RECent
Center for Economic Research

WORKING PAPER SERIES

Informing DSGE Models through Dynamic Factor Models

Mario Forni, Luca Gambetti, Marco Lippi, Luca Sala

Working Paper 160

March 2025

www.recent.unimore.it

RECent: c/o Dipartimento di Economia Politica, Viale Berengario 51, I-41100 Modena, ITALY
Phone +39 059 2056856, Fax +39 059 2056947, E-mail: RECent@unimore.it

Informing DSGE Models through Dynamic Factor Models

Mario Forni*
Università di Modena e Reggio Emilia, CEPR and RECent
Luca Gambetti†
Universitat Autònoma de Barcelona, BSE, Università di Torino, CCA
Marco Lippi
Einaudi Institute for Economics and Finance
Luca Sala‡
Università Bocconi, IGIER and Baffi Carefin

Abstract

Structural Dynamic Factor Models (SDFM) represent a reliable tool to inform the construction of Dynamic Stochastic General Equilibrium (DSGE) models. The reason is that the log-linear solution of a DSGE model has a factor structure which ensures consistency between the representations of the two models. We assess the usefulness of SDFM for DSGE analysis by means of simulations. Using a standard DSGE model as the Data Generating Process, we show that the factor model always provides accurate estimates of the impulse response functions. As an application, we reassess the literature studying the response of hours to technology shock. An additional application studies the effects of monetary policy.

JEL classification: C32, E32.

Keywords: DSGE models, structural VAR, structural factor models.

*Financial support from FAR 2017, Department of Economics "Marco Biagi", is gratefully acknowledged.

†Luca Gambetti acknowledges the financial support of the Spanish Ministry of Economy and Competitiveness through grant ECO2015-67602-P and through the Severo Ochoa Programme for Centres of Excellence in R&D (SEV-2015-0563), and the Barcelona Graduate School Research Network.

‡Corresponding author. Address: Università Bocconi, via Roentgen 1, 20136 Milano. E-mail: luca.sala@unibocconi.it. Phone: +39 02 58363062.

1 Introduction

Dynamic Stochastic General Equilibrium (DSGE) and Structural Vector Autoregressions (SVAR) models play a complementary role in modern macroeconomic analysis. In particular, DSGE analysis relies on SVAR evidence for several purposes.

We see three ways the literature has used VARs to inform DSGE models. The first is to use SVARs as a benchmark to assess a DSGE's ability to fit the data. This can be done using the DSGE-VAR approach in Del Negro and Schorfheide (2004) and Del Negro et al. (2007), where the value of the VAR parameters implied by the DSGE model is used as a prior in a Bayesian VAR and the marginal likelihood is used to measure the goodness of the DSGE model. The second is to use SVARs in the estimation of the parameters of DSGE models, either via direct or indirect inference, see Smith (1993), Del Negro and Schorfheide (2004), Guerron-Quintana et al. (2017) and the vast literature cited therein. The third, and most popular, is to compare SVAR and DSGE impulse responses to some shock of interest, in order to verify if the theoretical transmission mechanisms is empirically plausible.[1]

This paper argues that Structural Dynamic Factor Models (SDFM) are better at this third task than SVARs and discusses the advantages of using SDFM as an empirical tool when the goal is to compare impulse responses to some structural shocks.[2] With this goal in mind, a necessary requirement is that the effects of a shock of interest have to be well estimated in the empirical model.

From a theoretical viewpoint, the main reason why SDFM delivers reliable estimates and possibly outperform SVARs is the following. The log-linear representation of the DSGE model postulates a VAR for the state variables in terms of the structural

[1] For example, Galì (1999) investigates the response of hours to technology shocks estimated in a SVAR to assess the empirical support of RBC and New Keynesian models (see also Christiano et al. (2007)) and Canova and Paustian (2011) assess the validity of models with rule-of-thumb consumers.

[2] Consolo et al. (2009) introduce a FAVAR-DSGE model. Understanding how to use SDFM to globally inform the DSGE model as a whole is left for future research.

shocks.[3] So, if the state variables were observed, the empirical strategy closest to the DSGE model would simply be to identify structural shocks in a VAR for the state variables. This is obviously unfeasible because the state variables are generally not observed. The SDFM provides however a consistent estimator of a linear transformation of the states, the factors. The factors are statistical objects with the property of spanning the same space spanned by the state variables of the model. This implies that using a VAR for the factors, as done in the factor model, is equivalent to using a VAR for the states. On the contrary, there is no guarantee that the variables used in the SVAR contain all the information contained in the state variables. This lack of information is often referred to as nonfundamentalness (Lippi and Reichlin, 1993, 1994), noninvertibility or informational deficiency (Forni et al., 2019). When the variables are not sufficiently informative, the estimated impulse response functions are not reliable and cannot provide useful guidance to inform the DSGE on the transmission mechanism of a given shock.[4,5]

Informational deficiency can arise because of two main reasons. First, there are too many shocks relative to the number of variables included in the SVAR. Second, even if the number of shocks is equal or less than the number of variables, the variables considered might not convey enough information. The first case arises for instance in presence of measurement error, see Lippi (2020) and Forni et al. (2020) for a formal discussion.[6] The second case might arise for example in situations where a shock affects

[3]Aruoba et al. (2022) have extended VAR models to non-linear setups. Here we focus on log-linear solutions.

[4]This problem has received much attention in the literature. Fernández-Villaverde et al. (2007) and Forni et al. (2019) provide conditions to check for invertibility. The former states a verifiable condition to understand whether a set of n variables included in the DSGE admits a SVAR representation to recover n structural shocks. The latter provides a DSGE model-based measure to assess whether a given VAR specification can be used to estimate *any* individual shock of interest and its impulse response functions.

[5]When the goal is to inform the DSGE on the transmission mechanism of a certain shock, indirect inference might be problematic, because matching the indirect statistic does not imply matching the true impulse responses.

[6]Baumeister and Hamilton (2019) have studied VAR models in presence of measurement error.

the variables with a delay, so that the current value of the variables does not provide enough information about the current shock. News shocks or anticipated government spending shocks are two notable examples studied in the literature, see Beaudry and Portier (2006), Forni et al. (2014) or Leeper et al. (2013). Here we propose a way out: using a factor model in place of a VAR solves the problem, since factor models are not affected by informational deficiency (Forni et al., 2009, 2020).

We assess the performance of the SDFM relative to SVARs in recovering impulse responses by means of Monte Carlo simulations, using the DSGE model of Justiniano et al. (2010) as Data Generating Process. We focus on technology and monetary policy shocks, which have been extensively studied in the literature, as well as news shocks, which have been shown to likely raise noninvertibility problems in SVAR models. Technology and news shocks are identified using long run restrictions á la Galì (1999), while monetary policy shocks are identified using both the SVAR-IV *external* instrument approach discussed in Stock and Watson (2018) and Mertens and Ravn (2013), and the *internal* instrument approach discussed in Plagborg-Møller and Wolf (2021), which should be robust to noninvertibilities.

Regarding the (non-anticipated) technology shock, we show that a two-variables SVAR including labor productivity and hours worked is not able to recover the correct impulse response functions. Moreover, results depend on the treatment of hours worked (log-level vs. log-differences). Similarly, the SVAR is unable to recover the correct impulse response function of a news (anticipated) technology shock.

As for the monetary policy shock, the SVAR fails in presence of measurement errors. As observed in Plagborg-Møller and Wolf (2021), including the instrument in a non-invertible VAR solves the problem. However, there is a price to pay. As the authors point out, the method requires the instrument to satisfy the so-called lead-lag exogeneity condition, which is more stringent than the usual exogeneity condition required by an invertible VAR. Conversely, the DFM produces correct results even if the instrument is only simultaneously exogenous.

The above findings help us understanding the results of the empirical sections. In the empirical analysis, we revisit the old-standing but still unsolved controversy about the response of hours worked to technology shocks. Galì (1999), using a bivariate VAR with hours taken in differences, finds that a shock increasing productivity reduces hours. On the other hand, Christiano et al. (2004), using hours taken in levels, finds that hours increase. The result is relevant for discriminating between alternative economic models. We find that (a) the shocks estimated with both bivariate VAR specifications are predicted by the factors of a DFM, meaning that both VARs are informationally deficient. Therefore using the estimated impulse response functions from those VARs to inform the theoretical model can be misleading; (b) the SDFM produces the same results, irrespectively of the treatment of hours worked; (c) hours significantly increase after a positive technology shock. Results (a) and (b) are exactly what we find in the simulation of Section 3.2.

The above analysis is related to, but different from, Forni and Gambetti (2014) and Forni et al. (2014). In the former paper we also focus on standard technology shocks, but do not study their effect on hours worked. We start with a VAR with labor productivity and the unemployment rate; then we estimate a FAVAR to show that results change when adding factors. In the latter paper we focus on news technology shocks. We start with a VAR with TFP and stock prices. Again, we estimate a FAVAR to show that results change when adding information.

As an additional application, we identify the effects of monetary policy with the authoritative instrument proposed by Gertler and Karadi (2015). We show that informationally deficient VAR models may produce puzzling results, even if the instrument is included in the VAR as suggested in Plagborg-Møller and Wolf (2021). This is exactly what we find in the simulation of Section 3.5. By contrast, the SDFM produces results which are in line with the theory.

The remainder of the paper is organized as follows. Section 2 presents a theoretical discussion of the consequence of the presence of measurement error and nonfundamentalness in SVAR. Section 3 presents the results of the simulations. Sections 4 and 5 present the empirical analyses. Section 6 concludes.

2 DSGE and dynamic factor models

In this Section we introduce the factor model and its relationship with DSGE models.

2.1 The dynamic factor model

Let x_t be a n-dimensional vector of economic variables.[7,8] The variables x_{it} are stationary, possibly after detrending, and can be represented as

$$x_{it} = \chi_{it} + \xi_{it}, \quad i = 1, \dots, \infty, \tag{1}$$

where ξ_{it} is the idiosyncratic component, χ_{it} is the common component and the two are orthogonal to each other at all leads and lags. Given t, the χ's, for $i \in \mathbb{N}$, span a finite-dimensional space, whose dimension is r. This implies that there exists an r-dimensional vector F_t, weakly stationary, whose coordinates are the static factors, such that

$$\chi_{it} = \lambda_{i1} F_{1t} + \cdots + \lambda_{ir} F_{rt} = \Lambda_i F_t \quad \text{or} \quad \chi_t = \Lambda F_t, \tag{2}$$

[7] A rigorous definition of a High-Dimensional Dynamic Factor Model requires that the vector x_t is part of an infinite-dimensional vector, so that we can make assumptions by letting n tend to infinity, see Forni et al. (2000), Stock and Watson (2002a,b), Bai and Ng (2002). Here, making reference to the version in Forni et al. (2009), we limit ourselves to recalling the main features of the model.

[8] Bańbura et al. (2010) and Giannone et al. (2015) propose large-dimensional VAR models.

where χ_t is the n-dimensional vector of the χ's and Λ is the $n \times r$ factor loading matrix.

The common components χ_{it} and the factors F_t are driven by a q-dimensional vector of structural common shocks u_t, with $q \leq r$. More specifically, the r-dimensional vector F_t has the VAR representation[9]

$$Q(L)F_t = \varepsilon_t = Su_t, \tag{3}$$

where $Q(L)$ is a stable polynomial matrix of order p with $Q(0) = I$ and S is a $r \times q$ matrix of constants. Equation (3) implies that the structural shocks belong to the information space spanned by the VAR residuals, so that the factors are *informationally sufficient* for u_t (i.e. the structural shocks are fundamental for F_t).

By inverting (3) and using (2), we obtain the impulse response function representation

$$x_t = \Lambda Q(L)^{-1} S u_t + \xi_t = \Phi(L) u_t + \xi_t. \tag{4}$$

where $\Phi(L)$ is the matrix of structural impulse response functions. Estimation, identification and inference are discussed in the online Appendix A.

2.2 Factor model representation of a DSGE model

Suppose that the data generating process is a DSGE model which admits the following state-space representation, the ABCD representation, as defined in Fernández-Villaverde et al. (2007),

$$s_t = As_{t-1} + Bu_t \tag{5}$$

$$\chi_t = Cs_{t-1} + Du_t \tag{6}$$

where u_t is a q-dimensional vector of structural shocks, χ_t is a n-dimensional vector

[9]The existence of a VAR representation for the factors is perfectly compatible with cointegration among the variables in the panel, see Forni et al. (2020).

of economic variables, s_t is an m-dimensional vector of stationary state variables ($q \leq m$), A, B, C and D are conformable matrices of parameters and B has a left inverse B^{-1} such that $B^{-1}B = I_q$.

From the ABCD representation we make explicit the link between the DSGE model and the factor model discussed above. From equations (5) and (6) we get

$$\chi_t = G f_t \tag{7}$$

where $G = (DB^{-1} \quad C - DB^{-1}A)$ and $f_t = \begin{pmatrix} s_t' & s_{t-1}' \end{pmatrix}'$. The vector f_t has the VAR representation

$$f_t = \tilde{A} f_{t-1} + \tilde{B} u_t \tag{8}$$

with $\tilde{A} = \begin{pmatrix} A & 0_m \\ I_m & 0_m \end{pmatrix}$, $\tilde{B} = \begin{pmatrix} B \\ 0_m \end{pmatrix}$, I_m is the m-dimensional identity matrix and 0_m a $m \times m$ matrix of zeros.[10] Equations (7)-(8) reduce to the minimal representation

$$\chi_t = \Lambda F_t \tag{9}$$

$$F_t = Q F_{t-1} + S u_t \tag{10}$$

where $\Lambda = GP$, $Q = P^{-1}\tilde{A}P$, $S = P^{-1}\tilde{B}$, P^{-1} being a left inverse of P.[11] There are two important remarks to notice about representation (9)-(10). First, $F_t = P^{-1} f_t$ spans the same information space as $f_t = (s_t' \; s_{t-1}')'$, so that F_t contains all the relevant information about the DSGE dynamics. Second, F_t follows the VAR representation (10), where, in general, the residuals have reduced rank, i.e. $q < r$.

[10]Unlike representation (2)-(3), representation (7) is not necessarily minimal (i.e., the representation of the model with the smallest number of factors), since the covariance matrix of f_t, Σ_f, may have reduced rank $r < 2m$. We can easily derive the minimal representation. If (7) is not minimal, then there exist a $2m \times r$ matrix P such that $f_t = PF_t$, where F_t has a nonsingular covariance matrix.

[11]Boivin and Giannoni (2006) and Gelfer (2019) have proposed to estimate DSGE models in a data-rich environment. Their approach is to link a large dimensional dataset to the states of the model by assuming a linear relation between the states and the observables.

By inverting the VAR for F_t, we obtain the MA representation

$$\chi_t = \Phi(L)u_t = \Lambda(I - QL)^{-1}Su_t, \tag{11}$$

where $\Phi(L)$ is the matrix of impulse-response functions. Assuming that the economic variables are observed with error, we obtain

$$x_t = \Lambda F_t + \xi_t = \Phi(L)u_t + \xi_t, \tag{12}$$

where ξ_t is a vector of measurement errors. By comparing (4) and (12) it is seen that, when the variables are observed with error, the linearized DSGE model has a factor representation with $Q(L) = I - QL$.

In practice, one can compare the empirical impulse response functions estimated with the factor model for a given shock to those of the DSGE. Under the null hypothesis that the DSGE is the DGP, the two should be similar. If not, this suggests some sort of misspecification of the DSGE model.

2.3 Measuring informational deficiency

Consider again the structural model (11):

$$\chi_t = \Phi(L)u_t \tag{13}$$

where each individual variable is observed with measurement error $x_t = \chi_t + \xi_t$. Suppose, as it is done in SVAR analysis, that we consider a q-dimensional subvector y_t of x_t: $y_t = Wx_t$, where W is a selection matrix selecting elements from x_t. The relationship between observables and shocks is then

$$y_t = \begin{pmatrix} W\Phi(L) & W \end{pmatrix} \begin{pmatrix} u_t \\ \xi_t \end{pmatrix}. \tag{14}$$

It is immediately seen that the presence of measurement error makes (14) noninvertible simply because there are more shocks than variables. In this sense, measurement error can be simply considered a type of information deficiency. With no measurement error y_t can still be noninvertible, depending on the roots of the determinant of $W\Phi(L)$.

Notice however that even if the representation is noninvertible, some (but not all) of the shocks and their impulse response functions can still be estimated. This is a very important point since a large part of the contributions in the VAR literature aims at identifying typically a single shock. Thus, the relevant question is: is the informational content of vector y_t enough to correctly estimate a given shock of interest? Sims and Zha (2006) and Forni et al. (2019) have proposed a measure of informational deficiency of a set of variables y_t relative to a shock u_{it}. The measure, called δ_i, is defined as one minus the R^2 of a regression of the structural shock u_{it} onto the residuals of a VAR on y_t. If deficiency is zero, $\delta_i = 0$, the structural shock is a linear combination of the VAR residuals: the present and the past of y_t contain enough information to estimate u_{it}. If deficiency is maximal, $\delta_i = 1$, the present and the past of y_t do not provide any information about the structural shock.

In the simulations that follow we rely on δ_i in order to understand whether a VAR for y_t is enough informative. The δ_i for each shock will take into account both informational deficiency of the variables themselves (i.e. the χ_i) and the impact of measurement errors in reconstructing the shock and its impulse responses.

3 Simulations

In this section we assess the validity of the factor model and the VAR as empirical tools to inform on DSGE models using Monte Carlo simulations. We focus on technology and monetary policy shocks, two shocks which have been extensively studied in the literature.

First, we consider a (non-anticipated) technology shock. We find that a two-variables VAR including labor productivity and hours worked is unable to recover the correct impulse response functions. Moreover, results vary substantially depending on the treatment of hours (log-levels vs. log-differences). By contrast, the SDFM performs well, independently of the treatment of hours. These findings help us understanding the empirical application of Section 4: according to VAR estimation, technology shocks have different effects on hours worked, depending on whether hours are taken in levels or differences, whereas the treatment of hours does not affect SDFM estimation.

Concerning monetary policy, we find that the "internal instrument" method solves VAR invertibility problems, but only when the instrument fulfills the so called "conditional lead-lag exogeneity" property. If this property does not hold, the VAR model fails dramatically, whereas, again, the DFM model performs well. This helps us understanding the empirical results of Section 5, where internal IV identification produces puzzling impulse-response functions.

For our simulation exercises we use the medium-scale DSGE of Justiniano et al. (2010). The model has seven shocks and it is equipped with all the frictions that are considered necessary to capture the persistence of macro data: habit persistence, adjustment costs to investment, sticky prices, sticky wages, etc. We use the same specification and parameterization in Justiniano et al. (2010).

To estimate the factor model, we need a large cross section. We build our large dataset as follows.

1. Apply the Kalman smoother to the DSGE using as observables the most recent versions of the 7 data series used in the estimation of the DSGE in Justiniano et al. (2010), namely GDP, investment, consumption, wages, hours, interest rate and inflation and obtain smoothed estimates of all DSGE variables, states and controls.

2. Project the 229 series of our actual large N dataset (the data) on the smoothed states of the DSGE and collect the projection matrix, M. The matrix M captures the empirical relationship between the states of the DSGE and the data and it is used to generate the simulated large N dataset.[12]

The simulations are constructed as follows. From the DSGE model we first generate all states and stationary endogenous variables. The seven shocks are Gaussian i.i.d. with variance equal to the variance resulting from the estimates of the DSGE model in Justiniano et al. (2010). For each simulation, the large dataset is produced multiplying the matrix M just described by the simulated states. A measurement error of the appropriate size is added to all the series, where needed.

We use a very large number of time observations, $T = 5000$. We need a large sample to show that VAR models may produce asymptotically biased and therefore inconsistent estimates (whereas SDFM estimates are asymptotically correct). Using a number of observations comparable to those of actual data sets we could only show that there is a small sample bias. In the Appendix we report small sample simulations using $T = 235$. The basic result is that the DFM still performs better than the VAR, except that the differences are not as stark as with $T = 5000$. In each simulation we generate 100 artificial datasets.

[12] This is the same approach used in Boivin and Giannoni (2006) and Gelfer (2019) to match large cross-sections to DSGE models. In their approach, observed variables are linear and contemporaneous functions of the states.

3.1 Technology shocks with TFP

We first focus on the standard technology shock. As in Justiniano et al. (2010), $\log TFP_t$, denoted a_t, follows

$$a_t = a_{t-1} + T_t \tag{15}$$

$$T_t = \rho T_{t-1} + \eta_t \tag{16}$$

where T_t is the growth rate of technology, driven by the surprise shock η_t. We employ a simple bivariate VAR which includes the growth rates of TFP and the log of hours worked. Identification of the technology shock, along the lines of Galì (1999), is obtained using the model-based restriction that the only shock affecting TFP in the long run[13] is technology. We normalize the size of the shock by imposing that the estimated and the theoretical responses of TFP at horizon 25 are equal.

We start by analyzing the performance of the SVAR and the SDFM in absence of measurement error.[14]

The left column of Figure 1 reports the average of the estimated IRFs for the SDFM and the VAR (red with circles and blue with crosses lines, respectively) along with the 16th and 84th percentiles (dashed for SDFM, dotted for VAR) and the theoretical IRFs (black solid lines). The figure shows that both models perform very well in estimating the impulse responses. The result for the VAR is particularly interesting since it is a case where, even if the model is noninvertible because there are only two variables and seven shocks, still the impulse response functions of the technology shock can be correctly estimated. The result is in line with the δ_i associated to the technology shocks which is equal to zero. Next, we add a Gaussian i.i.d measurement error ξ_{it} to

[13]Long run here is computed using impulse responses 500 steps ahead.

[14]In all estimates in the main text, the SDFM is estimated with $r = 17$ and $q = 7$ and both the SDFM and the VAR are estimated with 4 lags, which is the standard number of lags used in empirical works with quarterly data. In the small sample analysis in the Appendix we use information criteria to select all the parameters of the two models. In the robustness section, online Appendix B, we show results with one lag, the number of lags in the VAR for the states in the DSGE solution.

each stationary variable χ_{it} of the DSGE model and build $x_{it} = \chi_{it} + \xi_{it}$. We then transform and cumulate the series entering the VAR in levels. The measurement error is therefore $I(0)$ for $I(0)$ variables and $I(1)$ for $I(1)$ variables. We scale the size of the variance of the measurement error by setting the ratio $k_i = \frac{Var(\chi_{it})}{Var(x_{it})} = 0.9$, so that the variance of the measurement error is 10% of the variance for each variable.

To get a first sense of how the introduction of the measurement errors affect information deficiency, we compute again the δ_i. Now the δ_i associated to the technology shock increases from 0 to 0.10. All in all, the effect of measurement error appears to be relatively moderate and the VAR is expected to perform reasonably well even in this case. Figure 1, right column, confirms our prior: measurement error alone does not contaminate substantially the performance of the two empirical models (yet the SDFM turns out to marginally improve upon the VAR in identifying the IRFs of hours after horizon 4).[15]

In the online Appendix B we do a "stress test" on the SDFM by increasing measurement error to the unrealistic figures 20%, 50%, 80% of the variance of each series. The SDFM performs reasonably well even in the worst case, while the VAR does not.

In conclusion, a simple bivariate VAR with TFP and hours is able to correctly estimate the technology shock and the correspopnding IRFs. Adding measurement errors of reasonable size produces modest effects. To better interpret these results consider that TFP is driven only by the technology shock, Δa_t being a simple AR(1) model, see equation (16). Hence TFP alone provides enough information to recover the shock. In other words, the VAR, albeit small, is not affected by informational problems (as far as the technology shock is concerned).

[15]Another way to interpret the results above is to notice that the TFP is the technology shock. Having TFP in the VAR is very similar to ordering it first in a VAR identified à la Choleski. This is the so-called *internal* IV procedure. Plagborg-Møller and Wolf (2021) show that this approach, provided some conditions on the instrument are satisfied (see the discussion in section 3.5), is delivering the correct IRFs up to an horizon equal to the number of lags in the VAR, even in presence of measurement error.

3.2 Technology shocks with labor productivity

In this subsection we replace TFP with labor productivity, defined as $\log(Y/L)$. This will be useful to better understand the empirical results of the following section, where we revisit the debate, originated by Galì (1999), about the effects of technology shocks on hours worked. In Figure 2 we show the impulse response functions obtained with a bivariate VAR on variables contaminated with measurement error (the δ_i for a VAR with 4 lags is 0.12). The left column shows IRFs with hours entering the VAR and the SDFM in levels, the right column with hours entering the models in first differences. We use two different treatments for hours worked in light of the discussion between Galì (1999) and Christiano et al. (2007) on the correct way to treat hours worked to identify a technology shock.

The SVARs with different hours specifications produce different and incorrect responses of hours. The comparison of these results with those of the previous subsection suggests the following interpretation. Unike TFP, labor productivity is driven by other shocks, in addition to the technology shock. Hence the bivariate SVAR with labor productivity and hours worked is informationally deficient, because there are more than two shocks driving these two variables. When hours are taken in differences, over-differencing represents an additional invertibility problem: this is the reason why results depend on the treatment of hours. By contrast, the SDFM captures the correct IRFs even when hours worked are taken in differences. An explanation of this last point and a thorough discussion of all these results is provided in Subsection 3.4 below.

3.3 News shocks

We now modify the model by replacing the standard technology shock with a news technology shock, a shock generating noninvertibilities even without measurement errors. The log-linear version of TFP, a_t, is modeled as

$$a_t = a_{t-1} + N_t \tag{17}$$

$$N_t = \phi N_{t-1} + \epsilon_{t-4} \tag{18}$$

where $\phi = 0.8$ and $V(\epsilon_t) = 0.04$. N_t is the news component of TFP and ϵ_t is the news shock with four periods of anticipation. In this model, the value of TFP observed up to time t does not provide information abut the current shock, but only on past shocks until time $t-4$. Hence the econometrician suffers from an informational mismatch.

We first estimate a VAR with TFP and hours. The news shock is identified as the only shock with a long run impact on TFP. We study large samples ($T = 5000$) without measurement error. The δ_i for the news shock is 0.99, so we expect to see that the VAR is unable to correctly capture the IRFs and the shock. Results are in the first column of figure 3: the IRFs from the VAR are far away from the true ones. On the contrary the factor model performs extremely well.

In this case informational deficiency is caused by the fact that TFP does not convey timely information about the shock, and is not necessarily solved by introducing additional series in the VAR. To show this, we focus on a VAR with 5 variables, similar to the one used by Barsky and Sims (2011). The variables are TFP, hours, GDP, consumption and investment. The value of δ_i is 0.986. Figure 3, right column, displays the impulse response functions for TFP and hours. Again, the VAR is unable to estimate the IRFs, whereas the SDFM can.

3.4 Interpreting the results: an illustrative example

The following stylized example may help understanding the above results. Let us assume that productivity changes, Δa_t, and hours, h_t, follow the model

$$\begin{pmatrix} \Delta a_t \\ h_t \end{pmatrix} = \begin{pmatrix} c + aL & f & b \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} s_t \\ d_t \\ e_t \end{pmatrix}, \tag{19}$$

where a, b, c and f are parameters, s_t is a technology shock, d_t is a demand shock and e_t can be interpreted either as an additional shock or a measurement error. Notice that $h_t = s_t + d_t$ is already stationary without differencing.

Let us consider first the SVAR of Subsection 3.1, the one with TFP and hours worked, which works reasonably well. The corresponding model is obtained by setting $f = b = 0$ and $|c/a| > 1$. Since $b = 0$, the third column of the matrix can be ignored, together with the measurement error e_t. Moreover, since $f = 0$, productivity is driven only by the technology shock. The determinant of the relevant 2×2 submatrix is $c + aL$, which vanishes for $L = -c/a$. Since $|c/a| > 1$, the model is invertible and the SVAR is successful. We have shown that a small value of b does not change things substantially.

Now let us consider the case in which $f = b = 0$, as above, but $c = 0$. In this case the technology shock has a delayed effect on productivity; in other words, it is a news shock, just like the one of Subsection 3.3. The determinant vanishes in 0, so that the model is not invertible and the SVAR fails, even if there are no measurement errors.

Finally, let us turn to the general case $f \neq 0$ and $b \neq 0$. In this case the model is "short", i.e. we have more shocks than variables, so that the two variables together cannot convey enough information to recover s_t. This is the case of Subsection 3.2, where TFP is replaced by labor productivity and the SVAR does not work properly.

What happens if h_t is taken in first differences? The model becomes

$$\begin{pmatrix} \Delta a_t \\ \Delta h_t \end{pmatrix} = \begin{pmatrix} c + aL & f & b \\ 1 - L & 1 - L & 0 \end{pmatrix} \begin{pmatrix} s_t \\ d_t \\ e_t \end{pmatrix}, \tag{20}$$

Even assuming $b = 0$, we have a square model where the determinant vanishes for $L = 1$, since hours are over-differenced. Hence we have an additional invertibility problem. This is the reason why the VAR with h_t and the one with Δh_t deliver different results (both incorrect).

To get an intuition of the reason why the SDFM works well in all cases, assume that we have two factors, F_{1t} and F_{2t}, following equation (20), where however $b = 0$, since the factors are free of measurement errors. In addition, we have a third factor F_{3t}, conveying additional information on s_t and d_t. For example, let $F_{3t} = d_t$. In this case we have the system

$$\begin{pmatrix} F_{1t} \\ F_{2t} \\ F_{3t} \end{pmatrix} = \begin{pmatrix} c + aL & f \\ 1 - L & 1 - L \\ 0 & 1 \end{pmatrix} \begin{pmatrix} s_t \\ d_t \end{pmatrix}. \tag{21}$$

The situation is now reversed. The system is "tall", i.e. we have more variables than shocks. Intuition suggests that in this case we have enough information to find s_t. Indeed, tall systems always admit a finite order VAR representation, provided that the MA matrix is zeroless.[16] In our case the matrix is zeroless irrespective of the values taken on by a, c and f, so that the system is invertible. For instance, assuming

[16] Zeroless means that the column rank is maximal for any complex value of L, see Anderson and Deistler (2008).

for simplicity $f = 0$ and letting $\alpha = a/(c+a)$, $\beta = a^2/(c+a)$, $\gamma = 1/(c+a)$, the factors have the VAR(2) representation

$$\begin{pmatrix} 1-\alpha L & -\beta L & \beta L - \beta L^2 \\ \gamma L & 1+\alpha L & c\gamma L + \alpha L^2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} F_{1t} \\ F_{2t} \\ F_{3t} \end{pmatrix} = \begin{pmatrix} c & 0 \\ 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} s_t \\ d_t \end{pmatrix}$$

Hence s_t can be obtained from current and past values of the factors, even if F_{2t} is overdifferenced and even in the case of s_t being a news shocks, i.e. $c = 0$.[17]

On the other hand, if $F_{2t} = h_t = s_t + d_t$, the structural representation of the factors becomes

$$\begin{pmatrix} F_{1t} \\ F_{2t} \\ F_{3t} \end{pmatrix} = \begin{pmatrix} c+aL & f \\ 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} s_t \\ d_t \end{pmatrix}.$$

In this case the factors have the simple VAR(1) representation

$$\begin{pmatrix} 1 & -aL & aL \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} F_{1t} \\ F_{2t} \\ F_{3t} \end{pmatrix} = \begin{pmatrix} c & f \\ 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} s_t \\ d_t \end{pmatrix}.$$

In conclusion, the structural factor model is informationally sufficient for the shock, and therefore is able to deliver the correct impulse response functions, irrespectively of the values of a, c and f and the factors being overdifferenced or not.

3.5 Monetary policy shocks

We now focus on the performance of the SVAR and the SDFM in estimating the impulse response functions of the monetary policy shock. We specify the VAR with

[17] Notice that in the last case F_{1t} is an exact linear combination of the past of F_{2t} and F_{3t}.

4 variables, the interest rate, GDP, inflation and GDP in the flexible price allocation. These variables are the only variables entering the DSGE's Taylor rule. This implies that the four variables are informationally sufficient for the monetary policy shock, being the shock, by construction, a combination of these four variables. Indeed, the δ_i for the monetary shock in absence of measurement error is zero.

Identification of the monetary shock is achieved using an external proxy SVAR scheme. As shown by Stock and Watson (2018) and Mertens and Ravn (2013), if the VAR is *globally invertible*, that is, all the shocks display $\delta_i = 0$, and an instrument z_t for the monetary policy shock satisfying the two conditions discussed below is available, then it is possible to identify the IRFs to the monetary policy shock. The two conditions are: (a) *relevance*: the instrument must be contemporaneously correlated with the shock of interest, (b) *contemporaneous exogeneity*: the instrument must be contemporaneously *un*correlated with all the other shocks.

However, the 4-variables VAR above is not globally invertible: not all the shocks are linear combinations of the present and the past of the VAR variables.[18] Therefore the external proxy SVAR with an instrument satisfying only (a) and (b) does not work. Miranda-Agrippino and Ricco (2021) reports the validity condition of the external proxy estimator when only $m < n$ shocks are invertible. The instrument must fulfill (a), (b) and (c) *limited lead-lag exogeneity*, i.e. the instrument is uncorrelated at all leads and lags with the shocks that are not invertible.

We construct an artificial instrument z_t for the monetary policy shock $u_{mp,t}$ satisfying conditions (a), (b) and (c). Precisely

$$z_t = \alpha u_{mp,t} + v_t, \tag{22}$$

[18] It suffices to note that 4, the number of variables in the VAR, is less than 7, the number of structural shocks in the DSGE.

where $\alpha = 0.5$, $\text{var}(u_{mp,t}) = 0.048$, v_t being a Gaussian zero-mean process, independent of the structural shocks (at all leads and lags), with variance equal to 20% of the variance of $u_{mp,t}$. The vector of impact effects is obtained by regressing the VAR residuals onto the instrument. The dynamics is obtained by post-multiplying the reduced-form Moving Average obtained from the VAR by the vector of the impact effects. We normalize the IRFs by imposing that the impact effect on the interest rate is equal to the one of the DSGE model.

We first analyze the estimates in a world with no measurement error. The first column of Figure 4 displays the estimated impulse response functions of GDP, inflation and the interest rate. As expected, the responses are perfectly captured. The SDFM delivers the correct impulse responses simply because the VAR on the factors F_t is *globally invertible*, see equation (10): property (c) is not needed in the factor model. The importance of this will be clear in the next experiment.

Next, we add a 10% measurement errors to all variables but the interest rate.[19] The presence of measurement error, unlike for the technology shock, raises the δ_i substantially, from 0 to 0.73.[20] The second column of Figure 4 reports the responses of the three variables. The IRFs obtained from the VAR are clearly off, with a large real activity puzzle (top panel). On the contrary, the factor model performs very well delivering the correct response for each of the three variables.

Unlike Subsection 3.1, here the presence of measurement errors generates a large and empirically problematic deficiency in an otherwise informationally sufficient VAR. Why the consequence of measurement errors are so different in the two cases? Our explanation is the following. In the DSGE model the technology shock is very much similar to the TFP growth process, the coefficient ρ in (16) being small (0.23 in JPT).

[19] We do not add a measurement error to the interest rate since it can be argued that the interest rate is not affected by measurement errors.

[20] As the monetary policy shock now is not invertible, the instrument defined in (22) is not valid anymore in a VAR. The instrument is still valid in a SDFM, as the VAR on F_t is globally invertible.

Hence the shock is estimated with an error which is essentially the same error affecting TFP. By contrast, the dynamics of the monetary shock is important, involving several regressors and lags with large coefficients. The measurement errors affecting all regressors can therefore produce sizable biases in the estimates of the parameters and the shock itself.

Plagborg-Møller and Wolf (2021) propose to include the instrument into the VAR, ordered first, impose a Cholesky identification scheme and take the impulse response functions of the first shock, normalized according to their impact effect on a given variable. We refer to this procedure as the "internal proxy" procedure. This method works even if the shock of interest is not invertible, therefore even in the presence of measurement errors, provided that the instrument fulfills suitable conditions.

This procedure requires (a) *relevance*: the instrument must be contemporaneously correlated with the shock of interest, (b) *contemporaneous exogeneity*: the instrument must be contemporaneously uncorrelated with all the other shocks, and (d) *conditional lead-lag exogeneity*: the instrument must be uncorrelated with all shocks at all leads and lags, controlling for the lags of the VAR variables.

Note that conditions (a), (b) and (d) are true for the instrument of equation (22). The third column of Figure 4 shows the results obtained with the "internal instrument" procedure (blue lines with crosses). The factor model estimates (red lines with circles) are the same of the other columns and are obtained with the standard external instrument method. Now the SVAR works much better (even though the confidence bands for GDP are very large and the estimates are somewhat biased for horizons larger than the number of lags in the VAR).

As stated above, the "internal proxy" method requires the lead-lag exogeneity condition (d), which is not required when we have global invertibility, as is the case with the SDFM. Stock and Watson (2018) convincingly argue that conditions (a), (b) and (d) are much more demanding than conditions (a) and (b) alone. To see what happens when condition (d) is violated, we construct a different proxy, fulfilling (a) and (b), but not (d):

$$\tilde{z}_t = \alpha u_{mp,t} + \beta \hat{w}_{t-1} + v_t \tag{23}$$

where $\alpha = \beta = 0.5$, v_t is as in (22) and $\hat{w}_t$ is real wages, as generated by the DSGE.[21]

Figure 5 shows the results obtained using $\tilde{z}_t$ in place of z_t (the Data Generating Process being the same as above). The IRFs obtained with the *external proxy* SVAR method are shown in the left column (blue lines with crosses) and those obtained with the *internal proxy* SVAR method are shown in the right column (blue lines with crosses). The SDFM estimates are obtained with the external proxy method and are the same in the two columns (red lines with circles). The figure shows that the SDFM does very well, with the red lines perfectly overlapping the true IRFs (black solid lines), whereas the SVAR does dramatically bad, independently of the method used. In sum, even in the case of proxy identification, because of the global invertibility property of the SDFM, less stringent assumptions are needed on the instrumental variable.

3.6 Shocks estimates

So far, we have focused on estimation of the IRFs. However, in some cases it is useful to estimate the shock itself (mainly for historical decomposition). Hence, for some of the experiments studied in the previous subsections, we estimate the structural shocks. Table 1 shows the average correlations of our shock estimates with the true shocks, together with the 5-th and the 95-th percentiles of the distribution of the correlations.

For the technology shock, consistently with the IRF results, the shocks are very well

[21] We just selected one variable among those of the DSGE model. Any DSGE variable not present in the VAR would do the same job.

	Technology (TFP)		News (bivariate)		Monetary	
No error	SDFM	0.96 (0.9 0.99)	SDFM	0.98 (0.95 0.98)	SDFM	0.99 (0.99 0.99)
	VAR	1.00 (0.98 1.00)	VAR	0.01 (-0.02 0.03)	VAR	1.00 (1.00 1.00)
Error	SDFM	0.95 (0.88 0.98)	SDFM	- (-)	SDFM	0.97 (0.97 0.97)
	VAR	0.94 (0.92 0.94)	VAR	- (-)	VAR	0.50 (0.48 0.51)

Table 1: Correlation among actual (simulated) and estimated shocks. The top figure in each box is the average across 100 simulations. In parenthesis, the $5th$ and $95th$ percentiles. The VAR to identify the technology shock is estimated with TFP and hours in levels. The monetary VAR is estimated using the instrument in equation (22).

estimated by both models with and without measurement errors. For the news shock (second column) without measurement error, the SVAR fails, the average correlation being 0.01, as against 0.98 of the SDFM. Finally, for the monetary policy shock (third column), both models perform well without measurement error, but with error the SVAR performs very poorly, the correlation with the true shock being 0.50, as against 0.97 of the SDFM.

In summary, this section shows that (a) both measurement errors and delayed effects on the TFP can produce sizable invertibility problems, which are reflected in bad estimates of the shock and the IRFs of interest obtained with SVAR models; (b) different transformations of the variables significantly affect the estimates of IRFs in SVAR models; (c) the SDFM is not affected by these problems; (d) with a valid instrument, the internal instrument SVAR approach can perform well even under noninvertibility; however, the validity conditions for the instrument are more demanding than those required in the SDFM.

4 Empirics: Technology shocks and hours worked

In this empirical analysis we revisit the long-standing but never resolved empirical controversy about the effects of technology shocks on hours worked. Galì (1999), using a bivariate VAR with labor productivity and hours worked, finds that the shock driving productivity in the long run reduces hours worked. This result is relevant for discriminating between alternative theories, as it is at odds with real business cycle theory, which predicts an increase in hours, and in line with the standard New Keynesian model. Several papers, among which Basu et al. (2006) and Francis and Ramey (2005), confirm Galì's finding. In contrast, using alternative VAR specifications and identification strategies, other papers, among which Christiano et al. (2003, 2004), Dedola and Neri (2007) and Peersman and Straub (2009), argue that the empirical evidence on the effect of a productivity shock on hours worked is not very robust and could be consistent with a positive effect on hours worked. In particular, Christiano et al. (2003), CEV henceforth, show that, if hours worked are taken in levels, rather than in first differences, as in Galì (1999), technology shocks increase hours.

Here we show that small VARs are affected by noninvertibility problems and produce results that depend dramatically on the treatment and choice of the variables. By contrast, SDFM results are robust: hours increase (after a nearly zero impact effect) independently of the treatment. These findings are in line with the simulation results of Subsection 3.1.

4.1 SVAR results

To begin with, we estimate Galì's VAR and CEV's VAR with our sample, spanning from 1960:II to 2019:II. As for the measure of productivity, we use labor productivity, as in both Galì 's and CEV's VARs.[22] We estimate two bivariate VAR(4) specifications with labor productivity and per capita hours worked: in the former (VAR Ia) hours are taken in log-differences; in the latter (VAR Ib), hours are taken in log levels.[23] In both cases the technology shock is identified as the only shock that has long-run effects on labor productivity. Figure 6 shows the results. The black solid lines are the point estimates of the IRFs, the shaded areas are the 68% confidence intervals.

The controversy is confirmed: hours reduce when specified in growth rates (left panels) and increase when are taken in levels (right panels).

Since small VARs can in principle be affected by noninvertibility, it is natural to ask whether more information-rich VARs provide different results, possibly consistent with different treatments of hours worked. The choice of the variables, however, is far from obvious. Here we show results for three reasonable specifications.

In the first one we include productivity, hours, unemployment, inflation and the federal funds rate, with hours in differences (VAR IIa) and in levels (VARIIb). In the second one we include productivity, hours, GDP, consumption and investment, again with hours in differences (VAR IIIa) and in levels (VARIIIb). In the third one we include productivity and hours, plus three financial variables providing information about the term spread and the risk premium: the federal funds rate, the 10-years Treasury bond rate and the Moody's BAA rate (VAR IVa and VAR IVb). In VAR III, trending real activity variables are taken in levels to avoid possible cointegration problems. We use 4 lags for all models.

The point estimates are plotted in Figure 7, together with the bivariate VAR, which is our benchmark. The red lines, corresponding to specification IV, are roughly similar

[22]Using the (utilization adjusted) Total Factor Productivity developed by Fernald (2012) gives very similar results.

[23]Labor productivity is taken in log-differences, as in both Galì 's and CEV's VARs. The hours series is the HOANS series, relative to the nonfarm business sector, divided by civilian noninstitutional population aged 16 years or more, taken from the Fred database. The log is multiplied by 100 to get percentage variations for the IRFs. Labor productivity is the OPHNFB series, FRED data base.

to the baseline VAR. Specifications II and III (blue and green lines, in order) produce IRFs for hours which are positive and roughly consistent across different treatments of hours, but the effects estimated with specification II are larger, particularly when hours are taken in differences. In summary, results differ substantially across specifications, especially if hours are taken in first differences, and it is hard to tell which is the most reliable.

4.2 Assessing the empirical evidence

In this subsection we analyze whether the VAR specifications above contain enough information to estimate the technology shock. To this end, we perform the invertibility test suggested by Forni and Gambetti (2014). This test consists in regressing the estimated shock onto the lags of the principal components of a large data set and performing a standard F-test for the joint significance of the regressors. If the null is rejected, the shock is predicted by the principal components and informational sufficiency is rejected, i.e., the VAR does not contain enough information to estimate the structural shock of interest. To compute the principal components, we take the the quarterly FRED-QD data set and apply the transformation described in the next subsection.

Table 2 reports the p-values of the F-test obtained when regressing the estimated technology shock onto the lags of the first seven principal components of our data set (the one with hours in differences). The basic finding is that the null of invertibility is rejected at the 5% level for the shocks of both VAR Ia (Galì) and VAR Ib (CEV). Our conclusion is that neither of the bivariate VAR specifications contain enough information to recover the technology shock, so that the estimated IRFs are not reliable.

	1 lag	2 lags	3 lags	4 lags
VAR Ia	0.060	0.021	0.035	0.026
VAR Ib	0.000	0.000	0.000	0.002
VAR IIa	0.112	0.175	0.405	0.389
VAR IIb	0.051	0.215	0.335	0.447
VAR IIIa	0.011	0.056	0.192	0.173
VARIIIb	0.016	0.080	0.270	0.165
VAR IVa	0.258	0.079	0.053	0.087
VAR IVb	0.355	0.318	0.404	0.188

Table 2: p values of the F-test for the significance of the regression of the estimated VAR shocks onto the lags of the first 7 principal components of our large macroeconomic data set (hours in differences).

Regarding the other VAR specification, the result is mixed. VAR IIIa and IIIb are rejected at the 5% level; VAR IIb and IVa are rejected at the 10% level. VAR IIa (Figure 7, left panels, blue lines) and VAR IVb (Figure 7, right panels, red lines) are not rejected and with both specifications hours increase after a technology shock.

4.3 SDFM results

Let us now consider the results obtained with the SDFM model. Our analysis is based on the quarterly FRED-QD data set of McCracken and Ng (2021), extended to the most recent data. To get a balanced panel, we retained only series spanning from 1960:II to 2019:I, ending up with 229 series. We transformed each series to reach stationarity. For hours worked, we performed two different transformations: differences (DFMa) and levels (DFMb).[24] Then we estimated two factor models. In both cases we used $r = 13$ and three lags of the factors.[25] The productivity shock is identified, as in the VAR models above, as the only shock affecting productivity at the 10-year

[24] We applied these transformation to two hours worked series, per capita HOANBS and per-capita HOABS.

[25] The number of factors is set by using the log-criterion of Alessi et al. (2010); the number of lags is set according to the AIC. For simplicity, in the estimation we skipped the rank reduction step, by setting the number of shocks equal to the number of factors.

horizon.

Figure 8 shows the IRFs obtained with the factor model. The basic finding is that now results are robust to the treatment of hours, the IRFs of both TFP and hours being similar in the two columns. The shocks estimated with the two data treatments are almost identical, the correlation coefficient being 0.99.

According to the SDFM, hours increase after a productivity shock and the shape of the IRF is similar to that of VAR Ib (CEV). Does this mean that CEV's VAR is in line with the factor model? Not quite. In fact, we have seen that the test of Table 2 rejects VAR Ib. Firstly, the IRF of productivity is different. According to VAR Ib, the IRF is almost flat. According to the SDFM, the dynamics of productivity is given by an initial jump, followed by a slow diffusion process. Secondly, the shock estimated with the SDFM is very different from the one of VAR Ib, the correlation coefficient being just 0.26.

5 Empirics: Monetary policy shocks

In the last decade, instrumental variable identification has become the most popular method to study the effects of monetary policy. In this section we identify the monetary policy shock and estimate its impulse response functions by using the authoritative instrument proposed by Gertler and Karadi (2015). We compare the IRFs obtained with different SVAR specifications with those obtained with the SDFM. The basic result is that SVAR models can deliver puzzling results, even if we use the internal instrument method discussed in Plagborg-Møller and Wolf (2021). By contrast, SDFM results are in line with the theory. Our findings can be explained in the light of the simulations in Section 3.

5.1 SVAR results: external instrument

In this subsection we present results from different VAR specifications, using the standard, external IV method. Precisely, we estimate a VAR (VAR I) including the Industrial Production index and the CPI index, plus an interest rate; following Gertler and Karadi (2015), we use the 1-year Treasury Bond rate. The frequency is monthly. Then we estimate a VAR (VAR II) including the three variabls discussed above plus the spread between the 10-year Treasury Bond rate and the Moody's BAA bond rate, which can be regarded as a measure of the risk premium.[26] Finally, we estimate a VAR (VAR III) including all variables in VAR II, plus the S&P500 stock price index. The IP index, the CPI index and the stock market index are taken in log levels. The time span is 1960:1 — 2008:12, so that the zero lower bound period is excluded. For these three VAR specifications we identify the monetary policy shock by the external IV method; precisely, we identify the impact effects by regressing the VAR residuals onto the instrument. The IRFs are then normalized to get an impact effect of 100 basis points on the interest rate. All VAR models include 12 lags.

Figure 9 shows the results. With VAR I we have the price puzzle; moreover, we have a real activity puzzle, since industrial production increases significantly on impact following a contractionary shock. In the second line, reporting results for VAR II, both puzzles disappear; however, the monetary shock does not affect significantly real activity.[27] In the third line (VAR III) we have results in line with economic theory: both industrial production and prices reduce significantly after a contractionary shock.

[26]The Excess Bond Premium used by Gertler and Karadi (2015) is not available for our time span.

[27]The difference with respect to Gertler and Karadi (2015) is due to the different time span and the fact that we use for the risk premium a different indicator, since the ECB is not available until 1973.

5.2 Testing for invertibility

We conjecture that the results of the previous subsection are driven by information: the first two VAR specifications do not contain enough information to recover the monetary shock. To verify this, we perform the invertibility test of the previous section. We regress the estimated shock onto the lags of the first 8 principal components of our large monthly data set, described below. Results are reported in Table 3.

	1 lag	2 lags	3 lags	4 lags	5 lags	6 lags
VAR I	0.002	0.000	0.000	0.000	0.000	0.000
VAR II	0.001	0.000	0.001	0.000	0.000	0.000
VAR III	0.702	0.554	0.240	0.182	0.271	0.444

Table 3: p-values of the F-test for the significance of the regression of the estimated VAR shocks onto the lags of the first 8 principal components of the FRED-MD monthly macroeconomic data set.

As expected, the first two VAR specifications are deficient, as invertibility is rejected at the 1% level. By contrast, invertibility of the third VAR specification, including stock prices, is not rejected.

5.3 SVAR results: internal instrument

Plagborg-Møller and Wolf (2021) observe that the internal instrument method works even if we do not have invertibility, provided that the instrument satisfies the lead-lag exogeneity condition. In the simulation section we have shown that, if such condition is violated, the internal instrument method can fail.

In the present subsection we use the internal instrument method applied to VAR I. We include the instrument into the VAR, ordered first, and identify the monetary policy shock as the first one in the Cholesky scheme. Again, we use 12 lags in the VAR.[28]

[28]Notice that necessarily, with the internal instrument method, the time span must coincide with

The estimated impulse response functions are reported in Figure 10, upper line. The price puzzle is solved; however, the real activity puzzle is still there. Moreover, the IRF of the interest rate is very different from that of VAR III. Our explanation is that the instrument does not satisfy the lead-lag exogeneity assumptions, so that the internal instrument method fails.

5.4 SDFM results

Our SDFM analysis is based on the monthly FRED-MD data set of McCracken and Ng (2016). To get a balanced panel, we retained only series spanning from 1960:1 to 2008:12, ending up with 122 series. We transformed each series to reach stationarity. Then we estimated the model with 8 factors and 12 lags.[29] Identification is obtained by the external IV method. Results are shown in Figure 10, bottom line. The estimated IRFs are in line with the theory and similar, both qualitatively and quantitatively, to those obtained with VAR III (Figure 9, bottom line).

Finally, Figure 11 shows a few robustness checks for the SDFM: in the first column, we show the results for a different number of lags ($p = 6$ and $p = 9$); in the second column we change the number of factors ($r = 9$ and $r = 10$); in the third column we use the instrument of Miranda-Agrippino and Ricco (2021); in the fourth column we change the sample span, starting in 1979:8, the beginning of the Volcker mandate. We conclude that results are reasonably robust.

the time span of the instrument, so that, for this exercise, we are forced to use the time span 1990:1—2008:12.

[29]The number of factors is set by using the log-criterion of Alessi et al. (2010).

6 Conclusions

We have argued that Structural Dynamic Factor Models are better suited than Structural VARs to inform and guide the construction of shocks' transmission mechanisms in DSGE models. The reason is that SVARs may be informationally deficient. Our simulations show several cases where SVAR analysis produce inconsistent IRFs, whereas the IRFs estimated with the SDFM are always consistent. Of course, there can be cases where SVARs perform accurately and we have shown some of them. Overall, however, we believe that the SDFM is a safer choice.

In the empirical study, we first revisit the old controversy about the effects of technology shocks on hours worked and, second, we study the effects of a monetary policy shock. Using a bivariate SVAR with productivity and hours, the effect is positive when hours are taken in levels, negative when hours are taken in differences. By using the SDFM, technology shocks, after a nearly zero impact effect, persistently increase hours worked, irrespectively of their treatment (differences or levels). Both SVAR specifications are informationally deficient and produce misleading results.

We then identify a monetary shock using the instrument of Gertler and Karadi (2015), with different VAR specifications. We find that SVAR models may fail to deliver IRFs in line with the consensus economic view, even if we use the internal instrument method. By contrast, the SDFM produces credible results.

References

Alessi, L., M. Barigozzi, and M. Capasso (2010). Improved penalization for determining the number of factors in approximate static factor models. *Statistics and Probability Letters 80*, 1806–1813.

Anderson, B. D. O. and M. Deistler (2008). Generalized linear dynamic factor models–A structure theory. *IEE Conference on Decision and Control*.

Aruoba, B. S., M. Mlikota, F. Schorfheide, and S. Villalvazo (2022). SVARs with occasionally-binding constraints. *Journal of Econometrics 231*(2), 477–499.

Bai, J. and S. Ng (2002). Determining the number of factors in approximate factor models. *Econometrica 70*, 191–221.

Bańbura, M., D. Giannone, and L. Reichlin (2010). Large Bayesian vector auto regressions. *Journal of Applied Econometrics 25*(1), 71–92.

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

Basu, S., J. G. Fernald, and M. S. Kimball (2006). Are technology improvements contractionary? *American Economic Review 96*(5), 1418–1448.

Baumeister, C. and J. D. Hamilton (2019, May). Structural interpretation of vector autoregressions with incomplete identification: revisiting the role of oil supply and demand shocks. *American Economic Review 109*(5), 1873–1910.

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

Boivin, J. and M. P. Giannoni (2006). DSGE models in a data-rich environment. *NBER Technical Paper 0332*.

Canova, F. and M. Paustian (2011). Business cycle measurement with some theory. *Journal of Monetary Economics 58*(4), 345–361.

Christiano, L. J., M. Eichenbaum, and R. Vigfusson (2003). What happens after a technology shock? International Finance Discussion Papers 768, Board of Governors of the Federal Reserve System (U.S.).

Christiano, L. J., M. Eichenbaum, and R. Vigfusson (2004). The response of hours to a technology shock: Evidence based on direct measures of technology. *Journal of the European Economic Association 2*(2-3), 381–395.

Christiano, L. J., M. Eichenbaum, and R. Vigfusson (2007). Assessing structural VARs. In *NBER Macroeconomics Annual 2006, Volume 21*, pp. 1–106.

Consolo, A., C. A. Favero, and A. Paccagnini (2009). On the statistical identification of DSGE models. *Journal of Econometrics 150*, 99–115.

Dedola, L. and S. Neri (2007). What does a technology shock do? A VAR analysis with model-based sign restrictions. *Journal of Monetary Economics 54*, 512–549.

Del Negro, M. and F. Schorfheide (2004). Priors from general equilibrium models for VARs. *International Economic Review 45*(2), 643–673.

Del Negro, M., F. Schorfheide, F. S. Smets, and R. Wouters (2007). On the fit and forecasting performance of New Keynesian models. *Journal of Business & Economic Statistics 25*(2), 123–143.

Fernald, J. (2012). A quarterly, utilization-adjusted series on total factor productivity. *Federal Reserve Bank of San Francisco WP Series 2012-19*.

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*, 1021–1026.

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

Forni, M., L. Gambetti, M. Lippi, and L. Sala (2020). Common components structural VARs. *CEPR Discussion Paper No. 15529*.

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

Forni, M., L. Gambetti, and L. Sala (2019). Structural VARs and noninvertible macroeconomic models. *Journal of Applied Econometrics 34*(2), 221–246.

Forni, M., D. Giannone, M. Lippi, and L. Reichlin (2009). Opening the black box: Structural factor models versus structural VARs. *Econometric Theory 25*, 1319–1347.

Forni, M., M. Hallin, M. Lippi, and L. Reichlin (2000). The Generalized Dynamic Factor Model: identification and estimation. *The Review of Economics and Statistics 82*, 540–554.

Francis, N. and V. A. Ramey (2005). Is the technology-driven real business cycle hypothesis dead? Shocks and aggregate fluctuations revisited. *Journal of Monetary Economics 52*, 1379–1399.

Galì, J. (1999). Technology, employment, and the business cycle: Do technology shocks explain aggregate fluctuations? *The American Economic Review 199*, 249–271.

Gelfer, S. (2019). Data-rich DSGE model forecasts of the great recession and its recovery. *Review of Economic Dynamics 32*, 18–41.

Gertler, M. and P. Karadi (2015). Monetary policy surprises, credit costs, and economic activity. *American Economic Journal: Macroeconomics 7*(1), 44–76.

Giannone, D., M. Lenza, and G. E. Primiceri (2015). Prior selection for vector autoregressions. *The Review of Economics and Statistics 97*(2), 436–451.

Guerron-Quintana, P., A. Inoue, and L. Kilian (2017). Impulse response matching estimators for DSGE models. *Journal of Econometrics 196*(1), 144–155.

Hallin, M. and R. Liška (2007). Determining the number of factors in the general dynamic factor model. *Journal of the American Statistical Association 102*, 603–617.

Justiniano, A., G. E. Primiceri, and A. Tambalotti (2010). Investment shocks and business cycles. *Journal of Monetary Economics 57*(2), 132–145.

Leeper, E. M., T. B. Walker, and S. S. Yang (2013). Fiscal foresight and information flows. *Econometrica 81*, 1115–1145.

Lippi, M. (2020). Validating DSGE models with SVARs and High-Dimensional Dynamic Factor Models. Working paper. Einaudi Institute for Economics and Finance, Rome.

Lippi, M. and L. Reichlin (1993). The dynamic effects of aggregate demand and supply disturbances: Comment. *American Economic Review 83*, 644–652.

Lippi, M. and L. Reichlin (1994). Common and uncommon trends and cycles. *European Economic Review 38*, 624–635.

McCracken, M. and S. Ng (2016). FRED-MD: a monthly database for macroeconomic research. *Journal of Business and Economic Statistics*, 574–589.

McCracken, M. W. and S. Ng (2021, January). FRED-QD: A Quarterly Database for Macroeconomic Research. *Review 103*(1), 1–44.

Mertens, K. and M. O. Ravn (2013). The dynamic effects of personal and corporate income tax changes in the united states. *American Economic Review 103*(4), 1212–47.

Miranda-Agrippino, S. and G. Ricco (2021). The transmission of monetary policy shocks. *American Economic Journal: Macroeconomics 13*(3), 74–107.

Peersman, G. and R. Straub (2009). Technology shocks and robust sign restrictions in a euro area SVAR. *International Economic Review 50*(3), 727–750.

Plagborg-Møller, M. and C. K. Wolf (2021). Local projections and VARs estimate the same impulse responses. *Econometrica 89*, 955–980.

Rozanov, Y. A. (1967). *Stationary Random Processes.* San Francisco: Holden Day.

Sims, C. A. and T. Zha (2006). Does monetary policy generate recessions? *Macroeconomic Dynamics 10*, 231–272.

Smith, A. A. J. (1993). Estimating nonlinear time-series models using simulated vector autoregressions. , *Journal of Applied Econometrics 8*, 63–84.

Stock, J. H. and M. W. Watson (2002a). Forecasting using principal components from a large number of predictors. *Journal of the American Statistical Association 97*, 1167–1179.

Stock, J. H. and M. W. Watson (2002b). Macroeconomic forecasting using diffusion indexes. *Journal of Business and Economic Statistics 20*, 147–162.

Stock, J. H. and M. W. Watson (2018). Identification and estimation of dynamic causal effects in macroeconomics using external instruments. *The Economic Journal 128*(610), 917–948.

Appendix: small sample results

In this Appendix we study what happens in small samples. The sample size is equal to $T = 235$, matching the number of observations in the data. In all the experiments, a 10% measurement error has been added to all variables, including those entering the simulated large dataset used in the DFM estimation, but the interest rate. The number of factors r is selected using the criterion proposed by Alessi et al. (2010) and the number of shocks q using the criterion by Hallin and Liška (2007). The number of lags in both the VAR and the SDFM is selected by the BIC criterion.

Figure 12 displays the IRFs in response to a technology shock. The factor model behaves extremely well, with bands clearly wider than those in the large sample case, while the VAR has problems in identifying the impulse response of hours after horizon 4 and the confidence bands for the response of hours are very large.

Figure 13 displays the IRFs in response to a news shock. The SDFM have problems to correctly capture the 4 lags of delay of the shock. This is due to the limited number of estimated factors suggested by the information criterion used. Aside from that, IRFs are correctly captured. The VAR, once again, displays in small samples the same biases shown in large samples and very large confidence bands.

Figure 14, left column, displays the IRFs in response to a monetary shock with an external instrument (the instrument being generated by equation (22)). The factor model captures correctly all impulse responses, with bands clearly wider than those in the large sample case. The VAR on the other side is unable to capture the response of all variables, even with very large confidence bands.

The right column shows the IRFs in response to a monetary shock with internal instrument (again with the instrument being generated by equation (22)). Concerning the VAR, because of the presence of few lags selected by the BIC in the estimated model, impulse responses are in line with the true ones only up to horizon 1. After that horizon, biases are evident.

All in all, these results show that the SDFM does better than the VAR, except that the differences are not as stark as with $T = 5000$.

Figures

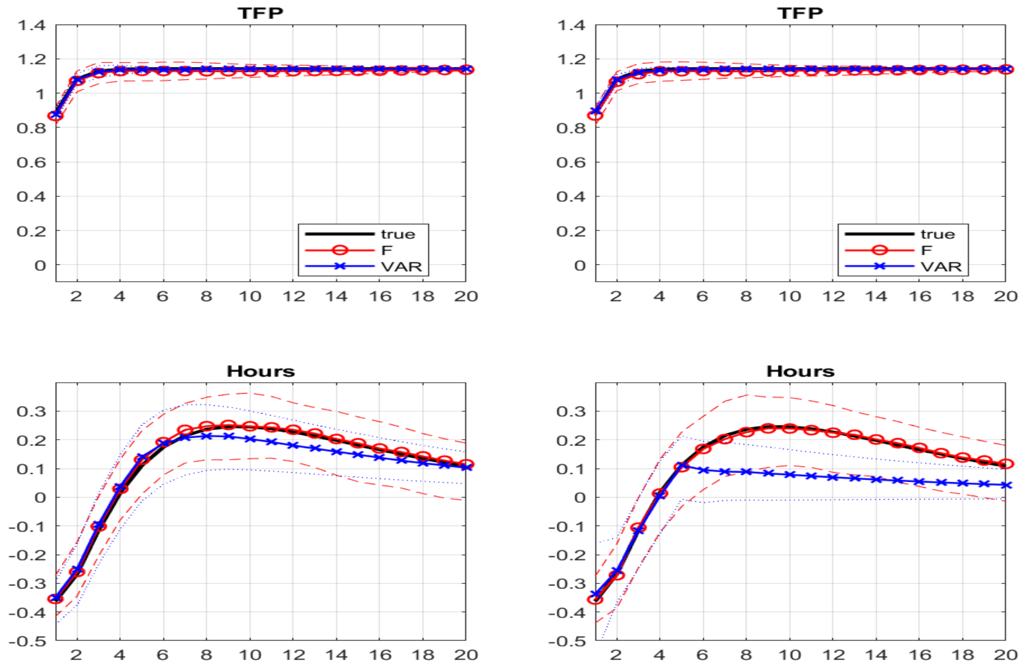

Figure 1: Identifying a technology shock - Bold line: true response. Circles: SDFM, $r = 17$, $q = 7$, $p_F = 4$. Crosses: VAR, $p_{VAR} = 4$. Dashed lines: 16-th and 84-th percentiles of the distribution of 100 simulations, SDFM. Dotted lines: 16-th and 84-th percentiles of the distribution of 100 simulations, VAR. $T = 5000$. Left: no measurement error. Right: measurement error ($k_i = 0.9$).

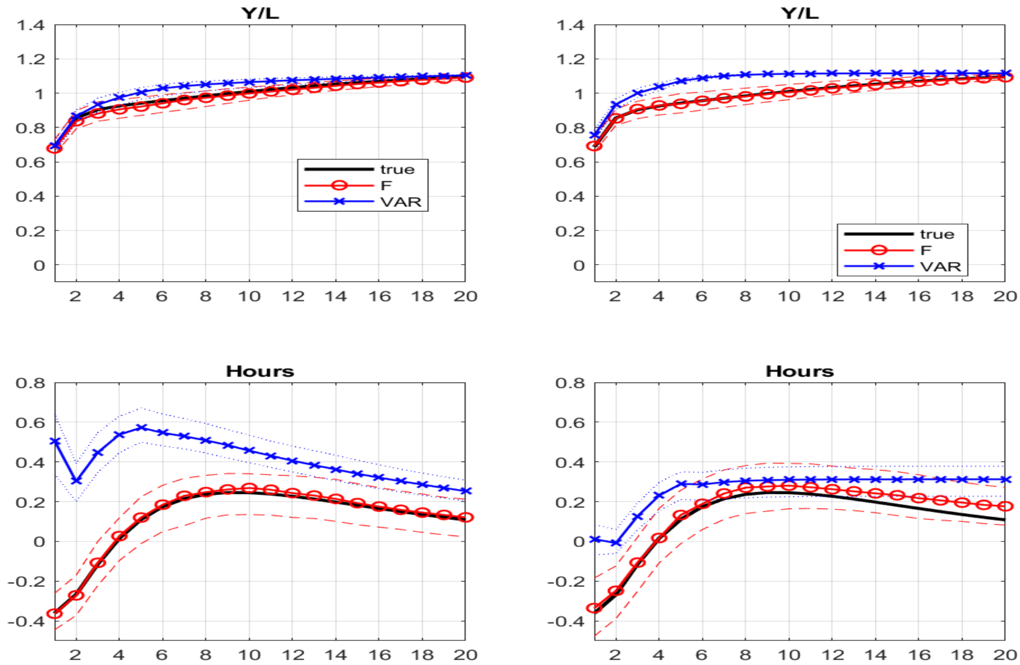

Figure 2: Identifying a technology shock - VAR with $\log(Y/L)$. Bold line: true response. Circles: SDFM, $r = 17$, $q = 7$, $p_F = 4$. Crosses: VAR, $p_{VAR} = 4$. Dashed lines: 16-th and 84-th percentiles of the distribution of 100 simulations, SDFM. Dotted lines: 16-th and 84-th percentiles of the distribution of 100 simulations, VAR. $T = 5000$. Left: measurement error ($k_i = 0.9$), hours in levels. Right: measurement error ($k_i = 0.9$), hours in first differences.

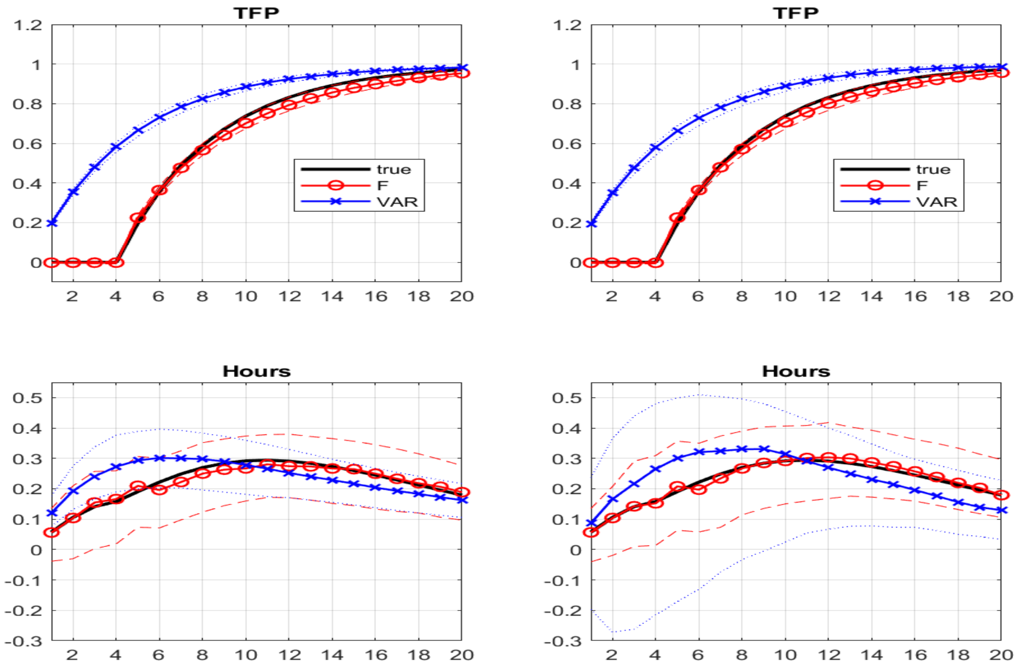

Figure 3: Identifying a news shock - Bold line: true response. Circles: SDFM, $r = 17$, $q = 7$, $p_F = 4$. Crosses: VAR, $p_{VAR} = 4$. Dashed lines: 68% percentiles of the distribution of 100 simulations, SDFM. Dotted lines: 68% percentiles of the distribution of 100 simulations, VAR. $T = 5000$. No measurement error. Left: two variables. Right: five variables.

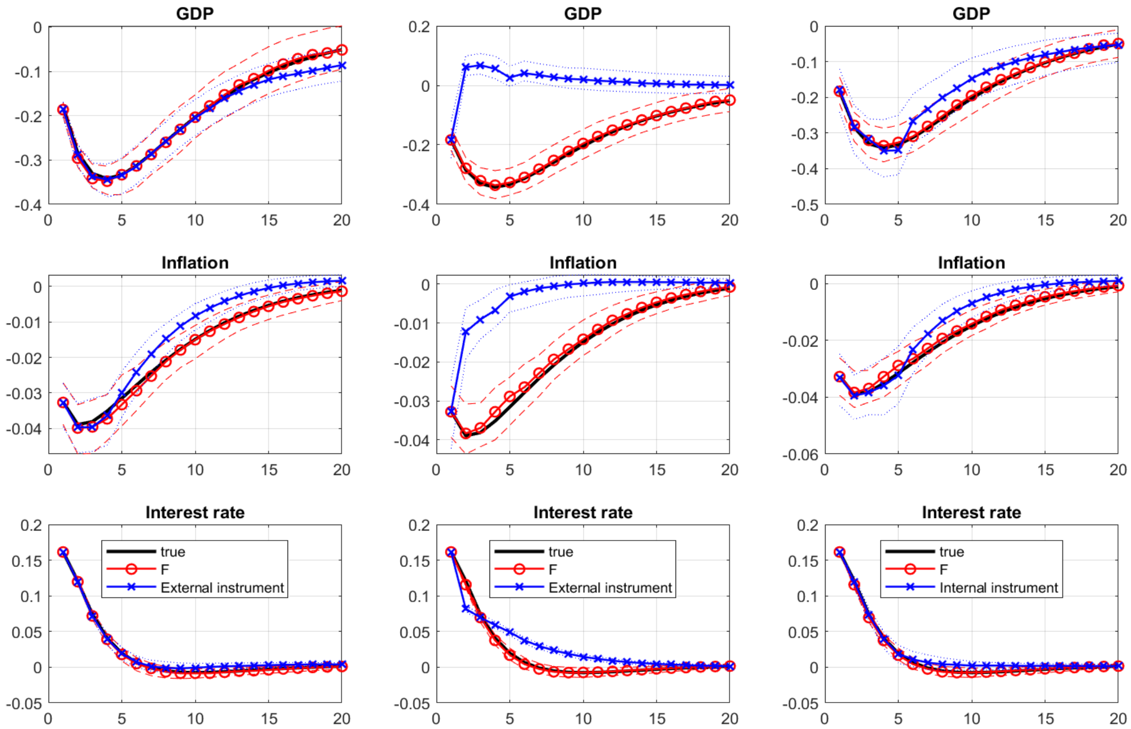

Figure 4: Identifying monetary shocks. Bold line: true response. Circles: SDFM, $r = 17$, $q = 7$, $p_F = 4$. Crosses: VAR, $p_{VAR} = 4$. Dashed lines: 68% percentiles of the distribution of 100 simulations, SDFM. Dotted lines: 68% percentiles of the distribution of 100 simulations, VAR. $T = 5000$. Left column: No measurement error, external instrument. Center column: measurement error ($k_i = 0.9$), external instrument. Right column: measurement error ($k_i = 0.9$), internal instrument. Instrument z_t generated as $z_t = \alpha u_{mp,t} + v_t$.

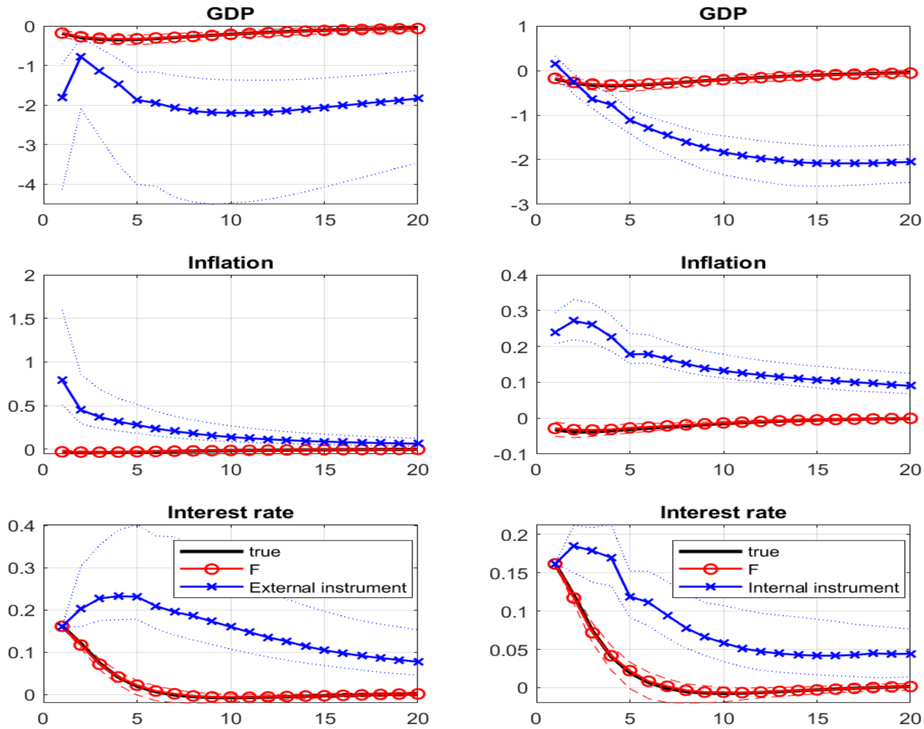

Figure 5: Identifying monetary shocks. Bold line: true response. Circles: SDFM, $r = 17$, $q = 7$, $p_F = 4$. Crosses: VAR, $p_{VAR} = 4$. Dashed lines: 68% percentiles of the distribution of 100 simulations, SDFM. Dotted lines: 68% percentiles of the distribution of 100 simulations, VAR. $T = 5000$. Left column: measurement error ($k_i = 0.9$), external instrument. Right column: measurement error ($k_i = 0.9$), internal instrument. Instrument generated as: $\tilde{z}_t = \alpha u_{mp,t} + \beta \hat{w}_{t-1} + v_t$.

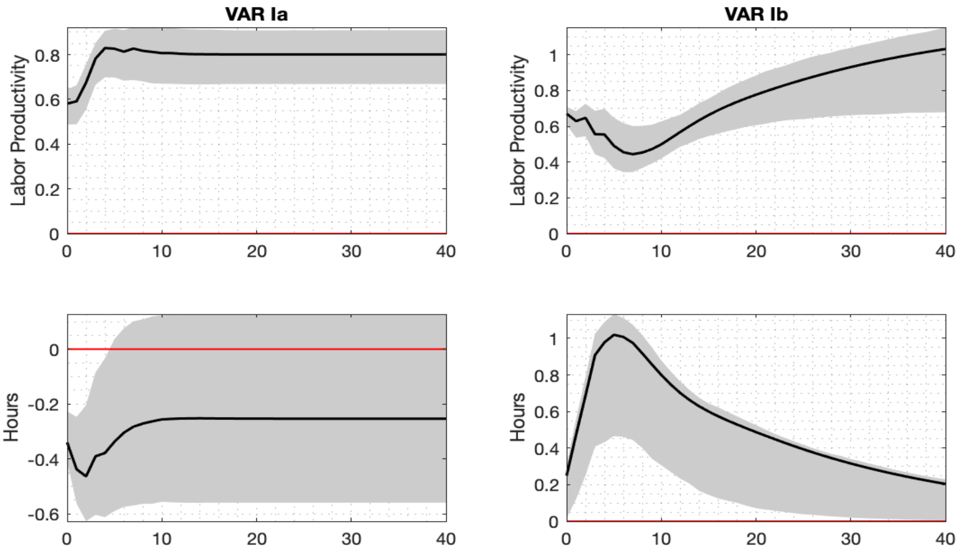

Figure 6: Impulse response functions to a productivity shock, identified as the only shock affecting productivity at the 10-year horizon. First column: VAR(4) with labor productivity and per-capita hours worked in log-differencies. Second column: VAR(4) with labor productivity and per-capita hours worked in log-levels. Black line: point estimate. Shaded areas: 68% confidence bands.

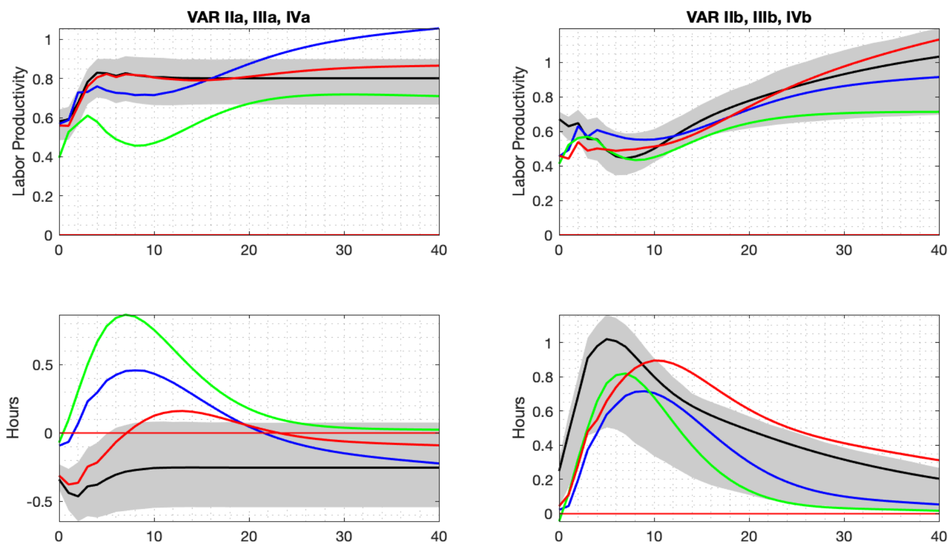

Figure 7: Impulse response functions to a productivity shock, identified as the only shock affecting productivity at the 10-year horizon. First column: per-capita hours worked in log-differencies. Second column: per-capita hours worked in log-levels. Black line: point estimate of the benchmark bivariate VAR I. Blue lines: point estimates of VAR II. Green lines: point estimates of VAR III. Red lines: point estimates of VAR IV. Shaded area: 68% confidence bands for the bivariate VAR I.

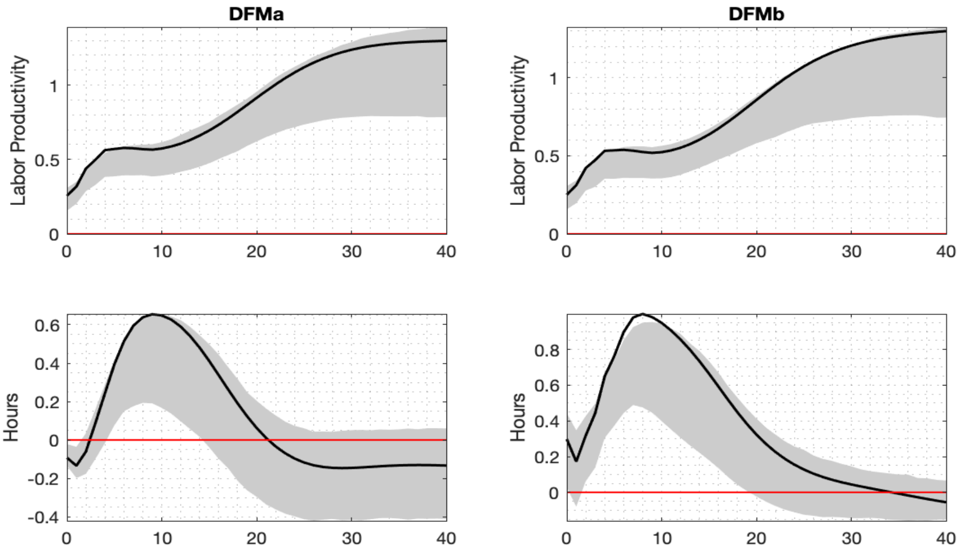

Figure 8: Impulse response functions to a productivity shock, identified as the only shock affecting labor productivity at the 10-year horizon. First column: SDFM(3) with per-capita hours worked in log-differences. Second column: SDFM(3) with per-capita hours worked in log-levels. Black line: point estimate. Shaded areas: 68% confidence bands.

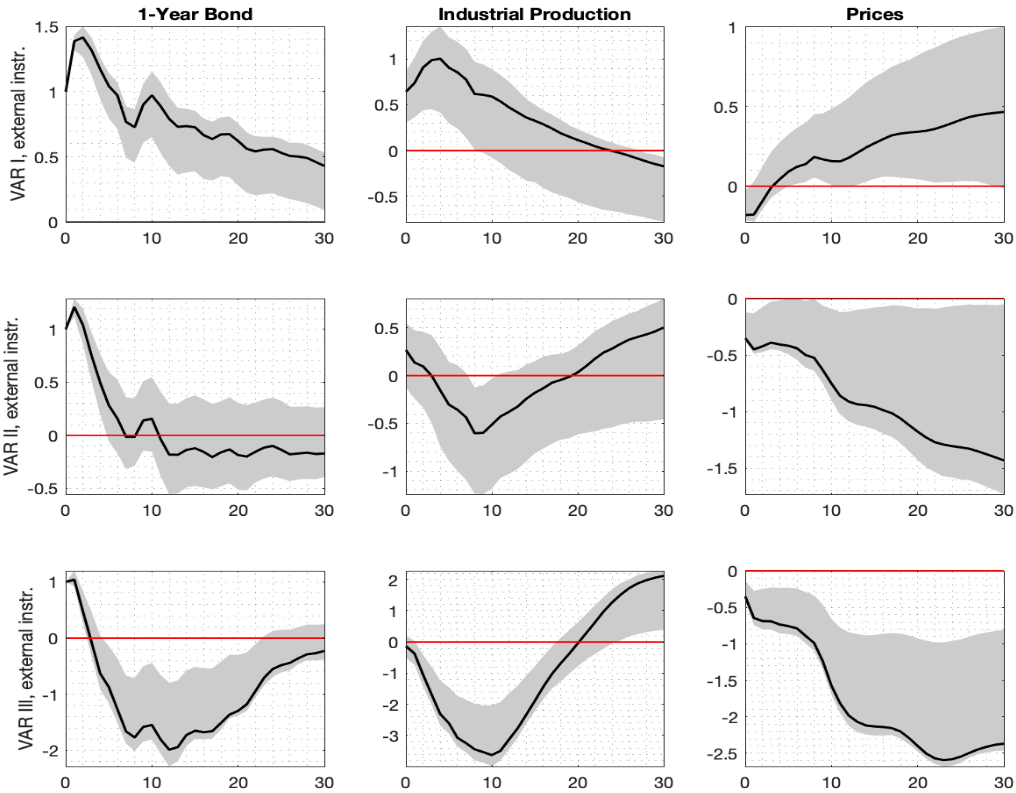

Figure 9: Identifying a monetary policy shock. First row: external IV, VAR I. Second row: external IV, VAR II. Third row: external IV, VAR III.

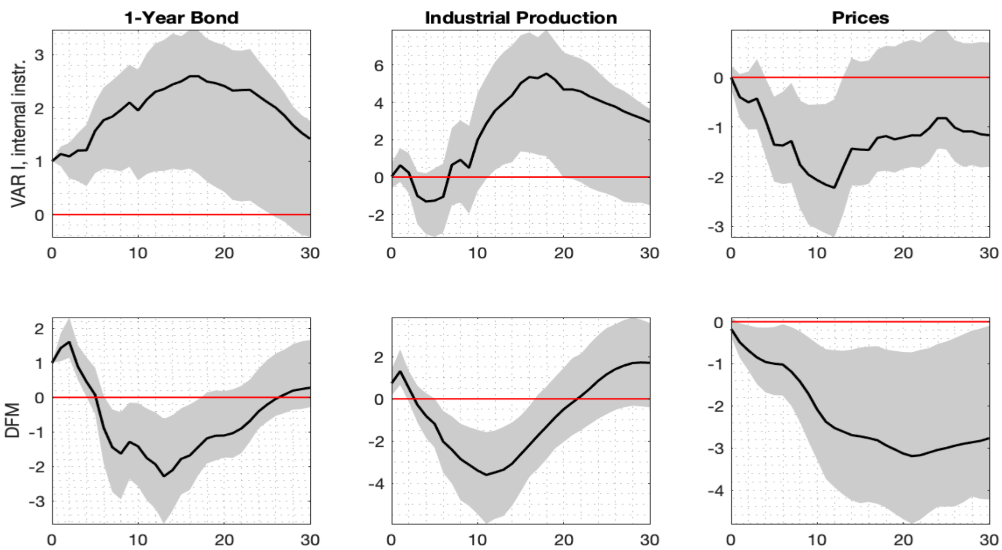

Figure 10: Identifying a monetary policy shock. First row: internal IV, VAR I. Second row: SDFM.

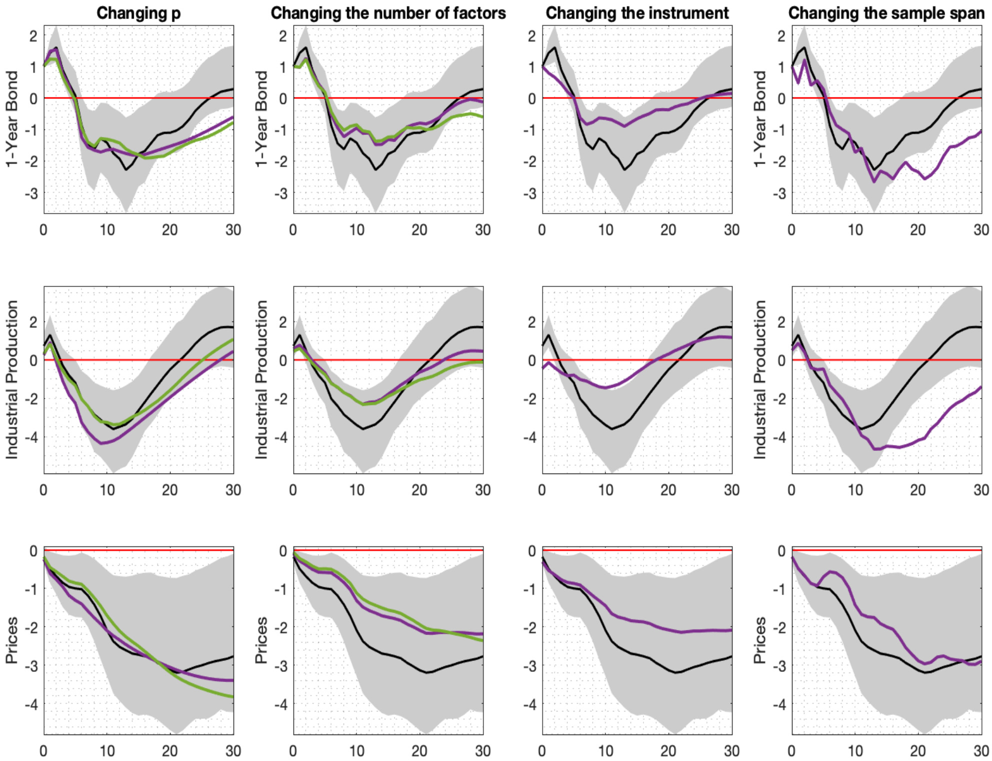

Figure 11: Robustness checks for DFM results. First column: $p = 6$ (purple line), $p = 9$ (green line). Second column: $r = 9$ (purple line), $r = 10$ (green line). Third column: instrument of Miranda-Agrippino and Ricco (2021) (purple line). Fourth column: time span 1979:8—2008:12 (violet line).

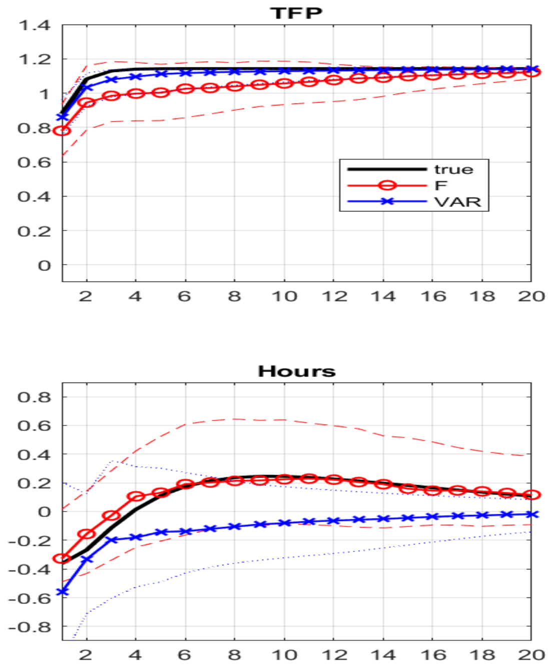

Figure 12: Identifying a technology shock - Circles: factor model. Crosses: VAR. Bold line: true response. Dashed lines: 68% percentiles of the distribution of 100 simulations, factor. Dotted lines: 68% percentiles of the distribution of 100 simulations, VAR. $T = 235$. Measurement error: $k_i = 0.9$.

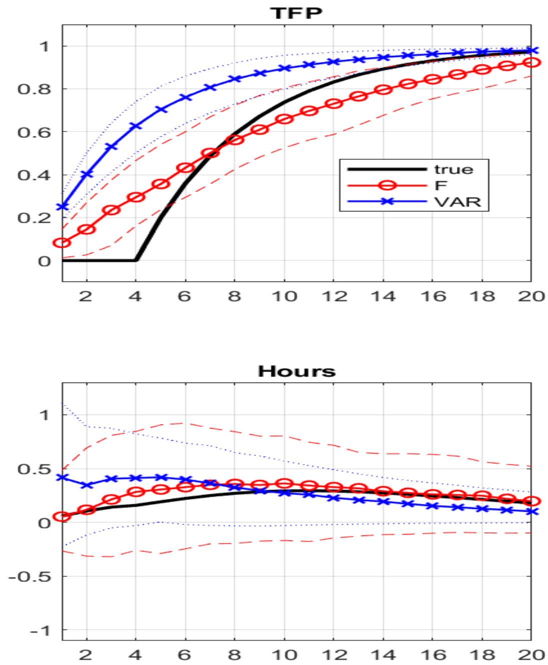

Figure 13: Identifying a news shock - Circles: factor model. Crosses: VAR. Bold line: true response. Dashed lines: 68% percentiles of the distribution of 100 simulations, factor. Dotted lines: 68% percentiles of the distribution of 100 simulations, VAR. $T = 235$. Measurement error: $k_i = 0.9$.

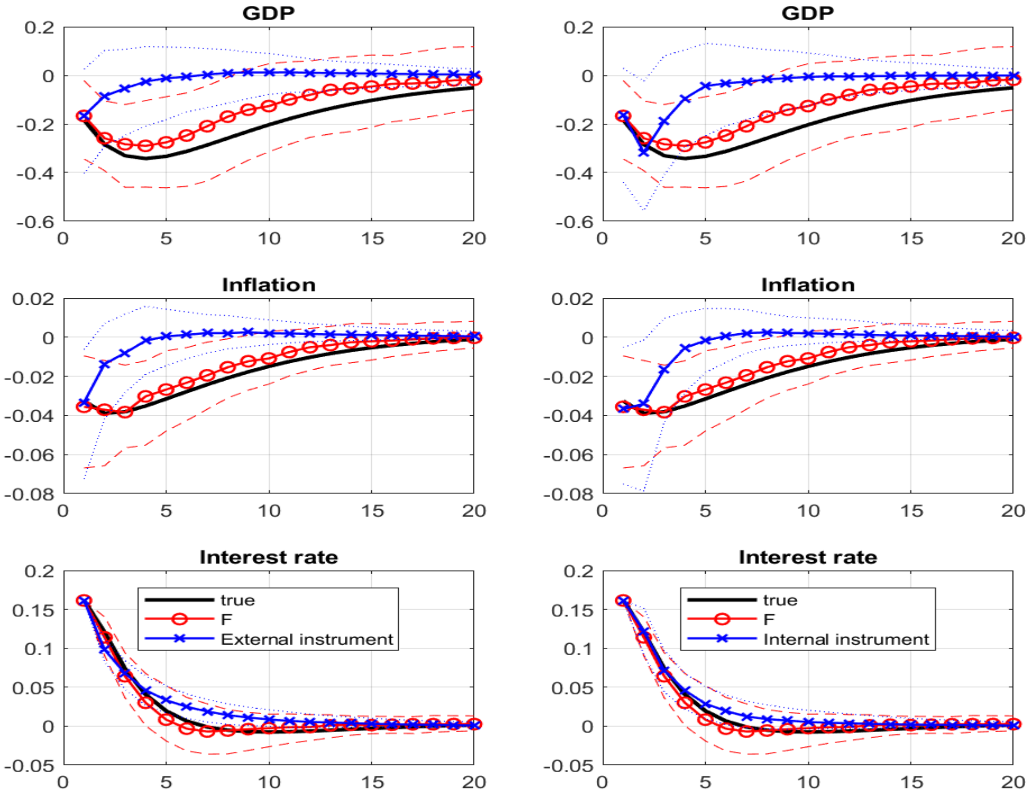

Figure 14: Identifying a monetary shock - Circles: factor model. Crosses: VAR. Bold line: true response. Dashed lines: 68% percentiles of the distribution of 100 simulations, factor. Dotted lines: 68% percentiles of the distribution of 100 simulations, VAR. $T = 235$. Measurement error: $k_i = 0.9$.

Online Appendix

A. Identification and estimation of the factor model

Although Λ and F_t are not unique, under the model assumptions, $\chi_t = \Lambda F_t = \Lambda Q(L)^{-1}\varepsilon_t$ is unique. To get an MA representation with q orthonormal shocks we observe that the covariance matrix of the VAR residuals can be represented as $\Sigma_\varepsilon = VMV'$, where M is the $q \times q$ diagonal matrix having on the diagonal the q non-zero eigenvalues of Σ_ε and V is the $r \times q$ matrix having on the columns the corresponding eigenvectors. Then, defining $W = VM^{-1/2}$, $v_t = W'\varepsilon_t$ and $R = VM^{1/2}$ we get the representation

$$\chi_t = \Lambda Q(L)^{-1} R v_t. \tag{24}$$

The above representation is a fundamental MA representation with orthonormal shocks. Starting from the above representation we can get the structural shocks as $u_t = H'v_t$, where H is a $q \times q$ orthogonal matrix (see Rozanov (1967), pp. 56-7; see also Section 3.2 in Forni et al. (2009)). The corresponding matrix of impact effects S is obtained as $S = RH$. Lastly, the determination of the matrix H can be obtained by imposing restrictions on the impulse response functions of the χ's, i.e.

$$\Phi(L) = \Lambda Q(L)^{-1} RH,$$

such as zero impact or long-run effects, just in the same way as in standard VAR analysis. Of course, we can identify a single shock along with its impulse response function by limiting ourselves to determining just a single column of the matrix H.

Coming to estimation, we first transform the variables to get a stationary vector x_t and estimate the number of static factors to get $\hat{r}$. Then we estimate the static factors themselves by means of the first $\hat{r}$ ordinary principal components of the x's, and the factor loadings by means of the associated eigenvectors. Precisely, let $\hat{\Sigma}_x$ be the sample variance-covariance matrix of x_t: our estimated loading matrix $\hat{\Lambda}$ is the $n \times r$ matrix having on the columns the normalized eigenvectors corresponding to the first largest $\hat{r}$ eigenvalues of $\hat{\Sigma}_x$, and our estimated factors are $\hat{F}_t = \hat{\Lambda}' x_t$.[30]

Second, we set a number of lags $\hat{p}$ and run a VAR($\hat{p}$) with $\hat{F}_t$ to get estimates of $Q(L)$ and the residuals ε_t, say $\hat{Q}(L)$ and $\hat{\varepsilon}_t$.

As a third step, having an estimate $\hat{q}$ of the number of dynamic factors, we obtain an estimate of the non-structural representation (24). Let $\hat{\Sigma}_\varepsilon$ be the sample covariance matrix of $\hat{\varepsilon}_t$. We first estimate the matrices V and M, by computing the largest $\hat{q}$ eigenvalues and the corresponding eigenvectors of $\hat{\Sigma}_\varepsilon$; then we compute $\hat{R} = \hat{V}\hat{M}^{1/2}$ and $\hat{v}_t = \hat{M}^{-1/2}\hat{V}'\hat{\varepsilon}$. When performing partial identification, this "rank reduction" step can be skipped (Forni et al., 2020).

Finally, we obtain an estimate of H by imposing suitable identification restrictions on the estimated impulse response functions. The estimates of the structural impulse response functions is given by $\hat{\Phi}(L) = \hat{\lambda}\hat{Q}(L)^{-1}\hat{S}$, where $\hat{S} = \hat{R}\hat{H}$. The structural shocks are estimated as $\hat{u}_t = \hat{H}'\hat{v}_t$.

For the consistency of this estimation procedure see Forni et al. (2009), Proposition 3 and (Forni et al., 2020), Propositions 1-3.

To get the confidence bands, we bootstrap the estimated VAR residuals $\hat{\varepsilon}_t$ and use $\hat{\Lambda}$ and $\hat{A}(L)$, along with the initial conditions $\hat{F}_1, \ldots, \hat{F}_p$, to construct the artificial series χ_t^1. Then we add the estimated idiosyncratic components $\hat{\xi}_t = x_t - \hat{\chi}_t$ to get the artificial series $x_t^1 = \chi_t^1 + \hat{\xi}_t$ and estimate the model to get the IRFs $\hat{\Phi}^1(L)$. We repeat the procedure m times to get $\hat{\Phi}^j(L)$, $j = 1, \ldots, m$. Finally we take suitable percentiles of the IRFs distribution for each horizon.[31]

[30] Notice that the factors F_t and the loadings Λ are not identified, since given any non-singular $r \times r$ matrix M, we have $\chi_t = \Lambda^* F_t^*$, where $\Lambda^* = \Lambda M^{-1}$ and $F_t^* = MF_t$. Hence, strictly speaking, we do not estimate F_t and Λ but a basis of the space spanned by F_t and the corresponding factor loading matrix. This however is not a problem in the present context since we are only interested in the product $\chi_t = \Lambda F_t$, which is identified.

[31] Estimation of the DFM entails the estimation of a VAR for the factors, which are stationary. One might wonder if this does not lead to cointegration problems. The answer is no. The reason is that in the DFM the spectral density matrix of the factors is singular. For singular vector variables, unlike the standard non-singular case, (a) I(1) variables (in our case the cumulated sum of the factors) are always cointegrated; despite this, (b) a (finite order) VAR for I(0) variables (in our case the factors) does exist, except very special cases. For a broader discussion see Forni et al. (2020).

B. Robustness

We show impulse responses by changing the size of the measurement error and the number of lags in the VAR and the SDFM. In Figure 1 we change the size of measurement errors from 20%, left column, to 50%, center column, to 80%, right column. The factor model performs surprisingly well, even if the volatility of the estimates widens as measurement error increases. The VAR, on the other side, has problems in capturing the responses.

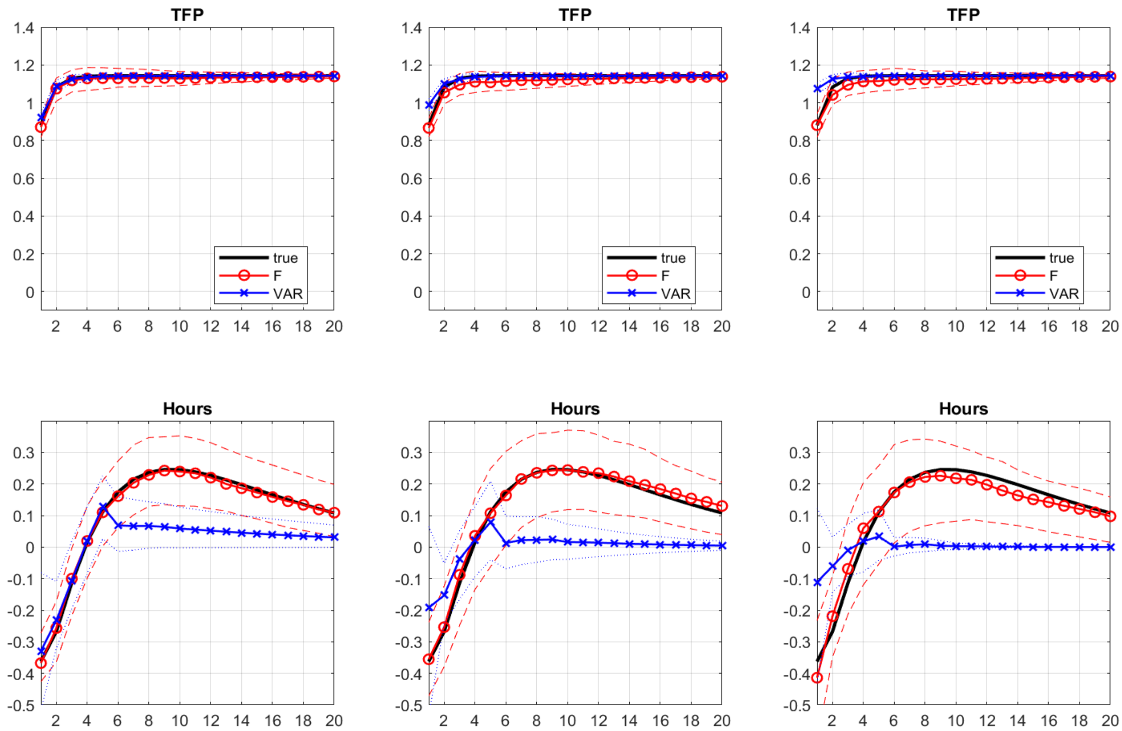

Figure 1: Identifying technology shocks. Bold line: true response. Circles: SDFM, $r = 17$, $q = 7$, $p_F = 4$. Crosses: VAR, $p_{VAR} = 4$. Dashed lines: 68% percentiles of the distribution of 100 simulations, SDFM. Dotted lines: 68% percentiles of the distribution of 100 simulations, VAR. $T = 5000$. Left column: measurement error ($k_i = 0.8$). Center column: measurement error ($k_i = 0.5$). Right column: measurement error ($k_i = 0.2$).

The VAR for the states in the DSGE solution has one lag. As long as the factors span the space of the states, the SDFM should perform well even with one lag as opposed to 4. In Figure 2 we show results with $p_F = p_{VAR} = 1$. The SDFM performs very well, while impulse responses estimated with the VAR are in line with the true ones up to horizon 1.[32]

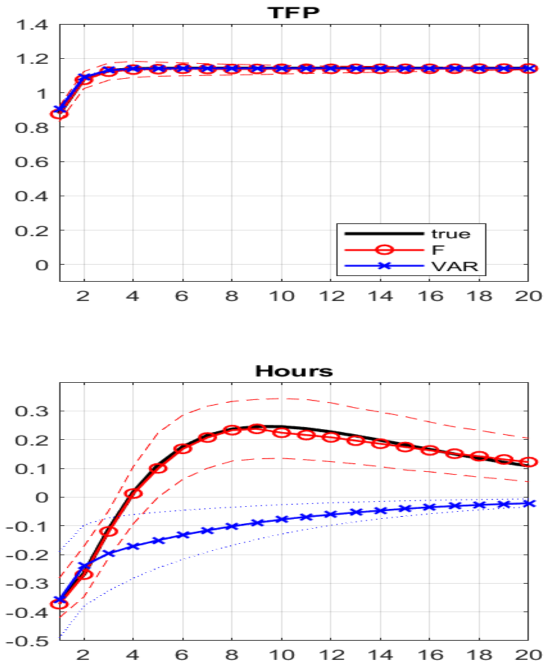

Figure 2: Identifying technology shocks. Bold line: true response. Circles: SDFM, $r = 17$, $q = 7$, $p_F = 1$. Crosses: VAR, $p_{VAR} = 1$. Dashed lines: 68% percentiles of the distribution of 100 simulations, SDFM. Dotted lines: 68% percentiles of the distribution of 100 simulations, VAR. $T = 5000$. Measurement error ($k_i = 0.9$).

[32]We thank a referee for suggesting the exercises in this Appendix.

RECent Working Papers Series

The 10 most RECent releases are:

No. 159	COMMON COMPONENTS STRUCTURAL VARS M. Forni, Luca Gambetti, M. Lippi, L. Sala
No. 158	MATH EXPOSURE AND UNIVERSITY PERFORMANCE: CAUSAL EVIDENCE FROM TWINS G. Bertocchi, L. Bonacini, M. Joxhe, G. Pignataro
No. 157	FAMILY PLANNING AND ETHNIC HERITAGE: EVIDENCE FROM SUB-SAHARAN AFRICA G. Bertocchi, A. Dimico, C. Falco
No. 156	TWO MAIN BUSINESS CYCLE SHOCKS ARE BETTER THAN ONE (2024) A. Granese
No. 155	TEMPERATURE AND GROWTH: A PANEL MIXED FREQUENCY VAR ANALYSIS USING NUTS2 DATA (2023) A. Cipollini, F. Parla
No. 154	NATURAL DISASTERS AND PREFERENCES FOR THE ENVIRONMENT: EVIDENCE FROM THE IMPRESSIONABLE YEARS (2022) C. Falco, R. Corbi
No. 153	DOES WAR MAKE STATES? MILITARY SPENDING AND THE ITALIAN STATE BUILDING, 1861-1945 (2022) A. Incerpi, B. Pistoresi, F. Salsano
No. 152	STRANGERS AND FOREIGNERS: TRUST AND ATTITUDES TOWARD CITIZENSHIP (2022) G. Bertocchi, A. Dimico, G. L. Tedeschi
No. 151	ADAMS AND EVES: THE GENDER GAP IN ECONOMICS MAJORS (2021) G. Bertocchi, L. Bonacini, M. Murat
No. 150	CHOOSE THE SCHOOL, CHOOSE THE PERFORMANCE. NEW EVIDENCE ON THE DETERMINANTS OF STUDENT PERFORMANCE IN EIGHT EUROPEAN COUNTRIES (2021) L. Bonacini, I. Brunetti, G. Gallo

The full list of available working papers, together with their electronic versions, can be found on the RECent website: http://www.recent.unimore.it/site/home/publications/working-papers.html