is the baseline case of Figure 7 with the reversed variable ordering.
Figure 12: Variance decomposition of VAR with $\Delta ffr$, $\Delta fff$, GDP, GDPDEF. One standard deviation bands in gray.
E.3.2 Sign restrictions.

Figure 13: Baseline VAR with $\Delta ffr$, $\Delta fff$, GDP, GDPDEF. News shock identified zero and sign restrictions. Results are compared to the Cholesky scheme. Restrictions on the news shock effects: (-) for GDP at h=5 and (+) for $\Delta ffr$ at h=5 and zero at h=0.

Figure 14: Variance decomposition of baseline VAR with $\Delta ffr$, $\Delta fff$, GDP, GDPDEF identified with sign restrictions. One standard deviation bands in gray.
Figure 15: Baseline VAR with FFR, FF6 GDP, GDPDEF, identified with zero and Sign restrictions. News shock identified with sign restrictions. Results are compared to the Cholesky scheme. Restrictions set are on the news shock effects: (-) for GDP at h=5; (-) for inflation at h=5; (+) for FF6 at h=2.

Figure 16: Variance decomposition of baseline VAR with FFR, FF6, GDP, GDPDEF, identified with zero and Sign restrictions. One standard deviation bands in gray.
E.3.3 SPF data.

Figure 17: Baseline VAR with GDP, GDPDEF, $\Delta ffr$ and FF6 (black line) or 1 and 2 quarters ahead SPF forecast of T-bill (green lines). The policy shock has been cumulated to obtain an I(1) series. 95% and 68% bands computed with bootstrap methods refer to the FF6 case.
Figure 18: Variance decomposition of baseline VAR with GDP, GDPDEF, \( \Delta ff \) and 1 and 2 quarters ahead SPF forecast of T-bill. One standard deviation bands in gray refer to SPF1.
References


57


