Evidence on News Shocks under Information Deficiency

Nelimarkka, Jaakko

University of Helsinki

10 August 2017

Evidence on News Shocks under Information Deficiency*

Jaakko Nelimarkka[†]
University of Helsinki

August 10, 2017

Abstract

News shocks about future productivity can be correctly inferred from a conventional VAR model only if information contained in observables is rich enough. This paper examines news shocks by means of a noncausal VAR model that recovers economic shocks from both past and future variation. As noncausality is implied by nonfundamentalness, the model solves the problem of insufficient information per se. By the impulse responses derived from the model, variables react to the anticipated structural shocks, which are identified by exploiting future dependence of investment with respect to productivity. In the U.S. economy, news about improving total factor productivity moves investment and stock prices on impact, but these responses are likely affected by a parallel increase in productivity. The news shock gradually diffuses to productivity and generates smooth reactions of forward-looking variables.

JEL classification: C18, C32, C53, E32
Keywords: News shocks, Structural VAR analysis, Nonfundamentalness, Noncausal VAR

1 Introduction

In news-driven business cycles, economic agents observe signals about future productivity, creating fluctuations in forward-looking variables before the news materialises. However, dynamics driven by these news shocks likely generates a situation where the information set of economic agents relevant for decision making is broader than what an econometrician observes. Consequently, a conventional vector autoregressive model (VAR) used for empirical validation may produce misleading results about the implications and significance of the news shocks.

Starting from the seminal paper of Beaudry and Portier (2006), the structural VAR (SVAR) methodology identifies the news shock as a shock that changes stock

*I thank Markku Lanne for his guidance and helpful suggestions. I am also thankful to Jani Luoto, Karel Mertens, Kristoffer Nimark, Antti Ripatti, Pentti Saikkonen and Aino Silvo as well as to participants at the HECER workshops, the Cornell Macro Lunch seminar, the DIW Macroeconometric workshop, the Spring Meeting of Young Economists and the Nordic Econometric Meeting for useful comments. Financial support from the Academy of Finland, the OP Group Research Foundation and the Yrjö Jahnsson Foundation is gratefully acknowledged.

[†]Department of Political and Economic Studies and Helsinki Center of Economic Research, P.O. Box 17, FI-00014 University of Helsinki. E-mail: jaakko.nelimarkka@helsinki.fi.

prices on impact but has delayed effects on total factor productivity (TFP).[1] However, by the expectations of forward-looking agents, the news shocks imply noninvertibility of the theoretical moving average (MA) representation for a typical small number of variables included to the VAR model (Leeper et al., 2013; Forni et al., 2014). The resulting nonfundamentalness problem prevents obtaining the structural shocks and their impulse responses from a causal autoregressive representation of the observed variables.[2]

Nonfundamentalness or noninvertibility eventually boils down to the fact that the observables do not contain all relevant state variables of the economy (Fernández-Villaverde et al., 2007). The problem can thus be avoided by including enough relevant variables to the VAR which then approximates the underlying MA representation. In the studies of Beaudry and Portier (2006) and Barsky and Sims (2011), nonfundamentalness is absent if the variables included are sufficiently forward-looking to achieve invertibility. Alternatively, augmenting the VAR with factors extracted from large-dimensional macroeconomic data can resolve the information deficiency (Forni and Gambetti, 2014; Forni et al., 2014).[3]

This study introduces a new approach to structural analysis of news shocks under insufficient information while still imposing few restrictions on the underlying economic process. In place of leaning on the information in observables only, I consider a noncausal representation that includes lead terms, arising as a result of nonfundamentalness. I make the representation operational by estimating the noncausal VAR model of Lanne and Saikkonen (2013).[4] Under nonfundamentalness, a causal model produces errors that are linear combinations of the past and present shocks (Lippi and Reichlin, 1994b), while the noncausal VAR model filters out, through its distinct lag and lead polynomials, an error term consisting of fundamental shocks anticipated by the economic agents.

As the noncausal VAR nests a causal VAR model by its lags terms, the approach conveniently complements the VAR analysis and facilitates checking fundamentalness of observables for the underlying economic process. If data suggest selecting a positive number of leads, nonfunfamentalness is present reflected by the lead terms of the model. However, to distinguish noncausal and causal representation, it is necessary to deviate from Gaussianity of the error term. As one particular deviation, I use a multivariate t-distribution, which adds a common volatility term to the Gaussian structural shocks. The noncausal VAR model can then be identified and

[1] Beaudry and Portier (2014) review the recent literature on news-driven business cycles.

[2] Hansen and Sargent (1991) and Lippi and Reichlin (1994b) provide earlier discussion on the topic. For a more recent review, see Lütkepohl (2014) and Beaudry and Portier (2014). In addition to technology-related news shocks, nonfundamentalness is in the fiscal policy under anticipation of economic agents. See Leeper et al. (2013) and the references therein.

[3] The issue of nonfundamentalness can also be avoided by imposing enough theoretical structure such as in Schmitt-Grohé and Uribe (2012) who directly estimate a theoretical model with news shocks. Plagborg-Møller (2016) and Barnichon and Matthes (2015) have proposed alternative methodologies to estimate possibly noninvertible representations directly. However, these strategies require strong prior information on the propagation mechanism of news shocks, making the results heavily dependent on the structural assumptions. A study of Arezki et al. (2017) derives the impulse responses using a narrative approach by collecting proxies for news shocks from oil discoveries. Kurmann and Sims (2017) use revisions in the TFP measure to identify the news shock.

[4] For statistical theory of noncausal autoregressions, see Rosenblatt (2000). To recent contributions on noncausal and noninvertible autoregressions belong, amongst others, Lanne and Saikkonen (2013), Davis and Song (2012) and Gouriéroux and Zakoian (2013).

estimated by maximum likelihood (ML) following Lanne and Saikkonen (2013).

I make two methodological contributions in the paper. First, I show how to conduct structural analysis under nonfundamentalness with the noncausal VAR model. In particular, the error term can be mapped into anticipated structural shocks, and the impulse response analysis is based on the two-sided MA representation of the model. Second, I propose an identification scheme where the shocks are recovered by finding a rotation that moves productivity and the noncausal, anticipatory component of a forward-looking variable the most. Specifically, the identification is based on the observation that nonfundamentalness induces the lead terms of the model to matter for a forward-looking variable such as investment.

I examine the performance of the approach by means of Monte Carlo simulations of a New Keynesian model augmented with news shocks. When the model implies fundamentalness for observables, the noncausal model is outperformed by the causal model that is capable of replicating the true impulse responses. On the other hand, noncausality arises through choosing lead terms to the model for variables inducing nonfundamentalness. Identifying the news shock in the selected noncausal VAR with the proposed strategy, the model reproduces the theoretical impulse responses while a causal VAR fails to reveal the initial reactions as the identification is based on a misspecified error term strongly weighted by the lagged news shocks. Hence, although the nonfundamentalness issue may only slightly distort the results on news shocks in situations with a persistent tehnological process and the discount factor close to unity as argued by Sims (2012), Beaudry and Portier (2014) and Beaudry et al. (2015), the performance of a causal model may well be considerably deteriorated when not all relevant variables are included to the VAR.

In the postwar U.S. data, I find support for non-Gaussianity of the error term of the estimated VAR, which allows me to compare causal and noncausal models. Moreover, noncausality is found to be a strong feature of data, facilitating the identification of a news shock on TFP with the proposed strategy. In response to the news shock, investment, hours, consumption and output increase on impact and inflation turns negative. The news shock prompts a continuous, steady improvement of TFP and smooth responses of macroeconomic variables, in contrast to the stronger reactions found by Beaudry and Portier (2006). In this respect, my results are in line with Barsky and Sims (2011) and Forni et al. (2014) who measure limited role for the news shock in the short run.[5] Hence, under information deficiency, a causal VAR model may overemphasise the relevance of news shocks, as also found by Forni et al. (2014), as a result of nonfundamental errors capturing the timing of the shock incorrectly.

The rest of the paper is organised as follows. Next, in Section 2, I review general results on noncausality and nonfundamentalness, present the methodology based on the noncausal VAR, and propose the identification scheme for finding news shocks. Section 3 presents Monte Carlo simulation results. In Section 4, I consider the empirical evidence on news shocks. The last section concludes.

[5] For further empirical results, see Schmitt-Grohé and Uribe (2012), Beaudry et al. (2015), Barsky et al. (2015) and Kurmann and Otrok (2013).

2 Theory

When news shocks drive the economy, the past observables may not contain sufficient information to recover the structural shocks of interest, which eventually leads to the existence of a noncausal representation. This section presents an approach to study the effects of news shocks based on a noncausal VAR model of Lanne and Saikkonen (2013). The implied two-sided MA representation of the model facilitates the derivation of impulse responses to the shocks that affect the current state of the economy before being observed by the econometrician but already anticipated by the forward-looking economic agents.

2.1 Nonfundamentalness and Noncausality

I start by reviewing general results on noninvertibility and nonfundamentalness, and demonstrate how they give rise to noncausality.[6] Consider the equilibrium of a linearised macroeconomic model for k observed variables in y_t with a vector autoregressive moving average (VARMA) representation

$$A(L)y_t = B(L)u_t, \tag{1}$$
$$A(L) = I - A_1L - \ldots - A_pL^p,\ B(L) = B_0 + B_1L + \ldots + B_dL^d,$$

where u_t is a vector containing the k uncorrelated structural shocks driving the economy, and $\mathrm{E}_t[u_{t+j}] = 0$ when $j > 0$ and $\mathrm{E}_t[u_{t+j}] = u_{t+j}$ for $j \leq 0$. $\mathrm{E}_t[\cdot]$ denotes the expectation conditional on the information set of the agents. $A(L)$ and $B(L)$ are $(k \times k)$ matrix polynomials, with L the usual lag operator, that determine the unique equilibrium of the model in terms of finite lags up to a truncation. $A(L)$ is assumed to be stable, implying an MA representation $y_t = A(L)^{-1}B(L)u_t$.

When the MA polynomial $B(L)$ in (1) is invertible in the past, i.e., $|B(z)|$ has no roots inside the unit circle, the structural shocks and the impulse responses can be obtained with a conventional causal VAR(p) model,

$$C(L)y_t = \varepsilon_t,\ C(L) = I - C_1L - \ldots C_pL^p, \tag{2}$$

from the reduced-form error term $\varepsilon_t = B_0u_t$ after imposing identifying restrictions on matrix B_0.[7] However, under nonfundamentalness, the polynomial $B(L)$ is noninvertible in the past, implying that there exists no VAR(∞) representation to recover the shocks u_t from the history of y_t only. In that case, fitting a conventional VAR model to y_t produces a nonfundamental error term which is a linear combination of the past shocks (Lippi and Reichlin, 1994b; Fernández-Villaverde et al., 2007), distorting conclusions drawn from the estimated impulse responses.

Noninvertibility of the MA polynomial is potentially caused by the news shocks, when the forward-looking agents see exogenous changes not contained in the empirical model. To illustrate the issue, consider for simplicity that the observables y_t contain all state variables except k uncorrelated exogenous variables in z_t.[8] By Sims

[6] Throughout, I use terms noninvertibility and nonfundamentalness interchangeably.

[7] There may be a truncation error when the inverse of $B(L)$ is of infinite order, which can be diminished by increasing lag order p.

[8] This illustration follows Leeper et al. (2013). The nonfundamentalness issue is also discussed in Walker and Leeper (2011); Leeper et al. (2013); Forni and Gambetti (2014); Forni et al. (2014); Beaudry and Portier (2014) and Sims (2012).

(2002), y_t has a forward-looking solution[9]

$$y_t = \Theta_1 y_{t-1} + \Theta_c + \Theta_0 z_t + \Theta_y \sum_{s=1}^{\infty} \Theta_f^{s-1} \Theta_z \mathrm{E}_t z_{t+s}. \quad (3)$$

The exogenous variables are driven by unanticipated shocks when $z_t = u_t$, and the last term vanishes, leading directly to a VAR(1) representation. In contrast, when agents have foresight on the exogenous variables q periods ahead, $z_t = u_{t-q}$, and the equilibrium is determined by

$$y_t = \Theta_1 y_{t-1} + \Theta_c + \Theta_0 u_{t-q} + \Theta_y \Theta_z u_{t-q+1} + \Theta_y \Theta_f \Theta_z u_{t-q+2} + \ldots + \Theta_y \Theta_f^{q-1} \Theta_z u_t, \quad (4)$$

which corresponds to the VARMA representation (1) of the model. Strikingly, even though the more distant expected events of z_t obtain a weaker weight in the forward-looking solution (3), the most recent news on the future event z_{t+q} in (4), u_t, is discounted the heaviest by factor $\Theta_y \Theta_f^{q-1} \Theta_z$. This reverse discounting easily causes noninvertibility of the MA polynomial $B(L) = \Theta_0 L^q + \Theta_y \Theta_z L + \ldots + \Theta_y \Theta_f^{q-1} \Theta_z$, the most recent shocks having the least influence on the overall dynamics of y_t.

The noninvertible solution prevents the recovery of the news shock contained in u_t based on the past and current values of y_t only. To gain fundamentalness for (4), an obvious strategy is to include in y_t variables such as measured proxies for news or other forward-looking variables that do not suffer from the inverse discounting of the shock term. Consequently, $B(z)$ would eventually become invertible. However, it may be difficult to come up with suitable forward-looking variables and ascertain their validity. Error term ε_t and the structural shocks u_t could also be restored from the nonfundamental error term of a causal VAR by a known Blaschke matrix (see, Lippi and Reichlin 1994b). This dynamic rotation is not unique, though, and the set of Blaschke matrices can be shrunk only by means of economic theory, which may be infeasible or set restrictive assumptions on the underlying structure.

As an alternative to the above approaches, it is possible to rewrite the VARMA model (1) under noninvertibility as a noncausal autoregressive representation for y_t, representing the structural shocks in terms of past and future terms. Let l roots of $|B(z)|$ lie within the unit circle. Then y_t is noncausal as

$$\bar{c}_l \beta(L)^{-1} B^{adj}(L) \alpha(L^{-1})^{-1} A(L) y_t = u_{t-l}, \quad (5)$$

where $\bar{c}_l$ is constant, $B^{adj}(z)$ is the adjoint matrix of $B(z)$, and scalars $\alpha(z^{-1})^{-1}$ and $\beta(z)^{-1}$ are convergent power series expansions in z^{-1} and z, respectively (see Appendix A.1 for details). Through the lead polynomial $\alpha(z^{-1})^{-1}$, the time-shifted structural shocks u_{t-l} are functions of the past, current and future terms of y_t. While the history of y_t lacks information to catch the variation of u_t, movements of the lagged shocks are captured by a linear weighted sum of the past and future values of y_t.[10] Hence, both lags and leads of observables are sufficient to recover the structural shocks that are now anticipated due to the time-shifting.

[9] Matrices Θ_1, Θ_c, Θ_0, Θ_y, Θ_f, Θ_z are functions of parameters of the model of dimensions $(k \times k)$, $(k \times 1)$, $(k \times k)$, $(k \times m)$, $(m \times m)$ and $(m \times k)$, respectively. m is a dimension of the unstable block of the system, defined in Sims (2002).

[10] To establish an exact mapping between (1) and (5), an infinite number of terms has to be

2.2 Noncausal VAR

Noncausality implied by nonfundamentalness facilitates the recovery of an anticipated but exogenous error term and the derivation of impulse responses to structural shocks. However, direct inference on the noncausal representation (5) is infeasible. In this section, I present the noncausal VAR model proposed by Lanne and Saikkonen (2013) which I use to make inference on responses to the news shock.

The noncausal VAR model of Lanne and Saikkonen (2013),

$$\Pi(L)\Phi(L^{-1})y_t = \epsilon_t, \tag{6}$$

includes distinct lag and lead polynomials $\Pi(z) = I_k - \Pi_1 z - \ldots - \Pi_r z^r$ and $\Phi(z^{-1}) = I_k - \Phi_1 z^{-1} - \ldots - \Phi_s z^{-s}$. The error term ϵ_t is independent and identically distributed (iid) with zero mean and positive definite covariance matrix. Hence, the observed variables are written in a form with separate past and future-relevant parts. In the rest of the paper, model (6) is referred to as VAR(r,s). If $s = 0$, the model reduces to a standard causal VAR(r).

To guarantee stationarity and the existence of an MA representation, the following stability conditions hold:

$$\det \Pi(z) \neq 0, |z| \leq 1 \text{ and } \det \Phi(z) \neq 0, |z| \leq 1,$$

i.e., the polynomials $\Pi(z)$ and $\Phi(z)$ have well-defined inverses convergent in the powers of z. The process $\Phi(L^{-1})y_t = y_t - \Phi_1 y_{t+1} - \ldots - \Phi_s y_{t+s}$ is by the former condition stationary and has an MA representation

$$\Phi(L^{-1})y_t = \Pi(L)^{-1}\epsilon_t = M(L)\epsilon_t = \sum_{j=0}^{\infty} M_j \epsilon_{t-j}. \tag{7}$$

The process y_t can also be decomposed as

$$y_t = \Phi_1 y_{t+1} + \ldots + \Phi_s y_{t+s} + \sum_{j=0}^{\infty} M_j \epsilon_{t-j} = f_t + \sum_{j=0}^{\infty} M_j \epsilon_{t-j}, \tag{8}$$

which highlights the dependence of y_t on the future through the lead terms Φ_i, $i = 1, \ldots, s$. Conveniently, when y_t is fundamental, the lead coefficients are zeros and the decomposition (8) reduces to an MA representation of the causal VAR model (2). A nonzero future-dependence f_t in (8) indicates instead the insufficiency of the lags to recover the structural shocks.

Finally, inverting the stable polynomial $\Phi(L^{-1})$ in (7) produces a two-sided MA representation for y_t:

$$y_t = \sum_{j=-\infty}^{\infty} \Psi_j \epsilon_{t-j}. \tag{9}$$

Thus y_t depends, in general, both on the past and future error terms. Given nonzero Ψ_j when $s > 0$ with nonzero lead terms in $\Phi(L^{-1})$, ϵ_t has an effect on y_t both before

included. However, as the more distant terms of the both scalar polynomials converge to zero, a finite number of terms are sufficient to obtain the structural shocks up to a truncation error. If all roots of $|B(z)|$ within the unit circle are equal to zero, the noncausal representation is finite in its leads.

and after period t. The effect of more distant error terms disappears since Ψ_j converges to zero when $j \to \infty$ or $j \to -\infty$.

If the error term of the model is Gaussian, a noncausal VAR(r,s) is observationally equivalent to a causal VAR($r+s$) model as they cannot be statistically distinguished by the properties of the first and second moments alone.[11] Therefore, estimation of the noncausal model necessarily requires departure from Gaussianity. In what follows, the error term of the model is assumed to be multivariate t-distributed, i.e. it can be characterised by $\epsilon_t = \omega_t^{-1/2}\eta_t$, $\eta_t \sim \mathrm{N}(0,\Sigma)$ and $\lambda\omega_t$ is χ^2_λ-distributed. As a consequence,

$$\omega_t^{1/2}\Pi(L)\Phi(L^{-1})y_t = \eta_t,$$

i.e., the non-Gaussian assumption adds a stochastic volatility factor $\omega_t^{1/2}$ that affects y_t in addition to the normally distributed error term η_t. Conditional on ω_t, the error term is Gaussian, but unconditionally, higher moments of data are relaxed to be determined by the degrees-of-freedom parameter λ. For small λ, the variables exhibit leptokurtic pattern, as also observed in economic time series.[12] On the other hand, when λ increases, the distribution approaches the Gaussian distribution. Estimating λ thus allows for checking the validity of the departure from normality.

The model can be estimated by maximising the log-likelihood function, as shown by Lanne and Saikkonen (2013) (See Appendix A.2 for details), and the unique maximum likelihood estimator is consistent and asymptotically normally distributed. Finally, selecting orders r and s is based on conventional information criteria by comparing all nested VAR(r,s) models satisfying $r+s \leq p_{max}$ with $s \geq 0$ and $r > 0$, i.e., causality is not ruled out in advance. Consequently, selecting $s > 0$ suggests, by (8), directly the inadequacy of the causal VAR and its invertible MA representation to capture the underlying shocks.

2.3 Noncausality in a Stylised Model with News Shocks

Next, I illustrate in a stylised, two-variable rational expectations model, how non-invertibility and noncausality arise when news shocks drive exogenous technology. This leads to an exact mapping between the noncausal representation (5) and the noncausal VAR model (6) that reproduces the true impulse responses. Let the equilibrium conditions for an endogenous variable x_t and a technology process a_t be

$$a_t = \rho a_{t-1} + \varepsilon^a_{t-2} \quad (10)$$

$$x_t = \beta \mathrm{E}_t[x_{t+1}] + a_t + \nu_t, \quad (11)$$

where $\mathrm{E}_t[\cdot]$ denotes conditional expectation with respect to the information set containing history of $\{a_t, x_t, \varepsilon^a_t, \nu_t\}$. ε^a_t and ν_t are mutually uncorrelated structural shocks with ε^a_t being a news shock affects productivity two periods later and ν_t an unexpected nominal shock on x_t.

[11] This nonidentifiability holds for noninvertible models as well. See, e.g., Rosenblatt (2000).

[12] Various studies consider non-Gaussian distributions using macroeconomic data. Distribution of growth rates of output in OECD have been observed to be fat-tailed by Fagiolo et al. (2008). The estimation results of Cúrdia et al. (2014) and Chib and Ramamurthy (2014) suggest a t-distribution for innovations in DSGE models in the low-frequency data. See also Ascari et al. (2015) for fat-tailed distributions in macroeconomic time series.

Assuming $\beta < 1$, x_t has a forward-looking solution

$$x_t = \sum_{j=0}^{\infty} \beta^j \mathrm{E}_t a_{t+j} + \nu_t \tag{12}$$

and together with (10),

$$x_t = \theta a_t + \theta\beta^2 \varepsilon_t^a + \theta\beta\varepsilon_{t-1}^a, \tag{13}$$

where $\theta = (1 - \beta\rho)^{-1}$. Thus, a_t and x_t have a VARMA representation

$$\begin{bmatrix} 1-\rho L & 0 \\ -\theta\rho L & 1 \end{bmatrix} \begin{bmatrix} a_t \\ x_t \end{bmatrix} = \begin{bmatrix} L^2 & 0 \\ \theta\beta^2 + \theta\beta L + \theta L^2 & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_t^a \\ \nu_t \end{bmatrix} = B(L)u_t, \tag{14}$$

which is noninvertible in the past since $|B(z)| = z^2 = 0$ for $z = 0$. Hence, the history of $y_t = (a_t, x_t)'$ is insufficient to recover the structural shocks $u_t = (\varepsilon_t^a, \nu_t)'$, and impulse responses cannot be derived from a causal VAR representation for y_t.

However, as in (5), y_t can be written as noncausal – in terms of observables and an anticipated error term. By (10), ε_{t-1}^a and ε_t^a reveal productivity perfectly two periods forward, and future terms of productivity can be directly substituted to x_t in (13) using $\varepsilon_t^a = -a_{t+2} + \rho a_{t+1}$ such that

$$x_t = \rho a_{t-1} + \beta a_{t+1} + \theta\beta^2 a_{t+2} + \varepsilon_{t-2}^a + \nu_t.$$

Hence, x_t is noncausal with finite leads. Together with (10), the noncausal representation for y_t is equivalently expressed as

$$\left(I_2 - \begin{bmatrix} \rho & 0 \\ \rho & 0 \end{bmatrix} L\right)\left(I_2 - \begin{bmatrix} 0 & 0 \\ \beta & 0 \end{bmatrix} L^{-1} + \begin{bmatrix} 0 & 0 \\ \theta\beta^2 & 0 \end{bmatrix} L^{-2}\right)\begin{bmatrix} a_t \\ x_t \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} \varepsilon_{t-2}^a \\ \nu_t \end{bmatrix}, \tag{15}$$

which is the noncausal VAR(1,2) model (6) with the right-hand side error term containing the anticipated shock ε_{t-2}^a.

Since both matrix polynomials on the left-hand side of (15) are stable, y_t has the two-sided MA representation (9). I show in Appendix A.3 that this representation collects the impulse response coefficients to the structural shocks ε_{t-2}^a and ν_t, which analytically coincide with their theoretical counterparts. In Figure 1, I plot the theoretical and empirical impulse responses of a_t and x_t in the upper and lower panels, respectively, when $\beta = 0.9$ and $\rho = 0.9$. Comparing the panels reveals that the noncausal VAR reproduces the theoretical impulse responses with respect to the anticipated shock ε_{t-2}^a. This shock realises in a_t at lag 0 but affects x_t at the negative lags due to terms ε_{t-1}^a and ε_t^a in (13). Importantly, the theoretical and empirical responses differ only in the timing of the shock ε_t^a, and the inclusion of negative lags allows to track the full propagation of the shock. Despite this time-shifting which occurs due to the absence of a causal representation, the impulse responses can be interpreted from the two-sided MA representation in a conventional manner, and they provide conclusions consistent with the underlying economic process.

This example established an exact mapping from the theoretical model under noninvertibility to the noncausal VAR (6). In Appendix A.4, I show that this mapping holds for a more general technology process as well. However, due to

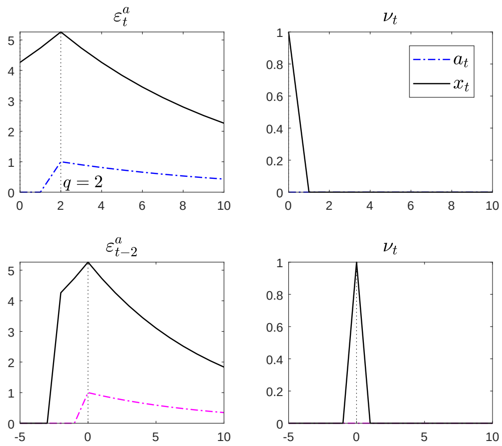


Figure 1: Theoretical and empirical impulse response functions of the example with $\beta = 0.9$ and $\rho = 0.9$

The upper row shows the theoretical impulse responses of a_t and x_t to shocks ε_t^a and ν_t. The lower row plots the impulse responses obtained from the MA representation of the noncausal VAR to the shocks of the error term, ε_{t-2}^a and ν_t.

the multiplicative structure of the noncausal VAR, the equivalence does not hold in general for all noninvertible representations. Nonetheless, as confirmed in the Monte Carlo simulations, the empirical model approximates the theoretical representation reasonably well when the orders r and s are optimally chosen, for instance, by an information criterion.

2.4 Identification of News Shocks

I turn now to the structural analysis and identification of a news shock with the noncausal VAR, which is the main contribution of the paper.[13] The causal SVAR methodology identifies the news shock as a shock with an immediate effect on forward-looking variables, investment and stock prices, followed by a delayed impact on total factor productivity (TFP). However, the identification strategy fails to recover the true shocks under nonfundamentalness. With the identification I propose, the news shocks can be obtained under nonfundamentalness from the reduced-form errors of a noncausal VAR model by extracting a shock that explains the most of the relation between productivity and the future-dependence of a forward-looking variable.

The structural analysis with the noncausal VAR is based on the anticipated, multivariate t-distributed error term, decomposed as

$$\epsilon_t = \omega_t^{-1/2} \eta_t = \bar{B} \bar{u}_t, \tag{16}$$

where the t-distributed structural shock vector, $\bar{u}_t = \omega_t^{-1/2} u_t^* = [\bar{u}_{1,t} \cdots \bar{u}_{k,t}]' \sim t_\lambda(I_k)$, is a product of two latent factors, a k-dimensional vector of Gaussian shocks

[13]Prior to this study, the noncausal VAR model has not been used for identifying structural shocks. Davis and Song (2012) also conduct impulse response analysis based on the two-sided MA representation. However, with the recursive identification through Cholesky decomposition they use, it is difficult to draw a structural interpretation.

$u_t^* \sim \mathrm{N}(0, I_k)$ and the volatility term $\omega_t^{-1/2}$. Hence, compared to the standard Gaussian setting, the distributional assumption adds the term $\omega_t^{-1/2}$ to the normally distributed structural shocks in u_t^* to control for overall volatility. As a whole, the structural shocks are mapped by a rotation matrix $\bar{B}$ to the reduced-form error term ϵ_t. From the two-sided MA representation (5), a reaction to a unit shock in $\bar{u}_{i,t}$ is

$$\frac{\partial y_{t+j}}{\partial \bar{u}_{i,t}} = \Psi_j \gamma_i, \ j = \ldots, -1, 0, 1, \ldots, \tag{17}$$

where γ_i denotes the ith column of matrix $\bar{B}$.

The identification of structural shocks follows the standard strategy by finding a nonsingular matrix $\bar{B}$ such that $\mathrm{E}[\epsilon_t \epsilon_t'] = \mathrm{E}[\bar{B}\bar{u}_t\bar{u}_t'\bar{B}']$, or

$$\Sigma = \bar{B}\bar{B}'. \tag{18}$$

Let further W be an orthogonal $(k \times k)$ matrix with w_i its ith column, and $\tilde{A}$ a Cholesky factor satisfying $\Sigma = \widetilde{A}\widetilde{A}'$. Assume now the news shock $\bar{u}_{1,t}$ is ordered first in $\bar{u}_t$. By rewriting the rotation matrix as $\bar{B} = \widetilde{A}W$, the impulse responses to the news shock are obtained by finding w_1 such that $\gamma_1 = \widetilde{A}w_1$.

Finding w_1 relies in turn on the relationship between movements in TFP and future-dependence of a forward-looking variable. The forward-looking variable, most notably investment or stock prices, is expected to respond to future increases in productivity. From (9), the MA representation for TFP, ordered as the first variable in y_t, is

$$y_{1,t} = \sum_{j=-\infty}^{\infty} e_1' \Psi_j \tilde{A} W \bar{u}_{t-j} \tag{19}$$

with the vector $e_i = [0 \cdots 1 \cdots 0]'$ having one in its ith element. On the other hand, by (8), y_t is a combination of the future, noncausal dependence and a weighted sum of past terms. The future dependence of the forward-looking variable, placed as the second in y_t, is

$$\begin{aligned} f_{2,t} &= e_2' \left(\Phi_1 y_{t+1} + \ldots + \Phi_s y_{t+s} \right) \\ &= y_{2,t} - \sum_{j=0}^{\infty} e_2' M_j \widetilde{A} W \bar{u}_{t-j} \\ &= \sum_{j=-\infty}^{\infty} e_2' \Psi_j^* \widetilde{A} W \bar{u}_{t-j}, \ \Psi_j^* = \begin{cases} \Psi_j - M_j, & j \geq 0, \\ \Psi_j, & j < 0 \end{cases} \end{aligned} \tag{20}$$

Notably, $f_{2,t}$ is the anticipatory component of $y_{2,t}$, which emerges due to nonfundamentalness. As $W\bar{u}_{t-j} = w_1\bar{u}_{1,t} + \ldots w_k\bar{u}_{k,t}$, $y_{1,t}$ and $f_{2,t}$ in (19) and (20) can also be written as sums of the contributions of the k structural shocks:

$$y_{1,t} = y_{1,t}^1 + \ldots + y_{1,t}^k, \ y_{1,t}^l = \sum_{j=-\infty}^{\infty} e_1' \Psi_j \widetilde{A} w_l \bar{u}_{l,t-j},$$

$$f_{2,t} = f_{2,t}^1 + \ldots + f_{2,t}^k, \ f_{2,t}^l = \sum_{j=-\infty}^{\infty} e_2' \Psi_j^* \widetilde{A} w_l \bar{u}_{l,t-j}.$$

Consequently, the covariance between these two measures is a sum of covariances contributed by the k distinct structural shocks,

$$\mathrm{Cov}(f_{2,t}, y_{1,t}) = \mathrm{Cov}(f_{2,t}^1, y_{1,t}^1) + \ldots + \mathrm{Cov}(f_{2,t}^k, y_{1,t}^k), \tag{21}$$

since $\mathrm{Cov}(f_{2,t}^i, y_{1,t}^j) = 0$ for $i \neq j$. The covariance due to the news shock $\bar{u}_{1,t}$ over horizon $[-H_1, H_2]$ reads as

$$\mathrm{Cov}(f_{2,t}^1, y_{1,t}^1) = \mu_{\omega^{-1}} \sum_{j=-H_1}^{H_2} e_2' \Psi_j^* \widetilde{A} w_1 w_1' \widetilde{A}' \Psi_j' e_1, \tag{22}$$

where $\mu_{\omega^{-1}} \equiv \mathrm{E}[\omega_t^{-1}] = \lambda E\left[\frac{1}{\chi^2_\lambda}\right] = \frac{\lambda}{\lambda-2}$.

To identify the news shock $\bar{u}_{1,t}$, I find the single component among the orthogonal structural shocks weighting the most on the covariance (21) in absolute value over the given horizon $[-H_1, H_2]$, i.e., the news shock explains the most of the correlation between movements in TFP and the future dependence of the forward-looking variable. In other words, w_1 is obtained from the optimisation problem

$$\max_{w_1} \left| \mathrm{Cov}(f_{2,t}^1, y_{1,t}^1) \right| \tag{23}$$

subject to

$$w_1' w_1 = 1.$$

w_1 then maximises the absolute value of covariance (22) subject to W being orthogonal. By construction, the covariance (22) and hence the objective function are finite, and Appendix A.5 shows the maximisation problem has a unique solution. This medium-run identification approach is similar to Uhlig (2004) and Francis et al. (2014), also used in studying news shocks by Barsky and Sims (2011) and Kurmann and Otrok (2013), but concerns covariance between the two above measures instead of forecast error variance of one variable.

In sum, the outlined identification strategy relies on the fact that the news shock is driving the anticipatory part $f_{2,t}$ of the forward-looking variable $y_{2,t}$ that cannot be captured by the backward-looking MA representation in (7). Therefore, once noncausality is present, the error term of the empirical model contains structural shocks that are anticipated by the economic agents, and with residuals exhibiting non-Gaussianity those shocks are identifiable from the future dependence $f_{2,t}$.[14]

3 Monte Carlo Simulation

In this section, I demonstrate with Monte Carlo simulations how causal and noncausal VAR models are able to identify news shocks and match the true impulse responses of a small-scale macroeconomic model.

[14]The news shocks identified by the strategy may also have impact effects on productivity and in fact nest both smoothly diffusing technology shocks as in Lippi and Reichlin (1994a) or Walker and Leeper (2011) as well as shocks with delayed effects on TFP as mostly considered in the VAR literature starting from Beaudry and Portier (2006).

3.1 New Keynesian Model with News Shocks

Consider a textbook New Keynesian (NK) model (See Galí, 2009). The equilibrium is characterised by the dynamic IS equation

$$\tilde{y}_t = -\frac{1}{\sigma}\left(i_t - \mathrm{E}_t\pi_{t+1} - r_t^n\right) + \mathrm{E}_t\tilde{y}_{t+1},$$

where $\tilde{y}_t$ is the output gap relative to the flexible price outcome, π_t is inflation, i_t is the interest rate and $r_t^n = \rho + \sigma\psi_{ya}^n \mathrm{E}_t\Delta a_{t+1}$ is the natural interest rate. The NK Phillips curve determines inflation:

$$\pi_t = \beta \mathrm{E}_t[\pi_{t+1}] + \kappa\tilde{y}_t,$$

and the central bank sets the nominal interest rate according to the Taylor rule

$$i_t = \rho(1-\rho_m) + \rho_m i_{t-1} + \phi_\pi\pi_t + \phi_y\tilde{y}_t + \varepsilon_t^m.$$

The interest rate is thus a state variable of the model due to the autoregressive term. The economy is driven by three exogenous shocks, the anticipated and unanticipated shocks on technology, ε_t^a and ε_t^u, respectively, and the monetary policy shock ε_t^m. Technology follows

$$a_t = \rho_a a_{t-1} + \varepsilon_{t-q}^a + \sigma_u\varepsilon_t^u \tag{24}$$

so through the news shock ε_t^a, agents are able to predict technology q periods forward. The surprise technology, or noise shock ε_t^u prevents the agents from perfectly observing future technology and σ_u determines its relative importance.

The inclusion of the news shock alters the dynamics of the sticky-price model as follows. In response to a positive news shock, firms anticipate the future improvement in productivity and have an incentive to decrease prices beforehand. As a result, the economy initially experiences a boom resulting in a positive output gap, which accelerates the inflation. To stabilise inflation and output gap, the Taylor rule responds by increasing the interest rate. Consistent with sticky-price models, the news shock at the time of its materialisation eventually turns the output gap and inflation negative, similar to the effects of a surprise technology shock in a basic NK model.

3.2 Simulation Design

I solve the model numerically using standard calibration (see Appendix B.1). In addition, I vary the significance of news shocks and the anticipation horizon in the model by considering values $\sigma_u \in \{0.5, 1\}$ and $q \in \{3, 16\}$. These values are motivated as follows. First, lowering the standard deviation of the noise shock from $\sigma_u = 1$ to $\sigma_u = 0.5$ increases the significance of news shocks in driving technology. The nonfundamentalness-inducing shock comprises then a larger share of exogenous variation, which may deepen the noninvertibility problem.[15] Second, compared to the benchmark anticipation horizon $q = 3$ also used by Sims (2012), the longer anticipation horizon $q = 16$ is more consistent with the empirical estimates starting from Beaudry and Portier (2006) who measure news shocks moving the economy

[15] However, see Sims (2012) for a contrasting view.

several years ahead. However, lengthening the horizon relates news to more distant changes in technology, which worsens the ability of current observables to recover the underlying shocks and to replicate the underlying impulse responses from the causal model.

I simulate from the model 1,000 samples of time series of length $T = 250$ with structural shocks $u_t = (\varepsilon_t^a, \varepsilon_t^m, \varepsilon_t^u)'$ drawn from multivariate t-distribution $t_{\lambda,}(\mu_{\omega^{-1}}^{-1/2} I_3)$ with $\lambda = 4$. The structural shocks are therefore uncorrelated and Gaussian conditional on the latent volatility factor $\omega_t^{1/2}$.[16] Due to the low degrees-of-freedom parameter λ, the distribution of shocks has more weight on tails than in the purely Gaussian case, extreme shocks being more frequent.[17] For each simulated sample, I estimate both causal and noncausal VAR models with two sets of observables, $y_t^1 = (a_t, \tilde{y}_t, \pi_t)$ and $y_t^2 = (a_t, i_t, \pi_t)$ as the number of included state variables influences the nonfundamentalness issue. The first set includes two forward-looking variables, inflation and output gap, whereas the latter includes two state variables, technology and the interest rate.

3.3 Nonfundamentalness and Model Selection

In Table 1, I summarise results from the simulated models. I report the maximum absolute eigenvalue of the invertibility condition of Fernández-Villaverde et al. (2007), computed from the theoretical state space representation for the observables. When this eigenvalue is greater than one in modulus, the observables suffer from nonfundamentalness.[18] Examining the nonfundamentalness issue further, I estimate all VAR(r,s) models satisfying $r + s \leq p_{max} = 6$ with $s \geq 0$ and $r > 0$ and compare them by minimising the Akaike Information Criterion (AIC) in each simulated sample.[19] The shares of selected specifications are collected in Table 1.

For the anticipation horizon $q = 3$, Table 1 shows that including two forward-looking variables, output gap and inflation, in observables is insufficient to attain invertibility because the maximum absolute eigenvalue is greater than one. The absence of an invertible MA representation is consequently reflected by primarily selecting models with lead terms. On the other hand, the invertibility condition is satisfied for observables y_t^2. In particular, by replacing one forward-looking variable by a state variable, the interest rate, in the observable vector, the causal VAR remains valid with respect to structural shocks. In line with this fact, fundamental observables y_t^2 induce the dominance of causal specifications over noncausal variants in the model selection. When lengthening the horizon to $q = 16$, the interest rate is, however, unable to recover the state space, which now includes all 16 lags of

[16] Term $\mu_{\omega^{-1}}^{-1/2}$ induces a unit variance for $(\varepsilon_t^a, \varepsilon_t^m, \varepsilon_t^u)'$.

[17] Since structural shocks are uncorrelated, the additional assumption compared to the Gaussian case has no consequences to the equilibrium conditions of the DSGE model up to the first order approximation.

[18] That is, under a state space representation with x_t, an $(n \times 1)$ vector of state variables,

$$
\begin{aligned}
x_t &= Ax_{t-1} + Bu_t \\
y_t &= Cx_{t-1} + Du_t,
\end{aligned}
$$

where A, B, C and D are $(n \times k)$, $(n \times k)$, $(k \times n)$ and $(k \times k)$ matrices, respectively, y_t is fundamental, if matrix $F = A - BD^{-1}C$ has all eigenvalues inside the unit circle.

[19] Increasing the maximum number of lags does not change the results

		Model							
q		3		16		3		16	
σ_u		1		1		0.5		0.5	
Observables		y_t^1	y_t^2	y_t^1	y_t^2	y_t^1	y_t^2	y_t^1	y_t^2
Maximum Eigenvalue		14.13	0.71	1.59	1.37	14.13	0.71	1.59	1.37
Orders		Specification Selected by AIC (%)							
$r = 1$	$s = 0$	1.20	0.0	0.8	0.0	0.0	0.0	0.0	0.0
$r = 2$	$s = 0$	0.60	0.0	2.2	0.1	0.6	0.0	0.0	9.7
$r = 3$	$s = 0$	2.00	0.0	1.7	0.1	1.2	0.0	0.3	0.0
$r = 4$	$s = 0$	0.20	9.8	1.5	0.0	0.2	1.6	0.9	0.0
$r = 5$	$s = 0$	0.00	74.0	0.9	0.0	0.0	73.0	1.4	0.0
$r = 6$	$s = 0$	0.10	6.6	1.3	0.0	0.1	6.5	3.2	0.0
$r = 1$	$s > 0$	30.80	0.3	35.6	78.5	37.6	0.0	23.0	89.9
$r = 2$	$s > 0$	15.70	0.4	21.6	19.7	8.9	0.9	14.5	0.4
$r = 3$	$s > 0$	41.80	0.1	23.6	1.1	40.4	1.7	23.0	0.00
$r = 4$	$s > 0$	7.20	8.8	7.9	0.2	10.5	2.4	14.5	0.0
$r = 5$	$s > 0$	0.40	0.0	2.9	0.3	0.5	13.9	24.2	0.0
Causal (%)		4.10	90.4	8.4	0.2	2.1	81.1	5.8	0.2
Noncausal (%)		95.90	9.6	91.6	99.8	97.9	18.9	94.2	99.8
(r^*_{AIC}, s^*_{AIC})		(1,2)	(5,0)	(1,3)	(1,2)	(1,2)	(5,0)	(4,1)	(1,2)
p^*_{AIC}		3	5	5	3	3	5	6	3

Table 1: Estimation Results for Simulated Models: Causal and Noncausal Models Selected by AIC

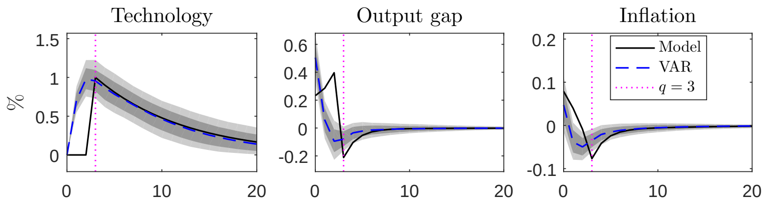
Percentages refer to the frequency of a VAR(r,s) model being selected from minimising the Akaike information criterion. Models are estimated by maximum likelihood with multivariate t-distributed residuals for 1,000 simulated Monte Carlo samples with $T = 250$. Noncausal specification with $r = r^*$,$s > 0$ nests all leads satisfying $r^* + s^* \leq 6$, (r^*_{AIC}, s^*_{AIC}) is the most selected model with $s \geq 0$, p^*_{AIC} the most selected causal model ($s = 0$).

ε_t^a, and the noninvertibility problem emerges irrespective of the used variables. In fact, avoiding nonfundamentalness can only be achieved, if feasible, by including a sufficiently precise proxy for the news shock ε_t^a, which would revert an invertible state space representation. As a consequence of nonfundamentalness, over 90 per cent of the selected models are noncausal.

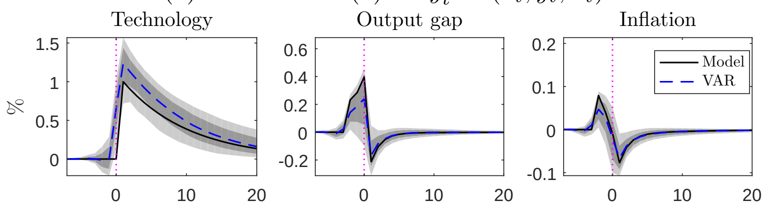
Although the above model selection performs well in checking the sufficiency of information, it has slight bias towards including additional lead terms under fundamentalness. Hence, the procedure cannot be regarded as a formal test for nonfundamentalness but it rather provides first-hand information on the fit and validity of a causal model compared to including lead terms. It is also worth noting that the chosen number of leads and lags, r and s, in Table 2 are evenly distributed under nonfundamentalness, suggesting that no particular noncausal VAR(r,s) is a direct empirical alternative to represent the structural shocks. This is in part due to numerical reasons since the multimodality of the nonlinear likelihood function is difficult to control for in the simulations. In practical implementation, these issues can be tackled by more scrutinised numerical algorithms and model diagnostics.

3.4 Impulse Responses

In the simulated samples, I estimate the impulse responses to the news shock from the causal VAR(p) and noncausal VAR(r,s) models using the orders of the most



(a) Causal VAR(3) on $y_t^1 = (a_t, \tilde{y}_t, \pi_t)$



(b) Noncausal VAR(1,2) on $y_t^1 = (a_t, \tilde{y}_t, \pi_t)$

Figure 2: Impulse Responses to a News Shock from a Causal and Noncausal VAR in the NK model with $q = 3$ and $\sigma_u = 1$

The dashed lines are the median estimated impulse responses from the Monte Carlo samples. Light and dark grey shaded areas border the middle 90 % and 68 %, respectively, of the distribution for estimated impulse responses. Black solid lines are the theoretical impulse responses, aligned in panel (b) with the estimated noncausal impulse responses according to the maximum impact on technology.

selected specification. If the selected orders correspond to a causal model, the news shock is identified by the Barsky-Sims strategy (Barsky and Sims, 2011), as a shock having no impact effect on technology but maximising its forecast error variance (FEVD).[20] The scheme is consistent with the theoretical model as long as observables are fundamental for the underlying process. If a noncausal VAR is selected, I identify the news shock as described in subsection 2.4: the news shock is a shock explaining the most of the comovement between technology $y_{1,t} = a_t$ and $f_{2,t}$, the future dependence of inflation with observables y_t^1 and interest rate when using y_t^2.

Figure 2 plots the estimated impulse responses of y_t^1 to a news shock from VAR(3) and VAR(1,2) models for $q = 3$ and $\sigma_u = 1$. As an immediate consequence of nonfundamentalness, the dashed median responses from the causal VAR in panel (a) fail to coincide with their theoretical counterparts. The distortion follows a pattern where the theoretical responses are lagged by two periods, originating from the nonfundamental error term that strongly weighs the past shocks.[21] As the causal VAR cannot then retrieve the most recent shock, the model is incapable to reproduce the initial reactions. This drawback disappears only if the observables are fundamental, which can be achieved by including the interest rate. The causal VAR with y_t^2 and the Barsky-Sims identification are then able to reproduce the true impulse responses, shown in Appendix B.2. Hence, although both observables could well be used for empirical analysis of news shocks, their capability to derive the true responses heavily depends on invertibility.

[20] In particular, the first column of matrix B_0 in (2), γ_1, is found by maximising $\sum_{j=0}^{20} \Psi_j A_0 \gamma_1 \gamma_1' A_0' \Psi_j'$ subject to $\gamma_1'\gamma_1{=}1$ and $\gamma_{1,1} = 0$, A_0 being a lower-triangular satisfying $\mathrm{E}[\epsilon_t \epsilon_t'] = A_0 A_0'$.

[21] See Lippi and Reichlin (1994b); Leeper et al. (2013) for deeper evaluation.

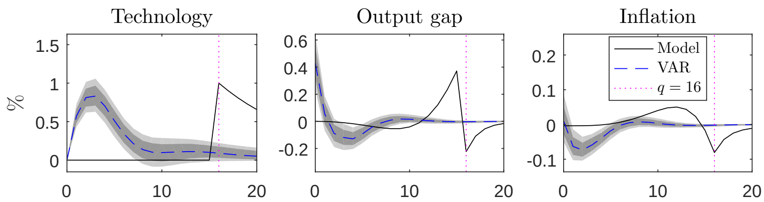
Instead, the impulse responses of y_t^1 can be recovered by the noncausal VAR selected by AIC, shown in panel (b) of Figure 2, where the timing of the theoretical responses are changed such that the peak response of technology is aligned with those from noncausal models. Notice that no information about this shifting of the error term was needed in the estimation. Evidently, the noncausal VAR is able to match the theoretical impulse responses and measures the early reactions to a news shock at negative lags, as a result of the two-sided MA representation including the leads of time-shifted shocks. In Appendix B.2, I additionally show the results when the relative weight of the news shock is greater, $\sigma_u = 0.5$. In that case, the impulse responses can be more accurately recovered from the noncausal VAR while the distortion in the causal VAR prevails.

Once the anticipation horizon is $q = 16$, fundamentalness becomes unattainable even with the inclusion of the interest rate. This information deficiency has adverse effects on the performance of causal VAR models as seen in panels (a) and (b) of Figure 3. For both y_t^1 and y_t^2, the estimated impulse responses lead the theoretical counterparts by several periods and are therefore unable to reveal the initial, smooth responses to the news shocks. Concluding from the causal VAR, a news shock would induce an immediate, strong response of output gap, interest rate and inflation. The risk of nonfundamentalness thus concerns missing the responses at the initial path of anticipation.[22] In contrast, the noncausal VAR models are robust with respect to the nonfundamentalness problem: their impulse responses shown in panels (c) and (d) are consistent with the theoretical model after shifting the timing of news shocks to the left, which does not alter interpretations about the news shock. Moreover, as reported in Appendix B.2, the performance of the noncausal VAR improves compared to the causal model when the relative weight of the noise shock is lowered to $\sigma_u = 0.5$

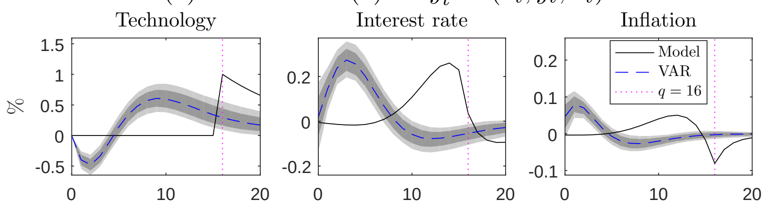
Finally, I study the sensitivity of the non-Gaussian distributional assumption. Originating from the fact that noncausal or noninvertible models cannot be statistically identified by the properties up to the second moments, the non-Gaussianity is the sole way out under nonfundamentalness. In this purpose, the multivariate t-distribution functions as a simple departure from Gaussianity, although the error term may then collect both anticipated and unanticipated shocks that share a common volatility term. Examining whether the assumed multivariate t-distribution is critical for the results, I draw Gaussian structural shocks but alternatively generate heavier tails by amplifying their size as described in Appendix B.2. Estimating now a noncausal VAR with multivariate t-distributed errors produces, first, a median estimate of 4.3 for degrees of freedom parameter λ and, second, recovers the underlying impulse responses.[23] Hence, in practice, the multivariate t-distribution is able to control for leptokurtosis originating from an alternative process.

[22] Distortion in causal VAR models is also independent of the chosen lag length.

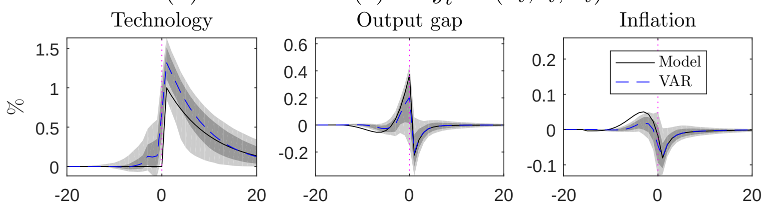
[23] On the other hand, when I simulate the theoretical model with Gaussian shocks, I estimate the degrees-of-freedom parameter to be considerably larger, above 30. Higher estimates of λ thus suggest failure of the distributional assumption and weak identification of the noncausal VAR.



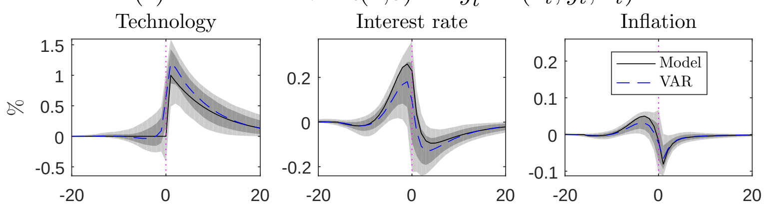
(a) Causal VAR(5) on $y_t^1 = (a_t, \tilde{y}_t, \pi_t)$



(b) Causal VAR(3) on $y_t^2 = (a_t, i_t, \pi_t)$



(c) Noncausal VAR(1,3) on $y_t^1 = (a_t, \tilde{y}_t, \pi_t)$



(d) Noncausal VAR(1,2) on $y_t^2 = (a_t, i_t, \pi_t)$

Figure 3: Impulse Responses to a News Shock from a Causal and Noncausal VAR in the NK model with $q = 16$ and $\sigma_u = 1$

The dashed lines are the median estimated impulse responses from the Monte Carlo samples. Light and dark grey shaded areas border the middle 90 % and 68 %, respectively, of the distribution for estimated impulse responses. Black solid lines are the theoretical impulse responses, aligned in panel (b) with the estimated noncausal impulse responses according to the maximum impact on technology.

4 News Shocks and U.S. Business Cycles

News shocks generate fluctuations in response to productivity changes materialising in the future. Beaudry and Portier (2006) concluded that they account for half of the output variation and increase stock prices, investment, consumption and hours on impact. Nonetheless, the estimated impulse responses and the significance of the news shock starkly depend on the choice of variables and identification strategy. Using the medium-run identification, Barsky and Sims (2011) find news shocks less important for the short-run fluctuations. Forni et al. (2014) reject the hypothesis of fundamentalness in their test for the small-scale VAR models. In a factor-augmented VAR they use, the news shock signifies less at the business cycle frequencies. According to Beaudry et al. (2015), nonfundamentalness does not alter the conclusions about the news but the identification scheme is a vital issue in reproducing the results of Beaudry and Portier (2006). However, the simulation results from the previous section suggest that the nonfundamentalness issue may be a consequential factor for the results drawn from the causal VAR.

Here, I revisit the evidence on news shocks in the U.S. economy with the noncausal VAR. The approach tackles the above issues with inference robust against the potentially insufficient information set.

4.1 Data

I use quarterly U.S. macroeconomic time series: total factor productivity (TFP_t) is the capacity-utilisation adjusted measure constructed by Fernald (2012). Consumption (C_t) is defined as a sum of consumption of nondurables and services, hours worked (H_t) are from the nonfarm sector, and investment (I_t) is the sum of fixed private investment and consumption of durables. Output (Y_t) is measured as the real gross domestic product (GDP). The variables are seasonally adjusted, expressed in per-capita terms by the civilian noninstitutional population and logarithmised. Annualised inflation (π_t) is computed from the consumer price index. Following the literature, stock price series (SP_t), measured as the log of the real S&P 500 index, is transformed to per-capita terms. The above data span quarters from 1948:1 to 2015:4, where the first quarter is lost because of differencing.[24]

I also consider two measures of the interest rate. First, the effective federal funds rate (FFR_t), available from quarter 1954:3 onwards, is aggregated to the quarterly frequency. As a long-term rate, I use the nominal 10-year interest rate (r_t^{10}) on government bonds.[25] Subsequently, I construct the term spread as difference between the 10-year yield and the policy rate.

4.2 Estimation and Identification

I estimate three different specifications summarised in Table 2 which all include TFP, consumption, hours and either investment or stock prices. The nonstationary

[24] Data are downloaded from the FRED database. The TFP series of Fernald (2012) is taken from the Federal Reserve Bank of San Francisco database, and the stock price data from Robert Shiller's webpage.

[25] I use as the 10-year rate the U.S. long-term rate constructed by OECD, based on the prices of government bonds traded on financial markets. The series, available from quarter 1953:2 onwards, is downloaded from the OECD database.

Specification	Variables	Periods
(1)	$(\Delta \log TFP_t, \Delta \log I_t, \Delta \log C_t, \log H_t, \Delta \log Y_t, \pi_t)$	1948:2–2015:4
(2)	$(\Delta \log TFP_t, \Delta \log SP_t, \Delta \log c_t, \log H_t)$	1948:2–2015:4
(3)	$(\Delta \log TFP_t, \Delta \log I_t, \Delta \log C_t, \pi_t, \log H_t, FFR_t, r_t^{10} - FFR_t)$	1954:3–2015:4

Specification	(r^*,s^*)	p_t^*	p_G^*	$\hat{\lambda}(r^*, s^*)$	$\hat{\lambda}(p_t^*)$	$\hat{\lambda}(p_G^*)$
(1)	(5,3)	3	4	2.0001	4.8372	4.7656
(2)	(3,5)	4	4	2.2020	4.6342	4.5350
(3)	(5,3)	3	6	2.00001	4.2340	5.0012

Table 2: Model specifications and Selection

(r^*, s^*) is the selected orders by minimising the Akaike Information Criterion. p_t^* and p_G^* are the lag lengths chosen by AIC among causal VAR models with multivariate t-distributed and Gaussian errors, respectively. $\lambda(\cdot)$ is the ML estimate of the degrees-of-freedom parameter for the VAR models with multivariate t-errors.

variables are differenced as the estimation of the noncausal VAR requires stability of the lag and lead polynomials. The baseline specification (1) is analogous that used in a number of studies examining news shocks. Specification (2) contains the stock prices, which according to Beaudry and Portier (2014) incorporate information relevant for the identification of the news shock. Finally, in specification (3), I analyse the reactions of monetary policy and interest rates by using a similar set of variables as Kurmann and Otrok (2013).

To select the orders r and s, I minimise the Akaike information criterion (AIC) among the ML estimates for VAR(r,s) models satisfying $r + s \leq 8$, $r \geq 0$ and $s > 0$ with a multivariate t-distributed error term.[26] In Table 2, I report the optimal orders (r^*, s^*), p_t^* and p_G^* among the noncausal VAR(r,s), non-Gaussian causal VAR and Gaussian causal VAR models, respectively. Accordingly, data support the assumption of the multivariate t-distributed error term by two aspects. First, in terms of likelihood and AIC, all non-Gaussian VAR(p_t^*) and VAR(p_G^*) models are superior to the corresponding Gaussian models. Second, the degrees-of-freedom estimate $\hat{\lambda}$ of the causal VAR model is low in all specifications. Hence, the non-normality of the error term leads to the improved empirical fit of the causal VAR compared to the Gaussian assumption. As non-Gaussianity is present, I proceed with the selection of lags and leads, and all best-fitting models are noncausal. In particular, the likelihood of a noncausal VAR(r,s) with $s > 0$ is considerably higher than that of a causal VAR($r + s$) model, indicating the insufficiency of the lags only to represent the structural shocks.

With these variable specifications, I continue the analysis using the VAR(r,s) models selected by AIC. Exploiting the noncausal structure, the identification of news shocks follows the strategy outlined in Section 2.4. In specifications (1) and (3), I use investment as target variable $y_{2,t} = \log I_t$ and, in specification (2), stock prices as $y_{2,t} = \log SP_t$. The news shock is then a shock explaining the most of the

[26] The estimation of the noncausal VAR models involves nonlinear optimisation, and the likelihood function is found by Lanne and Luoto (2016) to be multimodal in short time series. To find feasible initial estimates for the global maximum of the likelihood function, I use the posterior simulation algorithm of Lanne and Luoto (2016).

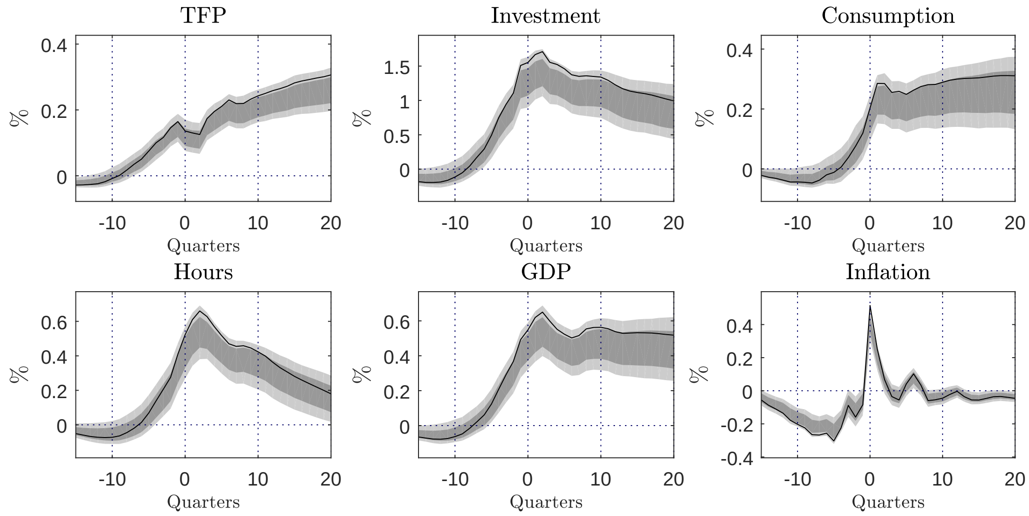


Figure 4: Impulse responses to a news shock from the noncausal VAR(5,3) model
The news shock is identified with investment as a target variable. The solid lines are the ML estimates of the model. All responses are shown in levels. Light and dark grey shaded areas are bordered by the middle 90 % and 68 % confidence regions of the bootstrap estimates. Confidence intervals are obtained from wild bootstrap with 1,000 replications.

covariation between TFP and future dependence of this target variable,

$$y_{1,t}^1 = \log TFP_t = \sum_{j=-H_1}^{H_2} e_1' \bar{\Psi}_j \tilde{A} w_1 \bar{u}_{1,t-j},$$

$$f_{2,t}^1 = \sum_{j=-H_1}^{H_2} e_2' \bar{\Psi}_j^* \tilde{A} w_1 \bar{u}_{1,t-j}$$

where $\bar{\Psi}_j = \sum_{k=-\infty}^{j} \Psi_k$ and $\bar{\Psi}_j^* = \sum_{k=-\infty}^{j} \Psi_k^*$ are cumulative sums that transform differences to levels. The news shock maximises the absolute value of covariance between these measures over the horizon $[H_1, H_2] = [-20, 20]$.[27]

4.3 Results from the Baseline Model

Figure 4 plots the impulse responses to a one-standard deviation news shock from the selected noncausal VAR(5,3) model for the baseline specification (1). All variables are shown in levels and as cumulative impulse responses, and the 68 and 90 percent confidence bands are derived by the wild bootstrap. The news shock triggers a steady increase of total factor productivity, with investment, consumption, hours and output having positive and persistent responses. On the other hand, inflation turns negative on the negative horizon before the realisation of the shock in TFP.

The news shock diffuses to the economy as follows. First, at the negative lags of the horizon, TFP starts to grow, followed by accelerating movements in the real variables. The earliest reactions to the news shock can be seen in inflation which turns to 0.3 percentage points after a steady decline of 15 quarters. At date 0, after the growth of the past 10 quarters, a half of the total impact of the news

[27]The results are insensitive to the choice of this horizon.

shock on TFP has materialised, investment peaks at 2 per cent above its initial level whereas hours, GDP and and consumption have increased by just over half per cent. Simultaneously, inflation becomes temporarily positive but converges to zero after a year. During the final phase, TFP continues its steady diffusion, investment, consumption and production slightly decrease and employment starts to return to its initial level.

Overall, it is difficult to find the news shock to be a signal from the future productivity with a long-delayed impact as observed by Beaudry and Portier (2006). Rather, TFP exhibits a slow, occasionally accelerating growth, and other variables follow a similar pattern. During the first 10 quarters of propagation starting from date -10, TFP grows fast, followed by strongly increasing investment, consumption and hours. In light of these reactions, the news shock cannot be seen to trigger short-run fluctuations contributed by not-yet-materialised long-run movements in technology. Instead, the initial phase of propagation is characterised by improvements in TFP as well. This comovement suggests that the forward-looking variables respond to the shock not only through expectations but also due to the simultaneous change of productivity, disputing the anticipating nature of the news shock.

For comparison, I draw the impulse responses to the news shock from the causal model in Figure 5, which are estimated by least squares. The news shock is now identified by the Barsky-Sims strategy where it maximises the forecast error variance of TFP over 20 quarters with no impact effect, using the MA representation of the variables in levels.[28] Compared to Figure 4, the news shock of the causal VAR has a stronger impact effect on investment, hours, GDP and inflation. In particular, while the impulse responses of the noncausal VAR document a contemporaneous increase of productivity and forward-looking variables, the causal VAR produces stronger immediate reactions of investment and hours in response to the future, long-run technological improvement.

In fact, the causal responses potentially misinterpret the impact effects under nonfundamentalness. When noninvertibility matters, a causal VAR depicts reactions to a shock from the incorrectly identified, nonfundamental error. As a result, the identified shock may produce excessively strong initial responses, reflecting the past shocks already anticipated by the economic agents.[29] Consequently, the most recent shocks are dismissed. In contrast, the inclusion of lead terms allows for the anticipated shock, and the resulting responses of the noncausal VAR show the earliest, smoother reactions at the negative lags.

How much do news shocks contribute to the fluctuations? In place of the forecast error variance decomposition of causal VAR models, the noncausal VAR measures the significance of the news shock by the relative variance due to news shocks over the horizon from $-H_1$ to H_2:

$$\rho_1(y_{i,t}, H_1, H_2) = \frac{\text{Var}(y_{i,t}^1)}{\text{Var}(y_{i,t})} = \frac{\sum_{j=-H_1}^{H_2} e_i' \bar{\Psi}_j A_0 \gamma_1 \gamma_1' A_0' \bar{\Psi}_j' e_i}{\sum_{j=-H_1}^{H_2} e_i' \bar{\Psi}_j \Sigma \bar{\Psi}_j' e_i} \tag{25}$$

[28] Results remain similar when estimating a causal VAR in levels with a constant term as well as estimating a causal VAR with multivariate t-distributed errors. In particular, the risk of distortion caused by differencing variables, pointed out by Beaudry and Portier (2014), can be regarded as low.

[29] In particular, this observation is also consistent with the findings of Forni et al. (2014) who measure the reactions to a news shock to become smoother after conditioning on more information.

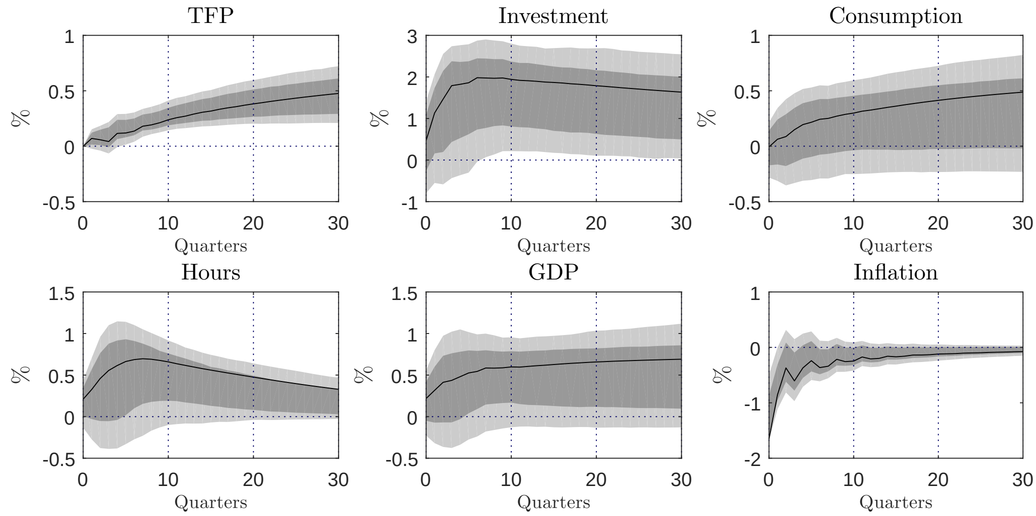


Figure 5: Impulse responses to a news shock from the causal Gaussian VAR(4)
The solid lines are the least squares estimates of the VAR(8) models. All responses are shown in levels. Light and dark grey shaded areas are bordered by the middle 90 % and 68 % confidence regions of bootstrap estimates. Confidence intervals are obtained from wild bootstrap with 1,000 replications. The news shock is identified as in Barsky and Sims (2011).

Hence, the news shock explains the share $\rho_1(y_{i,t}, H_1, H_2)$ of the movements of $y_{i,t}$. In Figure 6, the shares along with confidence bands are plotted for the baseline noncausal VAR model. The identified shock determines the major part of TFP at longer horizons, which is consistent with the observations of Beaudry and Portier (2006) who find the link between long-run movements of productivity and the news shock. The identified shock is also central in explaining fluctuations in investment, consumption, employment, output and inflation. However, the evidence is at odds with the news shock view in one important aspect: the shock explains now a considerable part of movements in TFP also in the short run.

Figure 7 plots the identified news shock $\bar{u}_{1,t}$ along with a cycle component of investment from the Hodrick-Prescott (HP) filter. In addition, NBER recessions are drawn as shaded grey areas. A direct linkage between these two time series is difficult to draw, but periods of positive news shocks tend to be followed by positive investment. Apart from the latest recession, the news shock does not take large values during the recessions but it rather comoves with investment during booms.

In conclusion, the news shock generates a steady growth of TFP together with modest positive reactions of forward-looking variables. The results confirm the positive comovement of consumption and investment conditional on the news shock, in line with the previous literature but difficult to be theoretically generated in dynamic general equilibrium models. Inferring from the above results, the comovement can be explained by the contemporaneous improvement of TFP, which gives the news shock a pattern similar to a shock of a diffusing technology process as in Lippi and Reichlin (1994a) or Walker and Leeper (2011). Hence, as also interpreted by Kurmann and Sims (2017), the news shock is related to sustained productivity growth and affects TFP on impact and with a lag.

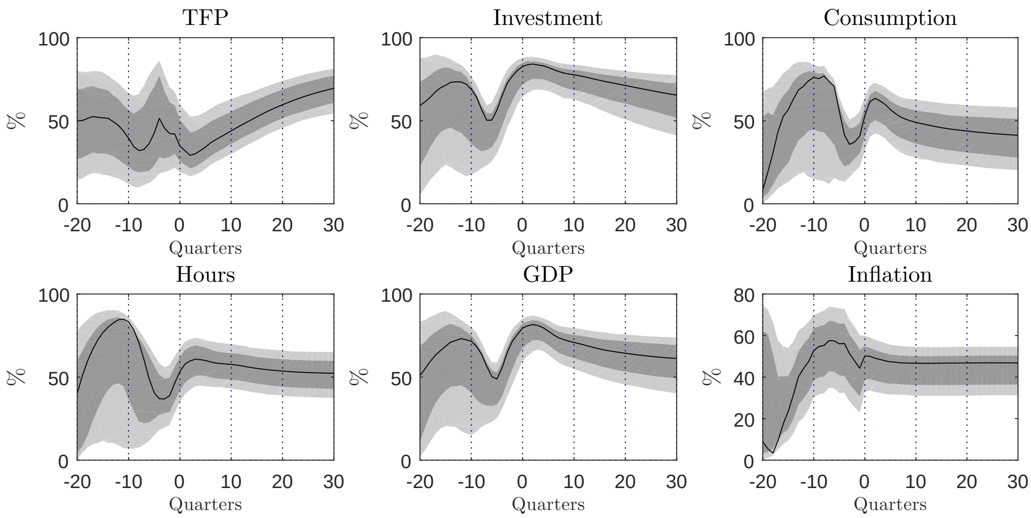


Figure 6: Share of variation for variables in levels due to identified news shocks, $\rho_1(y_{i,t}, H_1, H_2)$, from the noncausal VAR(5,3) model
Shares are computed from equation (25) with $H_1 = -20$ and H_2 is varied on the x-axes. The solid lines are the ML estimates. Light and dark grey shaded areas are bordered by the middle 90 % and 68 % confidence regions of the bootstrap estimates. Confidence intervals are obtained from wild bootstrap with 1,000 replications.

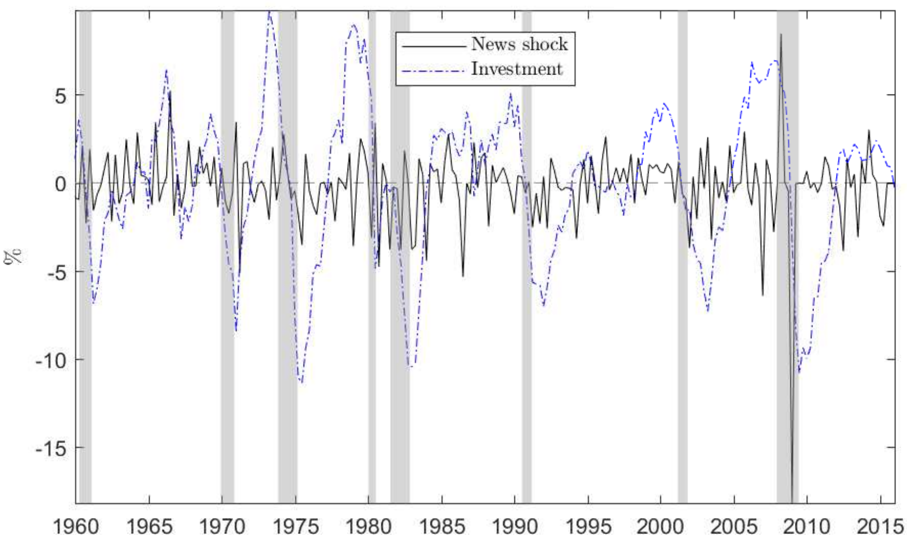
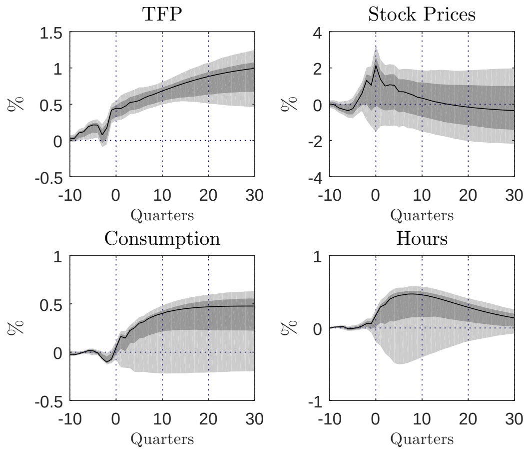
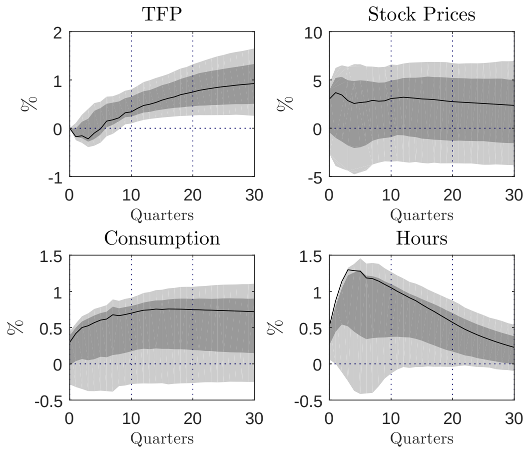


Figure 7: News shock and investment over time
The identified news shock is obtained from the noncausal VAR(5,3) model with variables of specification (1) and the cycle component of investment by Hodrick-Prescott (HP) filter. Investment is shown as percentage deviation from the HP-trend with smoothing parameter 1600. The shaded grey areas are the NBER recessions.





(a) Model (3): VAR(3,5) (b) Model (3): VAR(4)

Figure 8: Impulse responses to a news shock from the noncausal VAR(3,5) and Gaussian VAR(8) for specification (2)

The solid lines are the ML estimates of responses to a news shock, in panel (a) from the Gaussian VAR using the Barsky-Sims identification and in panel (b) from the noncausal VAR using the baseline identification explained in the text. All responses are shown in levels. Light and dark grey shaded areas are bordered by the middle 90 % and 68 % confidence regions of the bootstrap estimates. Confidence intervals are obtained from wild bootstrap with 1,000 replications.

4.4 Inclusion of Stock Prices and Interest Rates

Next, I report results related to specification (2) that replaces, for robustness, investment with the stock price index and identifies the news shock through its future-dependence with respect to TFP. In Figure 8, I plot the responses from the causal VAR(4) and noncausal VAR(3,5) models in panels (a) and (b), respectively. Conclusions from the noncausal responses stay similar to that from Figure 4. A news shock induces a modest increase of stock prices, consumption and hours, following the path of TFP. The responses peak at lag 0 when half of the news shock has incorporated to productivity, followed by a gradual fall in the other variables. Importantly, including stock prices does not alter the inference drawn from the noncausal VAR of specification (1). On the other hand, panel (b), similar to Figure 5, suggests that using the causal VAR(4) model would lead to conclusions where the initial reactions of stock prices and hours are stronger.

Finally, to analyse the behaviour of interest rates, specification (3) included federal funds rate and term spread between the 10-year and federal funds rate in addition to the baseline variables. In Figure 9, the impulse responses from the noncausal VAR(5,3) model are plotted with their confidence bands. I additionally plot the 10-year rate which is backed out as a sum of these two variables. Now, the conclusions about the responses of TFP, investment, consumption, hours and inflation are consistent with those from specification (1) shown in Figure 4 but with somewhat wider confidence bands. Accordingly, the diffusion of the news shock is characterised by smooth changes in the economic variables. Moreover, the federal funds rate initially declines in response to the deflationary effects, consistently with inflation-targeting monetary policy. This decline causes a slight increase of the term spread as the 10-year rate changes only mildly. At date 0, once TFP starts its steady

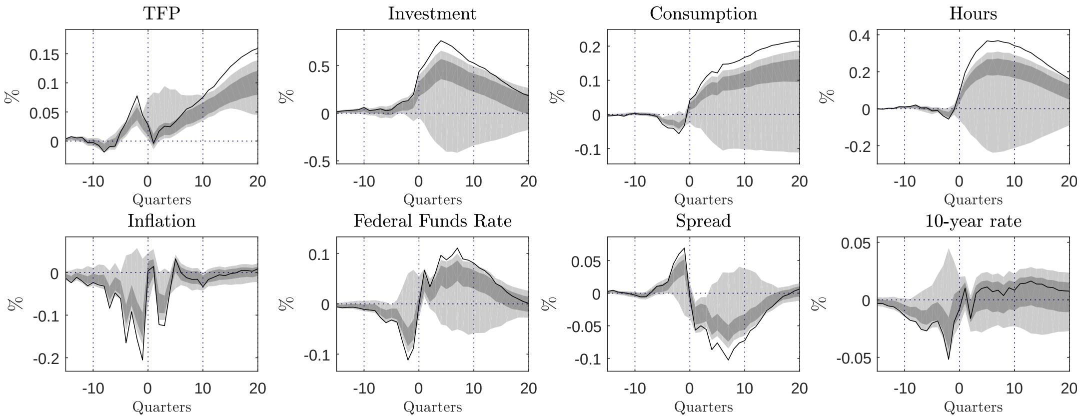


Figure 9: Impulse responses to a news shock from the noncausal VAR(5,3) – Specification (3)
The solid lines are the ML estimates of the model using the baseline identification. All responses are shown in levels. Light and dark grey shaded areas are bordered by the middle 90 % and 68 % confidence regions of the bootstrap estimates for the news shock. Confidence intervals are obtained from wild bootstrap with 1,000 replications. The 10-year rate is computed as a sum of spread and the federal funds rate.

growth followed by increasing investment, consumption and hours, the spread turns negative caused by the increasing federal funds rate. Similar to Kurmann and Otrok (2013), the news shock is measured to have has deflationary effects that decreases the federal funds rate resulting in a positive change in the term structure. However, the spread turns negative once the federal funds rate starts to increase.

5 Conclusions

The presence of news shocks permits economic agents to respond not only to current but also to future changes in productivity. However, causal VAR models fail to provide reliable evidence in validating these dynamics under nonfundamentalness, when economic agents possess more information than an econometrician. Consequently, the inference drawn from a VAR model may be based on misinterpreted structural shocks.

Allowing for noncausality instead, the nonfundamentalness issue is resolved by including future terms that capture variation omitted in an informationally deficient causal VAR. Structural analysis on the anticipated shocks can then be conducted using the two-sided MA representation with the identification scheme I introduced. Exploiting the multiplicative structure of the noncausal VAR model, I identified the news shock as the most significant factor moving productivity and future dependence of a forward-looking variable arising from nonfundamentalness. According to the Monte Carlo simulations, the approach is able to, first, detect nonfundamentalness in the form of noncausality and, second, recover the responses to a news shock under nonfundamentalness.

The evidence from the U.S. economy suggested that news shock induced contemporaneous increases in investment, hours, output and consumption as well as in total factor productivity, which may be the determining factor behind the strong anticipating responses observed earlier in the VAR literature. Rather than being a pure signal about productivity in the distant future, the news shock diffuses into TFP both in the short and long run. Mirroring this evidence with the news shock literature, the results confirm that the news shock explains the major part of long-run movements in TFP and is as such central in contributing to business cycle fluctuations. Nonetheless, it is difficult to view the news shock as triggering strong short-run reactions with materialisation in productivity only in the long run.

Finally, the analysis proceeded with the following limitations in mind. First, the noncausal VAR model I considered has a multiplicative form that may rule out certain noncausal representations. The link could be shown to exist in the Monte Carlo simulations, and the risk of misspecification can also be minimised by optimally choosing the lag and lead orders. However, it is not a direct empirical alternative to a noninvertible model but provides a sufficiently accurate approximation. Second, the multivariate t-distribution provides a simple departure from Gaussianity by adding a single volatility term to the error term that may contain both anticipated and unanticipated shocks. Hence, establishing identifiability and estimation theory for more general non-Gaussian distributions or conditional heteroskedasticity can be viewed as useful extensions. I leave these considerations for subsequent research.

References

Arezki, R., V. A. Ramey, and L. Sheng (2017). News shocks in open economies: Evidence from giant oil discoveries*. *The Quarterly Journal of Economics 132*(1), 103.

Ascari, G., G. Fagiolo, and A. Roventini (2015, 3). Fat-tail distributions and business-cycle models. *Macroeconomic Dynamics 19*, 465–476.

Barnichon, R. and C. Matthes (2015). Gaussian mixture approximations of impulse responses and the non-linear effects of monetary shocks. Working paper.

Barsky, R. B., S. Basu, and K. Lee (2015). Whither News Shocks? *NBER Macroeconomics Annual 29*(1), 225–264.

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

Beaudry, P., P. Fève, A. Guay, and F. Portier (2015). When is nonfundamentalness in VARs a real problem? an application to news shocks. Technical report, National Bureau of Economic Research.

Beaudry, P. and F. Portier (2006). Stock prices, news, and economic fluctuations. *American Economic Review 96*(4), 1293–1307.

Beaudry, P. and F. Portier (2014). News-driven business cycles: Insights and challenges. *Journal of Economic Literature 52*(4), 993–1074.

Chib, S. and S. Ramamurthy (2014). DSGE models with Student-t errors. *Econometric Reviews 33*(1-4), 152–171.

Cúrdia, V., M. del Negro, and D. L. Greenwald (2014). Rare shocks, great recessions. *Journal of Applied Econometrics 29*(7), 1031–1052.

Davis, R. A. and L. Song (2012). Noncausal vector AR processes with application to economic time series.

Fagiolo, G., M. Napoletano, and A. Roventini (2008). Are output growth-rate distributions fat-tailed? some evidence from OECD countries. *Journal of Applied Econometrics 23*(5), 639–669.

Fernald, J. (2012). A quarterly, utilization-adjusted series on total factor productivity. Working paper, Federal Reserve Rank of San Francisco.

Fernández-Villaverde, J., J. F. Rubio-Ramírez, T. J. Sargent, and M. W. Watson (2007). ABCs (and Ds) of understanding VARs. *American Economic Review 97*(3), 1021–1026.

Forni, M. and L. Gambetti (2014). Sufficient information in structural VARs. *Journal of Monetary Economics 66*(0), 124 – 136.

Forni, M., L. Gambetti, and L. Sala (2014). No news in business cycles. *The Economic Journal 124*(581), 1168–1191.

Francis, N., M. T. Owyang, J. E. Roush, and R. DiCecio (2014). A flexible finite-horizon alternative to long-run restrictions with an application to technology shocks. *Review of Economics and Statistics 96*(4), 638–647.

Galí, J. (2009). *Monetary Policy, inflation, and the Business Cycle: An introduction to the New Keynesian Framework*. Princeton University Press.

Gouriéroux, C. and J.-M. Zakoian (2013, February). Explosive Bubble Modelling by Noncausal Process. Working Papers 2013-04, Centre de Recherche en Economie et Statistique.

Hansen, L. P. and T. J. Sargent (1991). Two difficulties in interpreting vector autoregressions. In L. P. Hansen and T. J. Sargent (Eds.), *Rational Expectations Econometrics*, pp. 77–120. Westview Press Boulder, CO.

Kurmann, A. and C. Otrok (2013). News shocks and the slope of the term structure of interest rates. *The American Economic Review 103*(6), 2612–2632.

Kurmann, A. and E. Sims (2017). Revisions in utilization-adjusted TFP and robust identification of news shocks. Technical report, National Bureau of Economic Research.

Lanne, M. and J. Luoto (2016). Noncausal bayesian vector autoregression. *Journal of Applied Econometrics 31*(7), 1392–1406.

Lanne, M. and P. Saikkonen (2013). Noncausal Vector Autoregression. *Econometric Theory 29*(03), 447–481.

Leeper, E. M., T. B. Walker, and S. C. S. Yang (2013). Fiscal Foresight and Information Flows. *Econometrica 81*(3), 1115–1145.

Lippi, M. and L. Reichlin (1994a). Diffusion of technical change and the decomposition of output into trend and cycle. *The Review of Economic Studies 61*(1), 19–30.

Lippi, M. and L. Reichlin (1994b). VAR analysis, nonfundamental representations, Blaschke matrices. *Journal of Econometrics 63*(1), 307 – 325.

Lütkepohl, H. (2014). Fundamental problems with nonfundamental shocks. In N. Haldrup, M. Meitz, and P. Saikkonen (Eds.), *Essays in Nonlinear Time Series Econometrics*. Oxford University Press.

Plagborg-Møller, M. (2016). Bayesian inference on structural impulse response functions.

Rosenblatt, M. (2000). *Gaussian and Non-Gaussian Linear Time Series and Random Fields*. Springer Science & Business Media.

Schmitt-Grohé, S. and M. Uribe (2012). What's news in business cycles. *Econometrica 80*(6), 2733–2764.

Sims, C. A. (2002). Solving Linear Rational Expectations Models. *Computational Economics 20*(1-2), 1–20.

Sims, E. R. (2012). News, non-invertibility, and structural VARs. *Advances in Econometrics 28*, 81.

Uhlig, H. (2004). Do Technology Shocks Lead to a Fall in Total Hours Worked? *Journal of the European Economic Association 2*(2-3), 361–371.

Walker, T. and E. Leeper (2011). Information flows and news driven business cycles. *Review of Economic Dynamics 14*(1), 55–71.

A Appendix

A.1 Noncausal Representation of an Noninvertible MA Model

To derive (5), let l roots of $|B(z)|$ lie inside the unit circle with l_0 roots equal to zero. By standard algebra, $|B(z)|$ as scalar polynomial of order kd can be factorised to $|B(z)| = z^{l_0}(1 - z_{l_0+1}^{-1}z)\dots(1 - z_{kd}^{-1}z)$, and

$$|B(z)|^{-1} = z^{-l_0}\left(\prod_{i=l_0+1}^{l}(1 - z_i^{-1}z)\right)^{-1}\left(\prod_{i=l+1}^{kd}(1 - z_i^{-1}z)\right)^{-1} = z^{-l}\bar{c}_l\alpha(z^{-1})^{-1}\beta(z)^{-1}$$

with

$$\alpha(z^{-1}) = \prod_{i=l_0+1}^{l}(1 - z_i z^{-1}),\ \beta(z) = \prod_{i=l+1}^{kd}(1 - z_i^{-1}z),$$

and $\bar{c}_l = (-1)^{l-l_0}\prod_{i=l_0+1}^{l} z_i$. Now, scalars $\alpha(z^{-1})^{-1}$ and $\beta(z)^{-1}$ are well-defined power series expansions in z^{-1} and z, respectively, decaying to zero at geometric rate. Consequently, the inverse of $B(z)$ is

$$B(z)^{-1} \quad = \quad z^{-l}\bar{c}_l\alpha(z^{-1})^{-1}\beta(z)^{-1}B^{adj}(L),$$

where $B^{adj}(z)$ is the adjoint matrix of $B(z)$ of degree at most $(k-1)d$. Hence,

$$L^{-l}\bar{c}_l\beta(L)^{-1}B^{adj}(L)\alpha(L^{-1})^{-1}A(L)y_t = u_t$$

or

$$\bar{c}_l\beta(L)^{-1}B^{adj}(L)\alpha(L^{-1})^{-1}A(L)y_t = u_{t-l}$$

A.2 Estimation of the Noncausal VAR model

The log-likelihood function of the noncausal VAR is

$$l(\theta) = \sum_{t=r+1}^{T-s} \log f(\epsilon_t(\upsilon)'\Sigma^{-1}\epsilon_t(\upsilon); \lambda)$$

with density $f(\cdot; \lambda)$ of the multivariate t, $\upsilon = (\pi, \phi)$, $\pi = \text{vec}(\Pi)$, $\phi = \text{vec}(\Phi)$, scale matrix Σ, degrees-of-freedom parameter λ, and residuals

$$\epsilon_t(\upsilon) = v_t(\phi) - \sum_{j=1}^{r} \Pi_j(\pi) v_{t-j}(\phi),$$

where

$$v_t(\phi) = y_t - \Phi_1(\phi)y_{t+1} - \ldots - \Phi_s(\phi)y_{t+s}.$$

Derived by Lanne and Saikkonen (2013), the maximum likelihood estimator is asymptotically normal as

$$\sqrt{T}(\hat{\theta} - \theta_0) \xrightarrow{d} \text{N}(0, \mathcal{I}_{\theta\theta}(\theta_0)^{-1})$$

with $\mathcal{I}_{\theta\theta}(\theta_0) = -(T - r - s)^{-1}\text{E}\left[\frac{\partial^2 l(\theta_0)}{\partial\theta\partial\theta}\right]$. $\mathcal{I}_{\theta\theta}(\theta_0)$ can be consistently estimated by $-(T - r - s)^{-1}\frac{\partial^2 l(\hat{\theta})}{\partial\theta\partial\theta'}$.

A.3 Deriving the MA coefficients of the Stylised Example

The MA coefficients Ψ_j of the noncausal VAR (15) solve recursion

$$\begin{aligned} \Psi_j &= \Pi_1^j + \Phi_1\Psi_{j+1} + \Phi_2\Psi_{j+2}, \; j \geq 0 \\ \Psi_j &= \Phi_1\Psi_{j+1} + \Phi_2\Psi_{j+2}, \; j < 0. \end{aligned}$$

By solving for each j,

$$\begin{aligned} \Psi_j &= \begin{bmatrix} \rho^j & 0 \\ \theta\rho^j & 0 \end{bmatrix}, j > 0, \\ \Psi_0 &= \begin{bmatrix} 1 & 0 \\ \beta\rho\theta & 1 \end{bmatrix} \\ \Psi_j &= \begin{bmatrix} 0 & 0 \\ \theta\beta^{-j} & 0 \end{bmatrix}, j = -1, -2 \\ \Psi_j &= \mathbf{0}, j < -2 \end{aligned}$$

The impulse responses are then computed from

$$\begin{aligned} y_t &= \sum_{j=-\infty}^{\infty} \Psi_j\epsilon_{t-j} = \sum_{j=-\infty}^{\infty} \Psi_j \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_{t-2-j}^a \\ \nu_{t-j} \end{bmatrix} \\ &= \sum_{j=-\infty}^{\infty} \begin{bmatrix} \psi_{j,11}\varepsilon_{t-2-j}^a + \psi_{j,12}(\varepsilon_{t-2-j}^a + \nu_{t-j}) \\ \psi_{j,21}\varepsilon_{t-2-j}^a + \psi_{j,22}(\varepsilon_{t-2-j}^a + \nu_{t-j}) \end{bmatrix} \end{aligned}$$

as

$$\frac{\partial a_{t+j}}{\partial \varepsilon^a_{t-2}} = \psi_{11,j} + \psi_{12,j} = \psi_{11,j}, \qquad \frac{\partial a_{t+j}}{\partial \nu_t} = \psi_{12,j} = 0,$$
$$\frac{\partial x_{t+j}}{\partial \varepsilon^a_{t-2}} = \psi_{21,j} + \psi_{22,j}, \qquad \frac{\partial x_{t+j}}{\partial \nu_t} = \psi_{22,j},$$

where $\psi_{k,mn}$ is the (m,n) element of matrix Ψ_j. The impulse responses from the noncausal model to the anticipated shock ε^a_{t-2} are thus

$$\frac{\partial a_{t+j}}{\partial \varepsilon^a_{t-2}} = \begin{cases} \rho^j, & j \geq 0 \\ 0, & j < 0 \end{cases}, \quad \frac{\partial x_{t+j}}{\partial \varepsilon^a_{t-2}} = \begin{cases} \theta\rho^j, & j \geq 0 \\ \theta\beta^{-j}, & -2 \leq j < 0 \\ 0, & j < -2 \end{cases},$$

which coincide with the theoretical impulse responses.

A.4 Noncausal Representation in the Stylised Example for a More General Technology Process

Instead of (10), consider a more general technology process

$$a_t = \rho a_{t-1} + \varepsilon^a_{t-q} + \chi \varepsilon^a_t, \tag{26}$$

and assume it determines the equilibrim together with (11). Accordingly, a shock ε^a_t affects both a_t and a_{t+q} such that agents gradually learn about the future technology q periods forward.[30] If $\chi < 1$, ε^a_t has a greater contribution to a_{t+q} than to a_t, in which case the ARMA process for a_t is primarily driven by shocks observed by the economic agents before t.

Similarly as before, the solution can be derived from (12) such that $x_t = \theta z_t + \nu_t$ for $q = 0$ and

$$x_t = \theta a_t + \theta \sum_{j=0}^{q-1} \beta^{q-j} \varepsilon^a_{t-j} \tag{27}$$

for $q > 0$. Let now the anticipation horizon be $q = 2$. By inserting (10) to the latter equation and collecting terms, $y_t = (a_t, x_t)'$ follows

$$y_t = \begin{bmatrix} \rho & 0 \\ \theta\rho & 0 \end{bmatrix} y_{t-1} + \underbrace{\begin{bmatrix} \chi + L^2 & 0 \\ \theta(\beta^2 + \chi) + \theta\beta L + \theta L^2 & 1 \end{bmatrix}}_{=B(L)} \begin{bmatrix} \varepsilon^a_t \\ \nu_t \end{bmatrix}. \tag{28}$$

If movements in the exogenous process for a_t are dominated by the anticipated lag term, i.e. $\chi < 1$ and $q > 0$, $y_t = (a_t, x_t)$ is noninvertible in the past since $|B(z)| = 0$ for $z = \pm\sqrt{\chi}i$. In other words, as soon as the news shock contributes to a_t relatively more, $\chi < 1$, the observables suffer from nonfundamentalness and no causal VAR representation in terms of y_t exists for structural shocks $u_t = (\varepsilon^a_t, \nu_t)$.

[30] Walker and Leeper (2011) describe ε^a_t as a correlated news shock. An example with a similar process is also considered in Beaudry and Portier (2014).

As before, the solution can be written in terms of future observables. In particular, rewriting the noninvertible process (26) as

$$(1-\rho L)a_t = (L^2+\chi)\varepsilon_t^a, \tag{29}$$

and its right-hand side as $(1+\chi L^{-2})\varepsilon_{t-2}^a$, the modified lag polynomial, $1+\chi z^2$, has no roots smaller than one in modulus such that (29) implies

$$(1-\rho L)(1+\chi L^{-2})^{-1}a_t = \varepsilon_{t-2}^a. \tag{30}$$

Therefore, a_t is noncausal with a time-shifted error, and the shock ε_{t-2}^a is a function of the past and future values of a_t. Using the representation (30) to substitute the shocks ε_{t-1} and ε_t in (28), x_t has a noncausal form,

$$x_t = \rho a_{t-1} + \Big(\beta - \chi\rho + \theta(\beta^2+\chi)L^{-1}\Big)\sum_{j=0}^{\infty}(-\chi L^{-2})^j a_{t+1} + \varepsilon_{t-2}^a + \nu_t \tag{31}$$

consisting of an infinite number of leads of a_t. Furthermore, a_t and x_t in (30) and (31) satisfy

$$(I_2 - \Pi_1 L)(I_2 - \Phi_1 L^{-1} - \Phi_2 L^{-2} - \ldots)y_t = \epsilon_t, \tag{32}$$

with

$$\Phi_j = \begin{bmatrix}\phi_{j,11} & 0\\ \phi_{j,21} & 0\end{bmatrix},$$

$$\begin{aligned}\phi_{j,11} &= \begin{cases}0, & j=1,3,\ldots\\ -(-\chi)^{\frac{j}{2}}, & j=2,4,\ldots\end{cases},\\ \phi_{j,21} &= \begin{cases}\beta(-\chi)^{\frac{j-1}{2}}, & j=1,3,\ldots\\ \theta(\beta^2+\chi)(-\chi)^{\frac{j}{2}-1}, & j=2,4,\ldots\end{cases}.\end{aligned}$$

and the error term

$$\epsilon_t = \begin{bmatrix}1 & 0\\ 1 & 1\end{bmatrix}\begin{bmatrix}\varepsilon_{t-2}^a\\ \nu_t\end{bmatrix}.$$

In particular, the noncausal representation (32) is directly related to the underlying model when the infinite number of lead terms is truncated by large lead order s, and the noncausal VAR(1,s) recovers a linear combination of structural shocks ε_{t-2}^a and ν_t.

Finally, the MA coefficients are numerically obtained from recursions

$$\begin{aligned}\Psi_j &= \Pi_1^j + \Phi_1\Psi_{j+1} + \ldots + \Phi_S\Psi_{j+S},\ j \geq 0\\ \Psi_j &= \Phi_1\Psi_{j+1} + \ldots + \Phi_S\Psi_{j+S},\ j < 0,\end{aligned}$$

where S is a sufficiently large integer. The impulse responses of the noncausal VAR(1,s) and the theoretical model are now plotted in the lower and upper plots of Figure 10, respectively, for $\chi = 0.5$. Similar to Figure 1, the noncausal impulse responses replicate the theoretical counterparts shown in the upper panel.

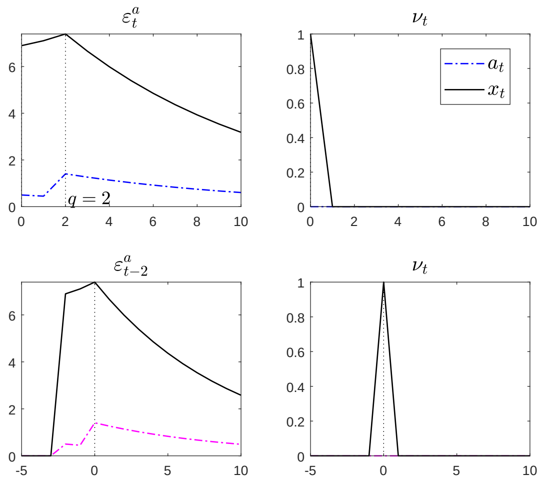


Figure 10: Theoretical and empirical impulse response functions of the model with $q = 2$, $\beta = 0.9$, $\rho = 0.9$ and $\chi = 0.5$
The upper figures show the theoretical impulse responses of a_t and x_t to shocks ε_t^a and ν_t. The lower row plots the impulse responses obtained from the MA representation of the noncausal VAR to the shocks in the error term, ε_{t-2}^a and ν_t.

A.5 Solution to the Identification Problem

w_1 solves

$$\max_{w_1} \left|\mathrm{Cov}(f_{i,t}^1, y_{2,t}^1)\right| \tag{33}$$

with

$$\mathrm{Cov}(f_{i,t}^1, y_{1,t}^1) = \mu_{\omega^{-1}} \sum_{j=-H_1}^{H_2} e_i' \Psi_j^* \widetilde{A}_0 w_1 w_1' \widetilde{A}_0' \Psi_j' e_1$$

subject to

$$w_1' w_1 = 1.$$

First, note that by Cauchy-Schwarz inequality, the objective function is bounded from above, which ensures with the constraint that the maximum exists. Partition the problem to maximising separately $\mathrm{Cov}(f_{i,t}^1, y_{2,t}^1)$ and $-\mathrm{Cov}(f_{i,t}^1, y_{2,t}^1)$ subject to the same constraint. By manipulating

$$\begin{aligned} e_i' \Psi_j^* \widetilde{A}_0 w_1 w_1' \widetilde{A}_0' \Psi_j' e_1 &= \mathrm{tr}\left(e_i' \Psi_j^* \widetilde{A}_0 w_1 w_1' \widetilde{A}_0' \Psi_j' e_1\right) \\ &= \mathrm{tr}\left(w_1' \widetilde{A}_0' \Psi_j' e_1 e_i' \Psi_j^* \widetilde{A}_0 w_1\right) \\ &= w_1' \widetilde{A}_0' \Psi_j' E_{1i} \Psi_j^* \widetilde{A}_0 w_1 \\ &= w_1' S_{1i}^j w_1, \end{aligned}$$

the covariance equals

$$\mathrm{Cov}(f_{i,t}^1, y_{1,t}^1) = w_1' \left(\mu_{\omega^{-1}} \sum_{H_1}^{H_2} S_{1i}^j\right) w_1 = w_1' \Omega w_1.$$

Now, the Lagrangian for maximising positive covariance is

$$\mathcal{L} = w_1'\Omega w_1 - \lambda(w_1'w_1 - 1)$$

with first-order condition

$$\frac{\Omega + \Omega'}{2} w_1 = \lambda w_1$$

so w_1 is an eigenvector of matrix $\frac{\Omega+\Omega'}{2}$ with only real eigenvalues. Since for any matrix Ω and vector w_1 holds that $w_1'\Omega w_1 = w_1' \frac{\Omega'+\Omega}{2} w_1$, inserting the solution to the covariance yields

$$\mathrm{Cov}(f_{i,t}^1, y_{1,t}^1) = \lambda,$$

i.e., the largest eigenvalue of $\frac{\Omega+\Omega'}{2}$ maximises the covariance. Analogously, finding w_1 that minimises the covariance corresponds to finding the smallest eigenvalue of matrix $-\frac{\Omega+\Omega'}{2}$. Finding the eigenvector of the largest eigenvalue in absolute value maximises (33).

B Monte Carlo Simulations

B.1 NK Model of Section 3

Equilibrium is determined by equations

$$\begin{aligned}
\tilde{y}_t &= -\frac{1}{\sigma}\left(i_t - \mathrm{E}_t \pi_{t+1} - r_t^n\right) + \mathrm{E}_t \tilde{y}_{t+1} \\
\pi_t &= \beta \mathrm{E}_t[\pi_{t+1}] + \kappa \tilde{y}_t, \\
i_t &= \rho(1-\rho_m) + \rho_m i_{t-1} + \phi_\pi \pi_t + \phi_y \tilde{y}_t + \varepsilon_t^m,
\end{aligned}$$

by $r_t^n = \rho + \sigma \psi_{ya}^n \mathrm{E}_t \Delta a_{t+1}$ and technology (24). Coefficients are functions of deep parameter of the model:

$$\begin{aligned}
\kappa &= \lambda(\sigma + \frac{\phi + \alpha}{1-\alpha}), \\
\lambda &= \frac{(1-\theta)(1-\beta\theta)}{\theta}\Theta \\
\Theta &= \frac{(1-\alpha}{1-\alpha+\alpha\varepsilon} \\
\psi_{ya} &= \frac{1+\phi}{\sigma(1-\alpha)+\phi+\alpha}
\end{aligned}$$

The calibration is summarised in Table 3.

Discount factor	β	0.99
Risk averion	σ	2
Frisch elasticity	ϕ	1
Calvo paremeter	θ	2/3
Capital share	α	1/3
Elasticity of substitution	ε	6
Taylor rule coefficient for inflation	Φ_π	1.5
Taylor rule coefficient for output gap	Φ_y	0.5/4
Persistence of interest rate	ρ_m	0.8
Persistence of technology	ρ_a	0.9

Table 3: Calibration of the New Keynesian Model

B.2 Further Simulation Results

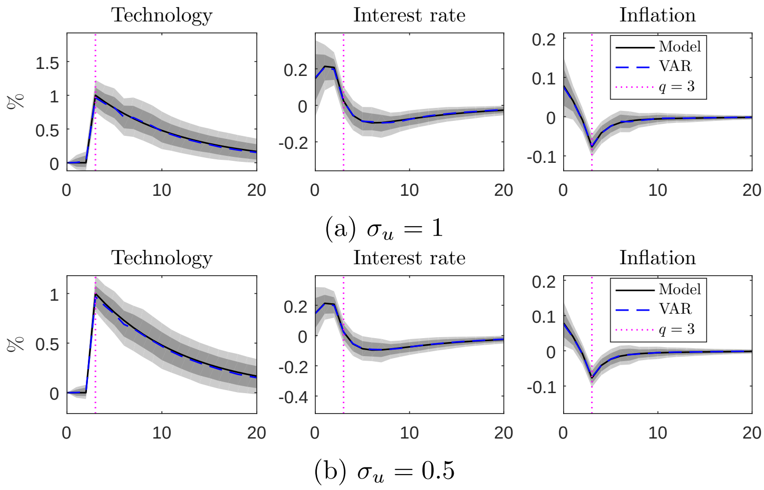
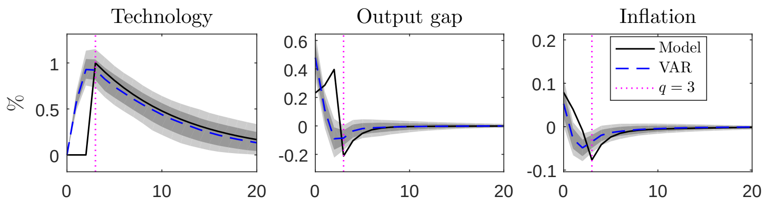
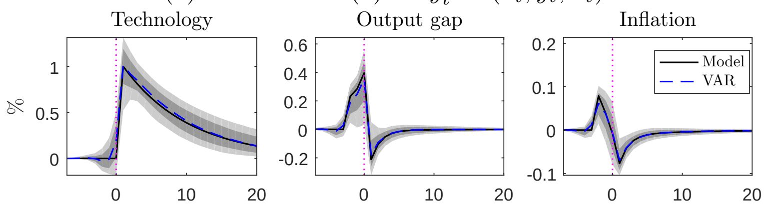


Figure 11: Impulse Responses to a News Shock from the Causal VAR(5) with Fundamental Observables $y_t^2 = (a_t, i_t, \pi_t)$ in the NK model with $q = 3$ and $\sigma_u \in \{0.5, 1\}$
The dashed lines are the median estimated impulse responses from the Monte Carlo samples. Light and dark grey shaded areas border the middle 90 % and 68 %, respectively, of the distribution for estimated impulse responses. Black solid lines are the theoretical impulse responses.



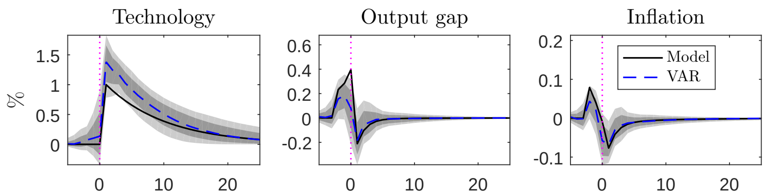
(a) Causal VAR(3) on $y_t^1 = (a_t, \tilde{y}_t, \pi_t)$



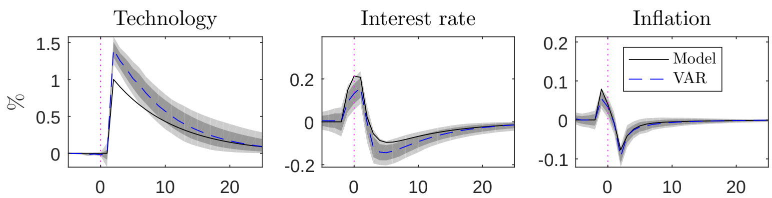
(b) Noncausal VAR(1,2) on $y_t^2 = (a_t, i_t, \pi_t)$

Figure 12: Impulse Responses to a News Shock from a Causal and Noncausal VAR in the NK model with $q = 3$ and $\sigma_u = 0.5$

The dashed lines are the median estimated impulse responses from the Monte Carlo samples. Light and dark grey shaded areas border the middle 90 % and 68 %, respectively, of the distribution for estimated impulse responses. Black solid lines are the theoretical impulse responses, aligned in panel (b) with the estimated noncausal impulse responses according to the maximum impact on technology.



(a) VAR on $y_t^1 = (a_t, \tilde{y}_t, \pi_t)$



(b) VAR on $y_t^2 = (a_t, i_t, \pi_t)$

Figure 13: Impulse Responses to a News Shock from a Noncausal VAR(3,3) in the NK Model with $\chi = 1$ and $q = 3$ and fat-tailed shocks

The model is simulated by structural shocks drawn in the following way. First, draw u_t from $N(0, I)$. Second, multiply u_t with probability 0.1 by 3 to put weight on tails. Last, standardise the new error series and simulate y_t. The dashed lines are the median estimated impulse responses from the Monte Carlo samples. Light and dark grey shaded areas border the middle 90 % and 68 %, respectively, of the distribution for estimated impulse responses. Black solid lines depict the theoretical impulse responses, aligned with the estimated noncausal impulse responses according to the maximum impact on technology.