WHAT DRIVES THE RECENT SURGE IN INFLATION? THE HISTORICAL DECOMPOSITION ROLLER COASTER*

DRAGO BERGHOLT†, FABIO CANOVA‡, FRANCESCO FURLANETTO§, NICOLÒ MAFFEI-FACCIOLI¶ AND PÅL ULVEDAL‖

This version: January 26, 2025

Abstract: What drives the recent inflation surge? To answer this question, one must decompose inflation fluctuations into the contribution of structural shocks. We document how whimsical an historical shock decomposition can be in standard vector autoregressive (VAR) models. We show that the deterministic component of the VAR is imprecisely estimated, making the shock contributions poorly identified under general conditions. Our preferred approach to solve the problem—the single-unit-root prior—massively shrinks the uncertainty around the estimated deterministic component. Once this uncertainty is taken care of, demand shocks are the main drivers of the inflation surge in the United States, the euro area, and in four small open economies.

Keywords: *Bayesian vector autoregression, deterministic component, single unit root prior, inflation dynamics.*

JEL Classification: *C11, C32, E32.*

*We are grateful to three anonymous referees, the editor of this journal, Dante Amengual, Juan Antolin-Diaz, Jonas Arias, Todd Clark, Olivier Coibion, Juan Dolado, Luca Fosso, Domenico Giannone, Michele Lenza, Francesca Loria, Elmar Mertens, Lorenzo Mori, Giorgio Primiceri, Juan Rubio-Ramirez, Samad Sarferaz, Frank Schorfheide and seminar participants of numerous seminars and conferences for comments and suggestions. This paper should not be reported as representing the views of Norges Bank. The views expressed are those of the authors and do not necessarily reflect those of Norges Bank.

†Norges Bank. P.O. Box 1179 Sentrum, 0107 Oslo, Norway. E-mail: drago.bergholt@norges-bank.no.

‡BI Norwegian Business School, Nydalsveien 37, 0484 Oslo, Norway. Corresponding author. E-mail: fabio.canova@bi.no.

§Norges Bank. P.O. Box 1179 Sentrum, 0107 Oslo, Norway. E-mail: francesco.furlanetto@norges-bank.no.

¶Norges Bank. P.O. Box 1179 Sentrum, 0107 Oslo, Norway. E-mail: nicolo.maffei-faccioli@norges-bank.no.

‖Nord University Business School P.O. Box 2501, 7729 Steinkjer, Norway. E-mail: pal.b.ulvedal@nord.no.

1 INTRODUCTION

MOTIVATION After a long period of low and stable inflation, the outlook has suddenly changed in the aftermath of the COVID pandemic, and most countries have seen inflation rates reach unprecedented levels since the late 1970s. In the US, the GDP deflator started rising at the end of 2020 and peaked at 7.7 percent in 2022:Q2. In the euro area, inflation was in negative territory by the end of 2020 before rising sharply in 2021 and 2022, peaking around 10 percent on an annual basis. In some European countries, such as the Netherlands or Estonia, the annual inflation rate exceeded 15 percent in 2022. Initially, supply disruptions associated with the pandemic-induced sectoral reallocation of economic activity were thought to cause the rise in inflation. However, the emergence of a historically tight labor market (as measured, e.g., by the vacancy-to-unemployment ratio) and the fast recovery of the employment-to-population ratio, as the result of the massive monetary and fiscal stimulus implemented in many countries, shifted attention towards demand factors. Needless to say, disentangling demand and supply factors is crucial, since a proper policy response depends on the nature of the impulses driving the inflation surge.

A natural way to separate demand and supply shocks is to rely on flexible time-series models such as Vector Autoregressions (VAR). These models are routinely used to understand key characteristics of the data and to assess the state of the macroeconomic environment, both retrospectively and prospectively (see Del Negro and Schorfheide (2011),Koop and Korobilis (2013), Carriero et al. (2015), Chan (2022), and citetcrump2021large among others). VAR models decompose the observable data into a deterministic and a stochastic component. The deterministic component is regulated by the VAR parameters and the initial values of the dynamic system. The stochastic component, in contrast, depends on the discounted sum of innovations hitting the system. Equipped with some identification restrictions, one can map the reduced form innovations into various structural shocks, with the most fundamental partition in our case being between shocks to demand and supply, respectively. The resulting *structural* VAR (SVAR) model can be used to decompose observed data changes over time into underlying economic drivers. In particular, such a historical shock decomposition can effectively quantify the relative importance of demand and supply factors during the recent inflation surge.

CONTRIBUTION This paper highlights an important flaw when computing historical shock decompositions. As a motivating example, suppose that a researcher wishes to quantify the role of demand and supply disturbances for inflation during the post-COVID period. She can estimate her preferred VAR model, split the observed inflation into its deterministic and stochastic components and, once certain identification restrictions are used, evaluate the role of supply and demand factors. Our key contribution is to show that there is considerable uncertainty in the splitting exercise. We establish that conditional likelihood-based estimation procedures deliver very dispersed estimates of the deterministic component, and even small perturbations of the estimated parameters of the VAR may result in large changes in the deterministic component, despite having only a small impact on the likelihood. Putting the same concept differently, the deterministic component of the model is poorly identified and poorly estimated. Because the deterministic and stochastic components sum to the observable data, a poorly estimated deterministic com-

ponent translates into very imprecise estimates of the shock path decomposition. To the best of our knowledge, this "excess" dispersion has been largely ignored in the literature. In fact, it is typical to compute historical shock decompositions conditioning on a point estimate of the deterministic component.

In contrast, most empirical SVAR applications report the uncertainty associated with estimated impulse response functions (IRFs). For presentation purposes, it is common to pay attention to a measure of central tendency of the IRFs. For example, Fry and Pagan (2011) select the draw closest to the pointwise median IRF, while Inoue and Kilian (2013) choose the most likely admissible draw within the set satisfying the imposed restrictions. We show that parameter draws exhibiting "good" impulse responses—meaning responses close to the pointwise median or to the most likely draw—may feature extreme deterministic components and thus provide unreasonable historical decompositions. Thus, there is no guarantee that "good" impulse responses are associated with "good" historical shock decompositions. In this sense, history narratives may become whimsical.

We highlight that the excess dispersion problem is related to, but distinct from the overfitting problem discussed in Sims (1996), Sims (2000) and Giannone et al. (2019) and shock decompositions may display excess dispersion even when overfitting is absent. Excess dispersion may occur in VARs featuring a different number of variables and estimated with different sample sizes. It appears both with relatively stable data, such as in our baseline sample (1983-2022) for the United States, and in longer samples subject to time series instabilities (e.g. 1949-2022). Importantly, it arises when estimating the reduced-form parameters. Thus, it is independent of the identification scheme used to transform reduced-form innovations into structural shocks.

After documenting how imprecise an estimated shock decomposition can be using a VAR with US data as illustrative example, and inspecting the data features that make the problem more serious, we turn our attention to potential solutions. We show that alternative prior assumptions often used in empirical work to reduce estimated parameter uncertainty (such as a Normal-Inverse Wishart or a Minnesota-like prior), are unable to solve the excess dispersion problem. Our approach considers a specific prior distribution, the single-unit-root prior, also known as the "dummy initial observations" prior, introduced by Sims (1993). Such a prior effectively shrinks the dispersion of the deterministic components by linking the constant of the VAR to some a-priori steady state measure. The tightness of the prior can be calibrated, if one wishes to do so. We treat it as a hyperparameter and estimate its posterior distribution, using the approach of Giannone et al. (2015). The modal value we obtain is generally small. As a result, a SVAR estimated with the single-unit-root prior features almost no uncertainty around the deterministic component and draws producing similar impulse response functions also produce similar historical decompositions. While the single unit root prior has been used in a few applications, see e.g. Sims and Zha (1998) and Antolin-Diaz et al. (2021), its ability to shrink the uncertainty around the estimated deterministic component and to improve historical narratives has not been highlighted before.

RESULTS We estimate a conventional SVAR endowed with sign restrictions to quantify the drivers of the recent inflation surge. When a diffuse prior is employed, there is substantial uncertainty around estimates of the deterministic component of inflation. In addition, the three draws delivering impulse responses (IRF) closest to the pointwise

median response imply vastly different historical decompositions, despite exhibiting almost identical impulse dynamics: one draw implies a dominant role for supply shocks; the next a dominant role for demand shocks; the third suggests that the two shocks are equally important.

When the single-unit-root prior is used, uncertainty around the deterministic component shrinks, and draws implying similar impulse responses are also associated with similar historical decompositions of inflation. In this case, demand shocks are prevalent over the sample and, in particular, over the last few years. In the US they account for more than 50 percent of inflation fluctuations in 2021 and almost 80 percent in 2022. The dominant role of demand factors in the post-COVID recovery is confirmed when we estimate a SVAR with a single-unit-root prior using euro area and four small open economies (Norway, Sweden, Canada and Australia) data. For example, demand shocks account for more than 50 percent of inflation fluctuations in 2022 in the euro area. However, while supply shocks play a limited role in the US since the middle of 2022, they still matter in the euro area until the middle of 2023 as the region is substantially more exposed to the Ukraine war, an important supply shock. In the other countries, the outcomes are surprisingly similar to the US: supply shocks drive inflation in 2020 but already in 2021 demand forces become prevalent.

For researchers who are reluctant to use our prior restrictions for inferential purposes, we offer two alternative pragmatic solutions, both of which make historical decompositions narrative more robust. The first is to demean the data prior to estimation. Such an approach can reduce the excess dispersion problem since poor estimates of the VAR constant contribute to making excess dispersion important. By reducing the dispersion of the deterministic component, one has the side benefit of reducing the uncertainty in the narrative, for any given historical decomposition. The second alternative treats the uncertainty in estimated VAR parameters symmetrically. That is, one draws from the (asymptotic or posterior) distribution of *all* model parameters, constructs the distribution of the historical decompositions and takes the pointwise median contribution of each shock, as in Bergholt et al. (2025). Such an approach does not reduce the excess dispersion of the deterministic component but takes it into account when computing a summary historical decomposition measure. It turns out that the alternative solutions we propose deliver a historical decomposition narrative for US inflation which is similar to the one obtained with the single-unit-root prior.

RELATED LITERATURE Our work contributes to three separate strands of the literature. First, we complement existing work by warning about an uncritical use of VARs for inference and for stylized fact collection, see e.g. Canova and Ferroni (2022) for a recent contribution. We point out that a realistic account of the uncertainty surrounding the deterministic component is crucial to produce credible historical decomposition narratives.

Second, we highlight an issue which, to the best of our knowledge, has not been discussed in the literature. One problem often touched upon is that the deterministic component tends to explain an implausibly large share of the low frequency variation of the data (the so-called overfitting problem), yielding implausible out-of-sample forecasts, see Sims (1996) and Sims (2000). Giannone et al. (2019) argue that this excessive explanatory power is due to transitional dynamics and occurs when the initial conditions are away from the steady state. They propose a prior specification, inspired by long-run the-

oretical predictions of macroeconomic models, that addresses the problem. We highlight that while both pathologies may cause poor out-of-sample forecasts, the excess dispersion problem is distinct from the overfitting problem. Thus, solving the latter does not ensure a resolution to the former and, as a consequence, does not necessarily produce less whimsical historical decompositions.

Finally, we contribute to rapidly growing literature on the drivers of the recent inflation increase, see e.g. Bernanke and Blanchard (2024), Beaudry et al. (2024), Gagliardone and Gertler (2023), Crump et al. (2024), Giannone and Primiceri (2024), Benigno and Eggertsson (2023), Schmitt-Grohé and Uribe (2023), and Bianchi et al. (2023) among many others. We stress the need of a robust interpretation of the historical decomposition of inflation and provide international evidence on the recent surge phenomenon.

OUTLINE The paper proceeds as follows. Section 2 illustrates the problem and Section 3 provides formal evidence on the sources of excess dispersion in the deterministic component. It also explains the differences between excess dispersion and overfitting. Section 4 proposes our solution. Section 5 presents estimates of the contribution of demand and supply disturbances for the US, for the euro area and for four small open economy countries. Section 6 illustrates two alternative pragmatic solutions to the whimsical historical decomposition problem. Finally, Section 7 concludes. The online appendix contains additional relevant material mentioned in the text.

2 WHIMSICAL HISTORICAL SHOCK DECOMPOSITIONS

This section shows, by way of an example, that the deterministic component of an estimated VAR may be subject to a large uncertainty and demonstrates the implications this excess dispersion has for historical shock decompositions. We stress that the issue is pervasive when VAR models are fitted to commonly used macroeconomic data. A more formal presentation of the problem appears in the next section.

2.1 AN OUTLINE OF THE PROBLEM

Consider the reduced-form VAR model with n variables and p lags:

$$Y_t = C + A_1 Y_{t-1} + A_2 Y_{t-2} + ... + A_p Y_{t-p} + u_t \tag{1}$$

where Y_t is the n x 1 vector of the endogenous variables, $u_t \sim N(0_n, \Sigma)$ is the $n \times 1$ vector of innovations, $A_1, ..., A_p$ are n x n coefficient matrices associated with lagged variables, and C is the n x 1 vector of constants. The reduced-form innovations u_t are linear combinations of underlying economic shocks: $u_t = Fe_t$, where F is a matrix restricted by some structural identification scheme. We focus on the system (1), as our problem arises in the estimation of a reduced-form model. The role played by structural identification schemes is discussed later.

We find it convenient to rewrite the system in companion form. Define the np x 1 vectors $\mathbf{Y}_t = (Y_t', Y_{t-1}', ..., Y_{t-p+1}')'$, $\mathbf{u}_t = (u_t', 0, ..., 0)'$, $\mathbf{C} = (C', 0, ..., 0)'$, and the np x

np companion matrix:

$$\mathbf{A} = \begin{bmatrix} A_1 & A_2 & ... & ... & A_p \\ I_n & \mathbf{0}_n & ... & ... & \mathbf{0}_n \\ \mathbf{0}_n & I_n & ... & ... & \mathbf{0}_n \\ . & . & . & ... & . \\ \mathbf{0}_n & \mathbf{0}_n & ... & I_n & \mathbf{0}_n \end{bmatrix}$$

With this notation the VAR is:

$$\mathbf{Y}_t = \mathbf{C} + \mathbf{A}\mathbf{Y}_{t-1} + \mathbf{u}_t \tag{2}$$

For the purpose of an historical shock decomposition, we express the right-hand side of (2) as the sum of a *deterministic* and a *stochastic* component, DC_t and SC_t:

$$\mathbf{Y_t} = DC_t + SC_t \tag{3}$$

Equation (3) follows from backward substitution, with DC_t and SC_t defined as:

$$\begin{aligned} DC_t &= \left(\mathbf{I} + \mathbf{A} + \cdots + \mathbf{A}^{t-1}\right)\mathbf{C} + \mathbf{A}^t\mathbf{Y}_0 \\ &= (\mathbf{I} - \mathbf{A})^{-1}\mathbf{C} + \mathbf{A}^t(\mathbf{Y}_0 - (\mathbf{I} - \mathbf{A})^{-1}\mathbf{C}) \qquad (4) \\ SC_t &= \mathbf{A}^{t-1}\mathbf{u}_1 + \cdots + \mathbf{A}\mathbf{u}_{t-1} + \mathbf{u}_t \qquad (5) \end{aligned}$$

The deterministic component DC_t depends on the model parameters ($\mathbf{A}$, $\mathbf{C}$) and the initial state vector $\mathbf{Y}_0$. It is the forecast of $\mathbf{Y}_t$ made in period 0, i.e. the counterfactual path of $\mathbf{Y}_t$ absent any innovations from period 1 to period t. The second line of (4) indicates that DC_t equals the steady state of $\mathbf{Y}_t$ plus a transitional dynamic term that depends on how far the initial $\mathbf{Y}_0$ is from the steady state. The stochastic component SC_t, instead, is a discounted sum of innovations between time 1 and t. Since $\mathbf{A}^t$ is decreasing in t for stationary and ergodic processes, shocks in the distant past play only a negligible role for the current observation as t grows.

Conceptually, the calculation of a historical shock decomposition requires two steps: (a) separating DC_t from SC_t, as in (3), and (b) disentangling the contribution of individual, structural shocks e_t. Our focus is on the first of these two steps. In particular, we establish that conditional likelihood-based estimation procedures provide a very imprecise estimate of DC_t: even small perturbations of the parametrization of (2) may result in large changes in DC_t, despite having only a minimal impact on the likelihood function. This is problematic for researchers who want to estimate a historical shock decomposition. Because the deterministic and the stochastic components enter linearly in (3), estimation uncertainty associated with DC_t must imply an equivalent estimation uncertainty associated with SC_t, with the paths for the structural shocks e_t adjusting accordingly.

The uncertainty in DC_t depends on the uncertainty in the estimates of $\mathbf{A}$ and $\mathbf{C}$. Because impulse responses are computed as the difference between two conditional forecasts, $\text{IRF}_t = E(\mathbf{Y}_t|\mathbf{u}_1 = 1, \mathbf{u}_k = 0,\ k > 1) - E(\mathbf{Y}_t|\mathbf{u}_k = 0,\ k > 0)$, the uncertainty present in the maximum likelihood estimates of $\mathbf{A}$ governs the dispersion of impulse response estimates. Thus, models associated with different draws of $\mathbf{A}$ but producing similar impulse responses may induce different narratives regarding the contribution of structural shocks to observed fluctuations whenever the uncertainty in $\mathbf{C}$ is significant. This is the sense in which historical shock decompositions become whimsical.

How pervasive is the issue? A few remarks serve to highlight the generality of the problem. First, given that DC_t is a reduced-form object, the problem occurs regardless of whether the VAR is estimated with frequentist or flat-prior Bayesian methods. In a Bayesian context, the problem leads to a wide posterior distribution for the deterministic component. A frequentist, instead, sees it in the form of large asymptotic confidence intervals around the deterministic component point estimate. In this paper, we take a Bayesian approach and our preferred solution to whimsical shock decompositions is inherently Bayesian in nature. An alternative solution we propose in Section 6 is designed for those who want to remain frequentist in their inferential approach. Second, whimsical shock decompositions occur even when $\mathbf{A}^t\mathbf{Y}_0$ tends to zero, suggesting that data transformations reducing the importance of the initial conditions for $\mathbf{Y}_t$ do not remove the problem. In fact, the problem emerges even when the VAR features ergodic, mildly persistent variables, displaying minor low frequency variations. Third and related, to the extent that dynamic stochastic general equilibrium (DSGE) models involve estimated parameters that enter the steady state, the issue may be relevant also for these models. After all, a linear solution to a DSGE generally admits a time series representation consistent with equations (2)-(5), and it is this representation that is confronted with data.

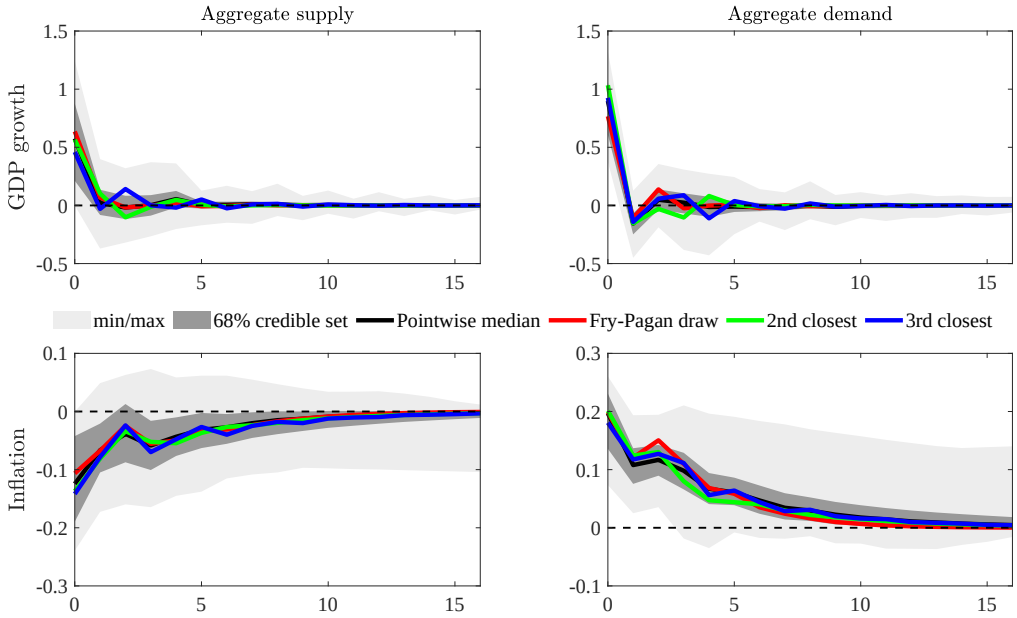
2.2 An illustration – A bivariate VAR

To illustrate the extent of the problem in practice, we estimate a bivariate VAR using US data on real GDP and the GDP deflator, both measured in log first differences. This is arguably the simplest possible setup to quantify the structural drivers of inflation. Our baseline sample runs from 1983:Q1 to 2022:Q4; we include four lags and a constant. Following the literature, we employ a diffuse prior (Jeffreys, 1946) on the VAR parameters and on the covariance matrix Σ.

We use standard sign restrictions to identify the two structural shocks. As in Canova and De Nicolò (2002), an aggregate demand shock is assumed to move GDP and inflation in the same direction, while an aggregate supply shock moves the two variables in opposite directions. These sign restrictions are imposed on impact only. Clearly, our estimated demand shock may bundle together various monetary and fiscal disturbances (see Giannone and Primiceri (2024) and Mori (2024)) and shocks related to the depletion of excess savings accumulated during the pandemic. Similarly, the supply shock we identify is likely a linear combination of disturbances to productivity, commodity prices, markups, and other supply-side factors. Kabaca and Tuzcuoglu (2023) disentangle an array of different supply shocks in a unified framework. Since our focus is on the role played by the deterministic component of a VAR in a historical decomposition exercise, a bivariate model, although minimal, is sufficient.

Figure 1 reports the impulse responses from the estimated structural model. We measure time in quarters on the x-axes and the estimated responses in percentage points on the y-axes. For each horizon, the black line represents the pointwise median response obtained using 1000 draws. The shaded areas instead represent the 68% and the min-max posterior credible sets (as in Baumeister and Hamilton (2015)), respectively. The pointwise median response is compared with three alternatives. As the pointwise median most likely mixes different draws across horizons, applied researchers often prefer to discuss the impulse responses obtained from one particular draw instead. A popular choice is the

Figure 1: Responses to identified aggregate demand and aggregate supply disturbances.



Note: The black line is the pointwise median and the shaded areas cover the 68 percent and min-max identified sets. The red line is the impulse response for the draw closest to the pointwise median; the green and blue lines are impulse responses for the 2nd and 3rd draws closest to the pointwise median, respectively.

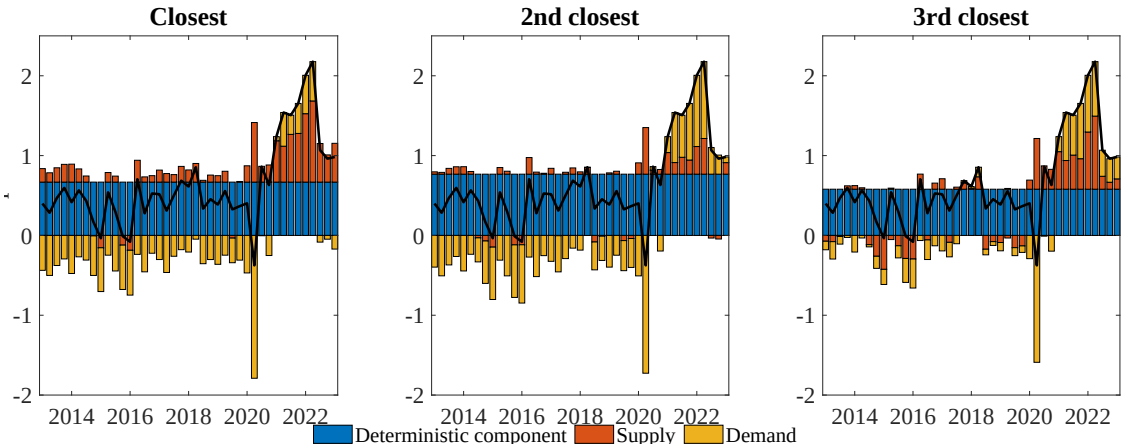
single draw that is closest to the pointwise median on average over horizons, see Fry and Pagan (2011). The impulse response obtained with this draw is reported as the red line. We also report the impulse responses obtained from the second and third closest draws to the pointwise median in green and in blue.

A couple of facts emerge from Figure 1. First, by construction both shocks are allowed to have a permanent effect on the level of GDP. Second, output growth responses are short-lived while inflation responses are more persistent. Third, the responses obtained from the three draws are almost indistinguishable from each other and close to those produced by the pointwise median. Thus, impulse responses are locally robust.

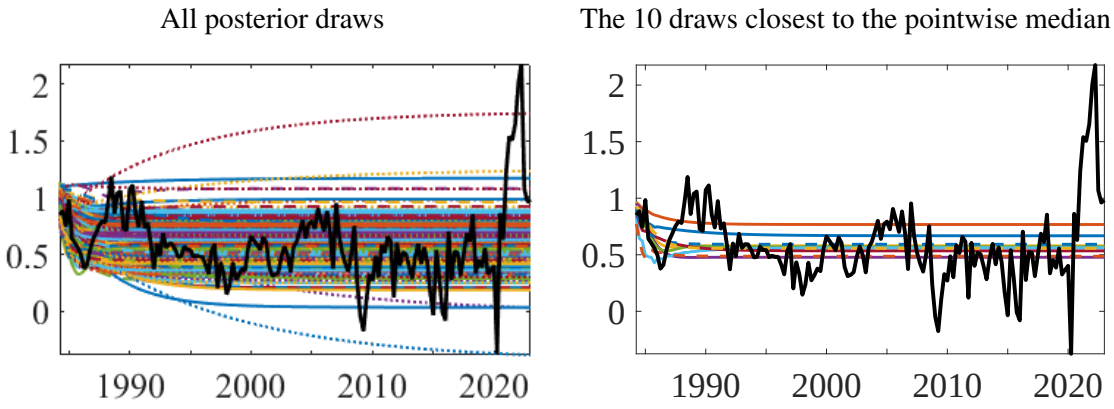
Panel (a) of Figure 2 plots the historical shock decomposition of inflation for the last eight years of the sample, conditioning on each of the three draws closest to the pointwise median impulse response. The blue bars correspond to the contribution of the deterministic component; the red and yellow bars represent the contribution of supply and demand disturbances, respectively. Combined, the contributions of supply and demand disturbances sums to the stochastic component SC_t. Interestingly, the three draws closest to the pointwise median impulse response produce different narratives about the post-pandemic inflation surge: the draw closest to the pointwise median implies that supply shocks explain more than two-thirds of the rise in inflation post-2020. The second closest draw, instead, implies that demand shocks account for more than two-thirds. The third closest draw suggests that demand and supply account for about fifty percent each. These differences are striking given that the three draws are associated with almost identical

Figure 2: Historical decompositions and deterministic components of US inflation

(a) Historical decomposition of inflation for the 3 draws closest to the pointwise median response.



(b) Deterministic components of inflation



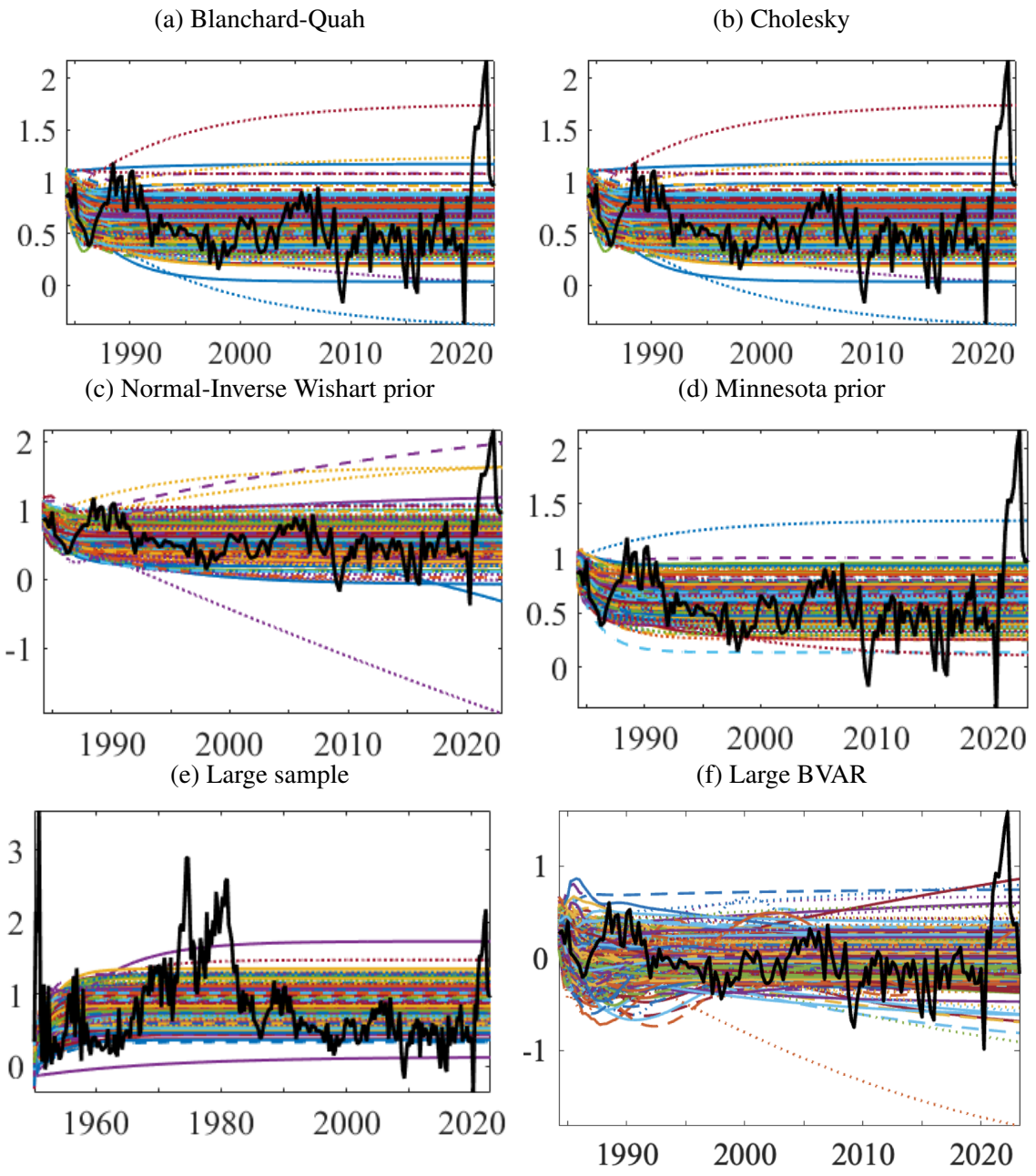
Note: The black line in all the panels shows quarterly percentage change of the GDP deflator. Panel (a) displays the historical shock decompositions of inflation from the 3 draws closest to the pointwise median response for the IRFs. The colored lines in panel (b) display the deterministic components of inflation for draws from the posterior distribution.

impulse responses. Panel (b) in Figure 2 plots the deterministic components of all draws on the left, and of the ten draws closest to the pointwise median on the right. Interestingly, the posterior dispersion of deterministic components is substantial even for the 10 draws closest to the pointwise median responses, and the time paths stabilize at very different levels. Thus, the uncertainty present in the blue bars drives the differences in the historical contribution of inflation across draws.

2.3 How pervasive is the roller coaster?

To document the generality of the problem, we show that a whimsical shock decomposition is produced in various contexts.

Figure 3: Estimated deterministic components of inflation.



Note: The figure reports the deterministic components for all draws from the posterior distribution. The black line in all the panels shows quarterly percentage change of the GDP deflator. The sample used in panel (e) (1949-2022) is different from the others (1983-2022).

2.3.1 The role of structural identification assumptions

First, we illustrate the irrelevance of structural identification assumptions, as the uncertainty in the deterministic component arises from the reduced-form estimation. Hence, irrespective of the identification scheme employed, and frequentist or flat-prior inference, a VAR is in principle susceptible to the problem highlighted here. We consider two of

the most popular identification strategies: the Blanchard-Quah decomposition, which imposes zero restrictions on the long-run cumulative sum of VAR coefficients (Blanchard and Quah, 1989), and the Cholesky decomposition, which imposes zero-restrictions on the impact coefficients (Sims, 1986). The Blanchard-Quah decomposition, in particular, is commonly used to separate demand from supply shocks under the assumption that demand shocks cannot affect the level of output in the long run.[1]

Because the Blanchard-Quah, Cholesky and sign-restricted decompositions share the same draws for the reduced form parameters, the uncertainty in the deterministic component is the same in all three schemes, see Figure 2 and the first row of Figure 3.

2.3.2 The role of reduced-form priors

Given that the dispersion we care about has to do with estimates of the VAR parameters, one may suspect that prior restrictions routinely employed to reduce estimated parameter uncertainty also may shrink the dispersion of the deterministic component. For illustration, we consider a standard conjugate Normal-Inverse Wishart prior and a Minnesota-like prior (see Doan et al. (1984)).[2]

As the second row of Figure 3 shows, the dispersion of the deterministic components remains large. The reason these standard priors are unable to reduce the dispersion of the deterministic components is that they are designed to shrink the uncertainty in VAR coefficients and in the covariance matrix of the disturbances, but they leave the prior on the constant diffuse.

2.3.3 The role of structural breaks

In our baseline estimation, we consider a relatively homogeneous sample period starting in 1983. Here, we re-estimate the model over the period 1949:Q1-2022:Q4. The results appear in the third row of Figure 3. The dispersion of the estimated deterministic component is substantially larger than in the baseline model. Thus, the massive uncertainty around estimates of the deterministic component is an additional argument for not using a long and dis-homogeneous sample when estimating a SVAR model (see Furlanetto et al. (2025) for a detailed discussion of this point).

2.3.4 The role of the VAR dimension

Next, we examine the role played by the dimensionality of the VAR system. Canova and Ferroni (2022) have shown that small-scale VARs may be prone to deformation prob-

[1] Neither the Blanchard-Quah decomposition nor the Cholesky decomposition guarantee sensible signs of the inflation and GDP growth responses to shocks. Indeed, Figure A-1 in the Appendix shows that Blanchard-Quah delivers positive co-movement of the two variables in response to a supply shock. Furlanetto et al. (2025) addresses this issue by combining Blanchard-Quah with sign restrictions.

[2] In the Normal-Inverse Wishart case, the prior for the autoregressive (AR) coefficients is normal, centered at zero with a diagonal covariance matrix of 10, while the prior for the covariance matrix of the residuals is inverse Wishart with a unitary diagonal matrix as scale and 3 degrees of freedom. The Minnesota-like prior is also normally distributed for all AR coefficients and is centered around zero for all parameters, including the variables' first own lag, as VAR variables appear in growth rates. We set the hyperparameters of the priors as in Giannone et al. (2015). The diagonal elements of the scale matrix of the inverse Wishart prior are set to the residual variance of an AR(1) process for each variables. In both cases, we still use sign restrictions to identify the shocks and compute historical decompositions.

lems, making it desirable to estimate models of a certain size to ensure that structural objects, such as impulse responses, make economic sense. Here, we estimate a 5-variable VAR model with real GDP, the GDP deflator, real gross private domestic investment, the Federal funds rate, and real wages. The latter is measured as average hourly earnings of production and non-supervisory employees, deflated with the GDP deflator. All variables except the Federal funds rate are measured in log first differences. The sample is 1983:Q1-2022:Q4, as in our baseline setup.

Estimates of the deterministic components of inflation are in panel (f) of Figure 3 (see Figure E-1 for the deterministic components of each VAR variable). Evidently, the excessive dispersion problem remains even in a larger scale VAR. If anything, it seems exacerbated and becomes rather dramatic for variables like inflation or the Federal funds rate.

2.3.5 THE ROLE OF PRELIMINARY DATA TRANSFORMATIONS

Our baseline result is obtained with a VAR where output and prices are in log first differences. Given that researchers sometimes prefer to study the transmission of shocks in VARs where the variables are measured in level (to retain cointegrating relationships), it is worth asking whether the large dispersion of the deterministic components is a specific feature of VARs in log first differences.

It is easy to show that if the variables are measured in log growth rate, the VAR in level corresponding to (1) is

$$\begin{aligned} Y_t &= C + (I + A_1)Y_{t-1} + (A_2 - A_1)Y_{t-2} + ... + (A_p - A_{p-1})Y_{t-p} - A_pY_{t-p-1} + u_t \\ &= C + \tilde{A}_1Y_{t-1} + \tilde{A}_2Y_{t-2} + ... + \tilde{A}_pY_{t-p} - A_pY_{t-p-1} + u_t \end{aligned} \tag{6}$$

Thus, a VAR in growth rates simply has a different companion matrix than a VAR in levels. Hence, the decomposition we have derived in (3) still holds with a different $\mathbf{A}$ matrix. Thus, a dispersed $\mathbf{C}$ estimate implies a dispersed deterministic component and a similarly whimsical historical decomposition.

3 A FORMAL LOOK AT EXCESS DISPERSION PROBLEM

In this section we formally analyze what drives the dispersion of the deterministic component, A simple Monte Carlo exercise complements the analytical derivation with some intuition. We also discuss how the deterministic component excess dispersion relates to the overfitting problem discussed in the previous literature.

3.1 ANALYTICAL DERIVATIONS

We study the dispersion determinants in terms of *mean squared error* (MSE), as this formulation is more natural in the Bayesian setting we employ. The details of the derivations are in Appendix B. There we show that the mean squared error of $\hat{\mathbf{C}}$ is:

$$\begin{aligned} \text{MSE}(\hat{\mathbf{C}}) &= E\Big(\big(\hat{\mathbf{C}} - \mathbf{C}\big)\big(\hat{\mathbf{C}} - \mathbf{C}\big)'\Big) \\ &= E\Big(\big(\bar{\mathbf{u}} - (\hat{\mathbf{A}} - \mathbf{A})\bar{\mathbf{Y}}_{-1}\big)\big(\bar{\mathbf{u}} - (\hat{\mathbf{A}} - \mathbf{A})\bar{\mathbf{Y}}_{-1}\big)'\Big) \end{aligned}$$

$$\approx \quad \frac{1}{T}\boldsymbol{\Sigma}_{\mathbf{u}} + \text{Bias}(\hat{\mathbf{A}})\Big(\frac{1}{T}\boldsymbol{\Sigma}_{\mathbf{Y}} + \boldsymbol{\mu}\boldsymbol{\mu}'\Big)\text{Bias}(\hat{\mathbf{A}})' \tag{7}$$

where $\bar{\mathbf{u}} = \frac{1}{T}\sum_{t=1}^{T}\mathbf{u}_t$, $\boldsymbol{\Sigma}_{\mathbf{u}} = \text{var}(\mathbf{u_t})$, $\bar{\mathbf{Y}}_{-1} = \frac{1}{T}\sum_{t=1}^{T}\mathbf{Y}_{t-1}$, $\boldsymbol{\Sigma}_{\mathbf{Y}} = \text{var}(\mathbf{Y_t})$ and $\boldsymbol{\mu} = (\mathbf{I}-\mathbf{A})^{-1}\mathbf{C}$. The bias in the conditional maximum likelihood estimate of $\mathbf{A}$ is:

$$\begin{aligned}\text{Bias}(\hat{\mathbf{A}}) &= E(\hat{\mathbf{A}} - \mathbf{A}) \\ &= E\Big(\big(\sum_{t=1}^{T}\tilde{\mathbf{Y}}_{t-1}\tilde{\mathbf{Y}}_{t-1}'\big)^{-1}\big(\sum_{t=1}^{T}\tilde{\mathbf{Y}}_{t-1}\tilde{\mathbf{u}}_t'\big)\Big) \approx -(\mathbf{B_A})/T\end{aligned} \tag{8}$$

where $\mathbf{B_A} = \boldsymbol{\Sigma}_{\mathbf{u}}\big((\mathbf{I}-\mathbf{A}')^{-1} + \mathbf{A}'(\mathbf{I}-\mathbf{A}'^2)^{-1} + \sum_\lambda \lambda(\mathbf{I}-\lambda\mathbf{A}')^{-1}\big)(\boldsymbol{\Sigma}_{\mathbf{Y}} + \boldsymbol{\mu}\boldsymbol{\mu}')^{-1}$, λ are the eigenvalues of $\mathbf{A}$ $\tilde{\mathbf{Y}}_t = \mathbf{Y}_t - \bar{\mathbf{Y}}$ and $\tilde{\mathbf{u}}_t = \mathbf{u}_t - \bar{\mathbf{u}}$, see Kilian and Lütkepohl (2017) for details. Generally, the direction of the bias depends on the eigenvalues of $\mathbf{A}$. For persistent processes, the bias in $\mathbf{A}$ tends to be negative.

Equation (7) highlights the sources of dispersion in the conditional maximum likelihood estimate of the constant term, $\hat{\mathbf{C}}$. First, holding $\mathbf{A}$ and T fixed, the dispersion of $\hat{\mathbf{C}}$ is increasing in $\boldsymbol{\Sigma}_{\mathbf{u}}$. Second, both the bias in $\hat{\mathbf{A}}$ and the scaled covariance matrix $T^{-1}\boldsymbol{\Sigma}_{\mathbf{u}}$ are larger when the sample size T is small. Thus, the dispersion of $\hat{\mathbf{C}}$ is decreasing in the sample size, given $\mathbf{A}$ and $\boldsymbol{\Sigma}_{\mathbf{u}}$. Third, the larger (in absolute value) the bias in $\hat{\mathbf{A}}$ is, the larger is the dispersion of $\hat{\mathbf{C}}$, given $\boldsymbol{\Sigma}_{\mathbf{u}}$ and the sample size T. In contrast, if this bias vanishes, the MSE of $\hat{\mathbf{C}}$ collapses to $T^{-1}\boldsymbol{\Sigma}_{\mathbf{u}}$.

These observations have important implications for the dispersion in the deterministic component and, thus, for historical shock decompositions. In Appendix B we show that:

$$\begin{aligned}\text{MSE}(\hat{DC}_t) &= E\left(\big(\hat{DC}_t - DC_t\big)\big(\hat{DC}_t - DC_t\big)'\right) \\ &= \Gamma(\mathbf{A})\text{MSE}(\hat{\mathbf{C}})\Gamma(\mathbf{A})' + \text{other terms}\end{aligned}$$

where $\Gamma(\mathbf{A}) = (\mathbf{I}-\mathbf{A}^t)(\mathbf{I}-\mathbf{A})^{-1}$ and the other terms are defined in the appendix. Moreover, we show that $\Gamma(\mathbf{A})\text{MSE}(\hat{\mathbf{C}})\Gamma(\mathbf{A})'$ is the *leading term* of this expression. Thus, given $\mathbf{A}$, the dispersion of the deterministic component increases in the dispersion of the constant term. It follows that imprecise shock decompositions tend to be a more important concerns in VAR estimated with smaller sample sizes or where the bias in $\hat{\mathbf{A}}$ is larger in absolute value.

3.2 Some simulation results

We perform a few simulation exercises to bring some intuition into the results of subsection 3.1. We generate data from a bivariate VAR(1):

$$Y_t = C + AY_{t-1} + u_t,$$

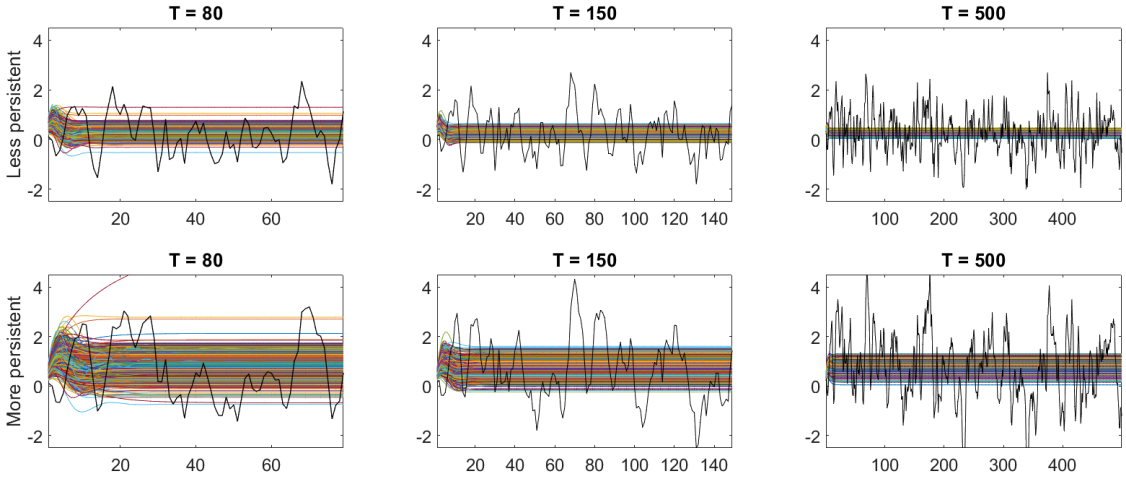
where $u_t \sim N(0, I)$. The parameters are calibrated as follows:

$$C = \begin{pmatrix}0.4\\0.5\end{pmatrix}, \quad \text{and} \quad A = \begin{pmatrix}a_{1,1} & -0.3\\0.3 & 0.4\end{pmatrix}.$$

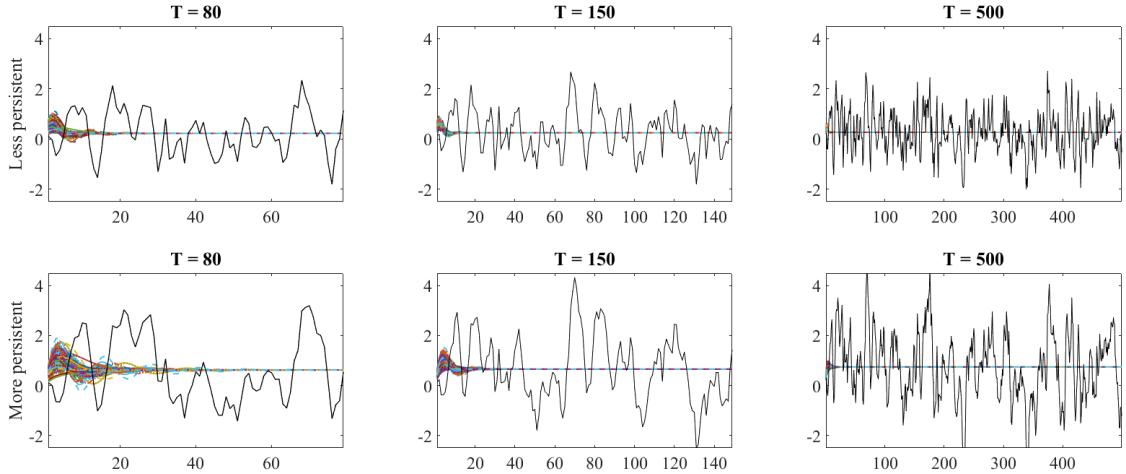
We limit the attention to the first variable of the system and consider different degrees of persistence as well as different sample sizes. The role of persistence is assessed by

Figure 4: Estimated deterministic components of $Y_{1,t}$.

(a) Diffuse prior



(b) Single-unit-root prior



Note: The figure reports the estimated deterministic components for $Y_{1,t}$ for 1000 draws from the posterior distribution. The black lines represent the simulated data.

varying $a_{1,1}$, the top-left element of A. We set $a_{1,1} = 0.6$ or $a_{1,1} = 0.95$.The constant term is the same in both cases. We compare samples with $T = 80$, 150, and 500 observations, respectively, and estimate the model with a diffuse prior. We measure the dispersion of the deterministic component using 1000 draws from the posterior distribution of the parameters.

Panel (a) of Figure 4 shows the estimated deterministic components. As expected, the more persistent the data or the shorter the sample, the greater is the dispersion in the estimated deterministic component. The intuition is straightforward: more persistent processes imply larger eigenvalues λ and larger biases in the estimates of the VAR coefficients. These biases, in turn, increases the dispersion of the estimated constant, making

the deterministic component estimates more uncertain, see (7). Similarly, when the sample size is small, biases in the VAR coefficients are significant for any level of persistence.

Interestingly, even when $T = 500$ and the persistence is moderate, posterior dispersion is still substantial. In fact, the dispersion is considerable even when persistence is low ($a_{1,1} = 0.2$) and the sample is moderate ($T = 150$), as shown in Figure C-1 in the appendix. Hence, while the data persistence and sample size matter, the flatness of the likelihood function in the constant space is crucial in driving the excess dispersion result.

3.3 Overfitting vs. excess dispersion

It useful to relate the well-known overfitting behavior of the deterministic component of a VAR to the more novel excess dispersion problem studied in this paper. As discussed in Sims (1996), Sims (2000) and Giannone et al. (2019), flat-prior VARs tend to attribute an implausibly large share of the variations in the observed time series to their deterministic components. The problem arises because the initial values of the variables are treated as non-stochastic when constructing the likelihood function – and this removes a penalty for them being away from their steady state – and because a flat prior on $(\mathbf{A}, \mathbf{C})$ implies an informative prior for $((\mathbf{I}-\mathbf{A})^{-1}\mathbf{C}, \mathbf{A})$, where there is little density mass for the eigenvalues of $\mathbf{A}$ in the proximity of one. To illustrate, take equation (4), which we reproduce for the reader's convenience:

$$DC_t = (\mathbf{I}-\mathbf{A})^{-1}\mathbf{C} + \mathbf{A}^t(\mathbf{Y}_0 - (\mathbf{I}-\mathbf{A})^{-1}\mathbf{C})$$

Overfitting may arise in a stationary environment whenever $\mathbf{Y}_0 \neq (\mathbf{I}-\mathbf{A})^{-1}\mathbf{C}$, as this induces transitional dynamics in the deterministic component that may persistently explain the evolution in $\mathbf{Y}_t$. Examples of such transitional dynamics are presented in Sims (2000) and Giannone et al. (2019), and are obtained, among others, in VARs with persistent labor market variables such as hours worked and unemployment. Clearly, if $\mathbf{A} = \mathbf{I}$, unit roots are present and $DC_t = \mathbf{Y}_0, \forall t$. On the other hand, if $\mathbf{Y}_0 = (\mathbf{I}-\mathbf{A})^{-1}\mathbf{C}$, the initial conditions are at the steady state, there is no overfitting.

To see that that excess dispersion is distinct from overfitting, consider this latter case, $\mathbf{Y}_0 = (\mathbf{I}-\mathbf{A})^{-1}\mathbf{C}$. Here no overfitting occurs. Still, if $\mathbf{C}$ is poorly estimated, historical decompositions may be whimsical in the sense that the narrative provided by impulse responses and by historical decompositions could be different. Thus, the elimination of overfitting does not necessarily remove excess dispersion.

However, there is a sense in which the two problems are related. They both concern the deterministic component of $\mathbf{Y}_t$ and both affect long-run forecasts. In addition, they are more pronounced in large systems, in VARs with persistent variables, and in samples of smaller size. A further interesting connection emerges in the case when the data-generating process is a random walk without drift. While inflation seems stationary in our sample starting in 1983, the case of a random walk without drift is conceivable for a less homogeneous sample (for example, a sample starting in the 1970s). In that case, overfitting and excess dispersion jointly appear, both in finite and large samples, as discussed in the Appendix (see Figure D-1).

4 SOLVING THE PROBLEM

Having established how and why an estimated shock decomposition may be whimsical, we next turn to possible solutions to the problem.

First, notice that a whimsical decomposition arises because, even though researchers follow a Bayesian approach when presenting impulse responses, they often forget their Bayesian point of view when it comes to the historical decompositions. Under a quadratic (absolute) loss function, one should report the mean (median) historical decomposition across draws, rather than the estimate obtained using a point estimate of the VAR parameters. Previous sections have shown that reporting the latter may be dangerous as the historical decomposition narrative may become erratic, even when impulse responses are not. Hence, one obvious solution to the problem is to perform the required integration and report the average (median) historical decomposition. We discuss this option in subsection 6.2. While current computer machines can perform such an exercise in a matter of seconds, most software packages are not setup to routinely produce it. In addition, there is no insurance that, once the exercise is performed, IRFs and historical decompositions will display compatible information. For this reason, we provide an alternative solution that adds prior restrictions. These restrictions robustify historical decomposition inference by making the deterministic component less dispersed.

We exploit the single-unit-root prior, also known as the "dummy initial observations" prior, first introduced by Sims (1993). We show that, when such a prior is properly set, it dramatically restricts the posterior dispersion of the estimated deterministic component. Given that the dispersion in the stochastic component is unaffected, our prior makes the uncertainty in the estimated historical decompositions and estimated impulse responses comparable. In this sense, the narrative historical decompositions provide becomes robust.

4.1 PRIOR SHRINKAGE OF THE DETERMINISTIC COMPONENT

We add an artificial observation to the data, where both the current and lagged data are given by $\frac{1}{\delta}\bar{\mathbf{Y}}_0$ and the intercept is set to $\frac{1}{\delta}$. The idea is that, if lags of $\mathbf{Y}_t$ are at some value $\bar{\mathbf{Y}}_0$, then also $\mathbf{Y_t}$ will tend to be a-priori close to $\bar{\mathbf{Y}}_0$. δ is a non-negative hyperparameter while $\bar{\mathbf{Y}}_0$ is a vector of size np x 1. The stochastic constraint imposed by our artificial observation on the VAR can be written as

$$(\mathbf{I} - \mathbf{A})\,\bar{\mathbf{Y}}_0 - \mathbf{C} = \delta \mathbf{u}_0, \tag{9}$$

where $\mathbf{u}_0 = (u_0', 0, ..., 0)'$ is a vector of size np x 1 (see Miranda-Agrippino and Ricco (2019) for details). δ governs the tightness of the constraint. At one extreme, as δ approaches infinity. the constraint becomes uninformative At the other, as δ approaches zero, the constraint implies either (i) a nonstationary VAR with at least one (common) unit root and $\mathbf{C} = 0$, or (ii) a stationary VAR, with $\mathbf{C} \neq 0$. If $\mathbf{u}_0$ is a random variable with zero mean and unit variance, then the dummy observation restricts the estimated unconditional mean of $\mathbf{Y}_t$, $(\mathbf{I} - \mathbf{A})^{-1}\mathbf{C}$, to be close to $\bar{\mathbf{Y}}_0$ as δ approaches zero.

To appreciate why (9) is useful, we manipulate equation (4) and substitute out $\mathbf{C}$, using the stochastic restriction (9), to arrive at the following expression:

$$DC_t = \mathbf{A}^t(\mathbf{Y}_0 - \bar{\mathbf{Y}}_0 + (\mathbf{I} - \mathbf{A})^{-1}\delta\mathbf{u}_0) + \bar{\mathbf{Y}}_0 - (\mathbf{I} - \mathbf{A})^{-1}\delta\mathbf{u}_0 \tag{10}$$

If the system is stationary and ergodic, $\mathbf{A}^t(\mathbf{Y}_0 - \bar{\mathbf{Y}}_0 + (\mathbf{I} - \mathbf{A})^{-1}\delta\mathbf{u}_0)$ approaches zero when the sample size grows large. Thus, asymptotically, equation (10) implies that the deterministic component of the data a-priori fluctuates around $\bar{\mathbf{Y}}_0$, with δ regulating the tightness of the constraint. If $\delta \to 0$, the deterministic component will approach $\bar{\mathbf{Y}}_0$, effectively eliminating any dispersion. Notably, this happens regardless of $\mathbf{A}$ and Σ. The stochastic component will still be uncertain—different draws of $\mathbf{A}$ and F will still potentially affect the estimated, historical importance of different shocks. However, this type of uncertainty is the same as the one present in the impulse responses, making the two statistics consistent.

The choice of $\bar{\mathbf{Y}}_0$ and δ are important for the properties of DC_t. When $\bar{\mathbf{Y}}_0$ is set to the (a priori) steady state, our prior has the same flavor as the steady state prior of Villani (2009), see Karlsson (2013). When it is set to zero, instead, it effectively demeans the data. Other choices are possible. $\bar{\mathbf{Y}}_0$ can be set to the pre-sample mean, as in Giannone et al. (2015), to a sample average, or to a value dictated by theory or by policy experience. For example, in the case of inflation, a value of 2 percent, corresponding to the inflation target, could be chosen. Thus, our specification is flexible as far as centering the prior restriction. We treat the hyperparameter δ, which regulates the dispersion of the prior, as a random variable and construct its posterior distribution jointly with the other VAR parameters. The posterior mode of δ will then give us a sense of how informative the data is about the dispersion of the deterministic component.

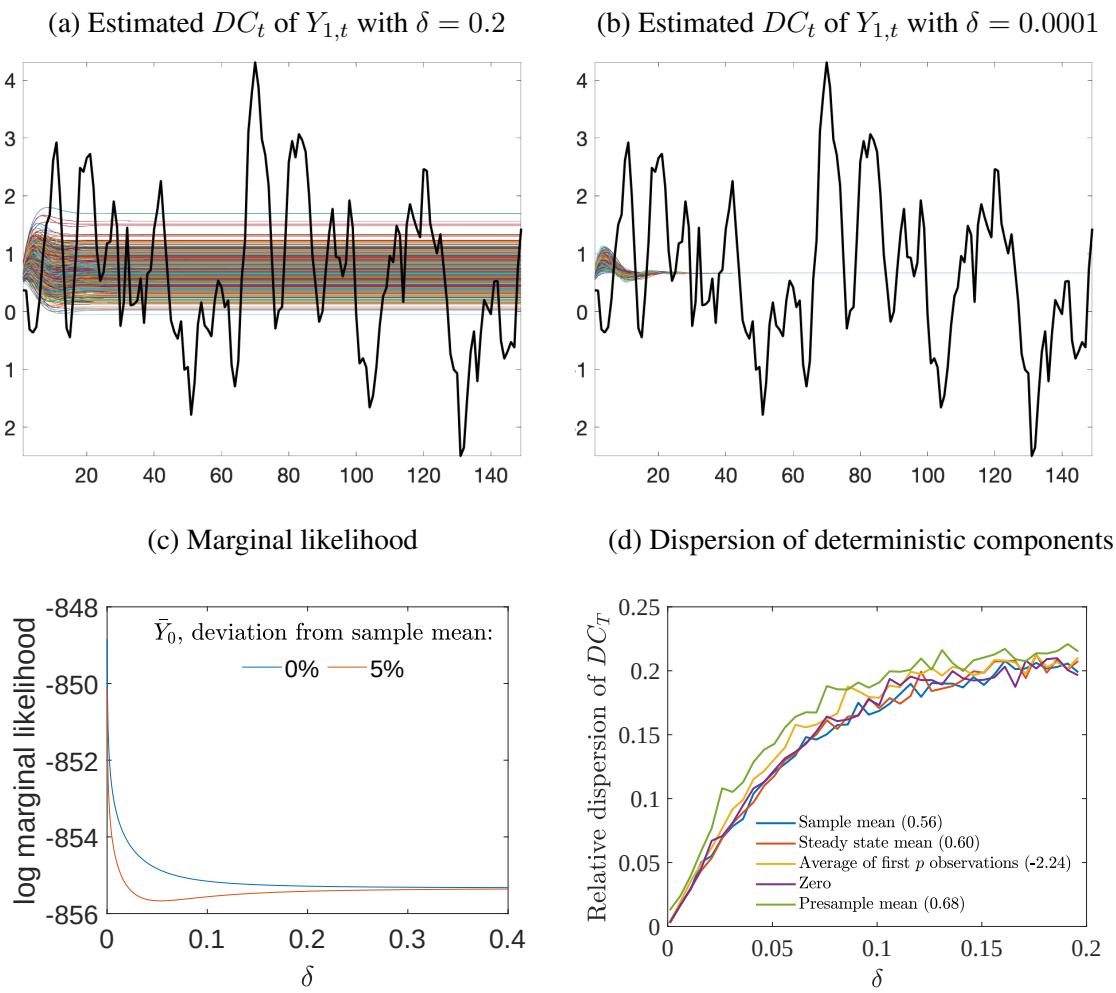
We would like to highlight a few important points. First, imposing an informative prior on the VAR constant is possible but it would not necessarily solve the problem as the steady state of the model would be a-priori centered around zero rather than $\bar{\mathbf{Y}}_0$ [3] Second, while Karlsson (2013) highlights the existence of a relationship between the single-unit-root and the steady state priors (see footnote 8, page 814), he does not use the former to solve the excess dispersion problem. Third, Antolin-Diaz et al. (2021) argue that the single-unit-root prior may help to address the problem of excessive explanatory power of initial conditions. However, they do not tackle the excess dispersion problem either, nor show how a single-unit-root prior could take care of it.

4.2 Implications of the single-unit-root prior in simulation

In order to illustrate how the single-unit-root prior affects the dispersion of the deterministic component, we repeat the simulation exercise estimating the VAR with the single-unit-root prior. Panel (b) of Figure 4 reports the deterministic component for 1000 parameter draws from the estimated posteriors. We set $\bar{\mathbf{Y}}_\mathbf{0}$ equal to the sample mean and we choose a Gamma prior density for δ with mode equal to 1 and standard deviation equal to 1, which is quite loose, and construct the posterior distribution of the VAR parameters in a standard hierarchial fashion (see Giannone et al. (2015)). The comparison with Panel (a) in Figure 4 is striking: after about 20 periods, all DC_t draws converge to virtually identical values, irrespective of the data persistence or the sample size. Because they all

[3] Imposing a prior on the VAR constant has two additional disadvantages are the following. The prior for the VAR coefficients and for the constant would be independent, complicating posterior sampling; these is no insurance that the prior on endogenous objects, such as the steady states, is meaningful, see e.g., Del Negro and Schorfheide (2011).

Figure 5: The role played by the hyperparameter δ



Note: Panels (a) and (b) report the estimated deterministic components for $Y_{1,t}$ for all draws from the posterior distribution, for different levels of δ. The black lines are the actual values. Panel (c) displays the log marginal likelihood for different values of δ when $\bar{Y}_0$ is equal to the sample mean (blue line) and when it deviates from the sample mean by 5% (red line). Panel (d) shows the dispersion of the estimated deterministic components at the end of the sample for different values of δ and different values of $\bar{Y}_0$.

converge to $\bar{\mathbf{Y}}_\mathbf{0}$, the uncertainty in the shock decomposition becomes consistent with the uncertainty in impulse responses. The key behind this result is the discipline imposed by δ, whose mode is estimated to be 0.0001.

Panels (a) and (b) in Figure 5 shed further light on the role played by δ. These panels plot the deterministic components when we redo the simulation exercise in subsection 3.2, but fix δ at $\delta = 0.2$ and $\delta = 0.0001$ respectively. Recall that the former value corresponds to the "rule-of-thumb" used in the literature (see, for example, Sims and Zha (1998)), while the latter is what we obtain as modal value. Note that such a low value of δ implies that the stochastic restriction $(\mathbf{I} - \mathbf{A})\,\bar{\mathbf{Y}}_0 = \mathbf{C}$ holds almost exactly. In Figure 5 we limit our attention to the case $a_{1,1} = 0.95$ and $T = 150$ observations.

The differences between Panels (a) and (b) illustrate the crucial role played by δ: with $\delta = 0.2$ considerable dispersion in the deterministic component remains, with the

long-run values settling at vastly different estimates. This may suggest that the "rule-of-thumb" value of 0.2 can be problematic for researchers interested in shock decompositions. $\delta = 0.0001$, instead, forces all 1,000 posterior draws to share the same deterministic component after relatively few periods. As we will see, a low δ is consistent with the modal estimate obtained using data for many countries.

Why is a very small value of δ preferred by data? Given that small values of δ imply less dispersion in the deterministic component, they may also deliver more precise forecasts. Panel (c) in Figure 5 reports the implied marginal likelihood as a function of δ. We consider two cases: when $\bar{\mathbf{Y}}_\mathbf{0}$ is equal to the mean $(\mathbf{I} - \mathbf{A})^{-1}\mathbf{C}$ (blue line), and when the deviation between the two is 5% (red line). In the first case, we note that the marginal likelihood is strictly decreasing in δ, consistent with greater precision in the one-step ahead forecasts of the model. In the second case the marginal likelihood is not strictly decreasing in δ. Still, even with a 5% deviation in the initial observation we obtain a much higher marginal likelihood for δ close to zero than, say, for $\delta = 0.2$. Thus, to optimally forecast one-step-ahead, it is preferable to considerably reduce the uncertainty around the deterministic component, and we achieve this when δ is close to zero.

Finally, Panel (d) in Figure 5 vizualises the trade-off one may face when jointly setting $\bar{\mathbf{Y}}_\mathbf{0}$ and δ. The figure reports the posterior, end-of-sample, dispersion in the deterministic component for different combinations of $\bar{\mathbf{Y}}_\mathbf{0}$ and δ. The end-of-sample dispersion is expressed as the standard deviation of the deterministic component, relative to the standard deviation of the variable. Three observations emerge. First, the choice of $\bar{Y}_0$ does not matter. Thus, issues of in-sample vs. out-of-sample predictability do not arise. Second, the value used by Antolin-Diaz et al. (2021) ($\delta = 0.05$), which the authors consider useful to reduce overfitting in their data, is insufficient to eliminate the excess dispersion problem. Third, it is only for small values of δ that impulse responses and historical decompositions uncertainty is comparable.

Since free lunches are rare in practice, one may wonder if there costs when using our prior. While our specification solves the excess dispersion problem, it could in principle have important side effects and, e.g., significantly alter the mean forecasts or the dynamics in response to shocks. Table C-1 in the online appendix shows that the forecasting performance, as measured by the root mean squared error, of a VAR with or without the single-unit-root prior is similar at different horizons. Thus our prior, while reducing out-of-sample uncertainty, does not affect the mean of the forecasts. On the other hand, Figure C-2 and Figure C-3 in the online appendix show that both the posterior distribution of the AR coefficients and the impulse responses are unaffected by our prior. Thus, our prior restriction makes the information contained in impulse responses and historical decompositions consistent without imposing significant distortions in other statistics.

5 THE DRIVERS OF THE POST-PANDEMIC INFLATION

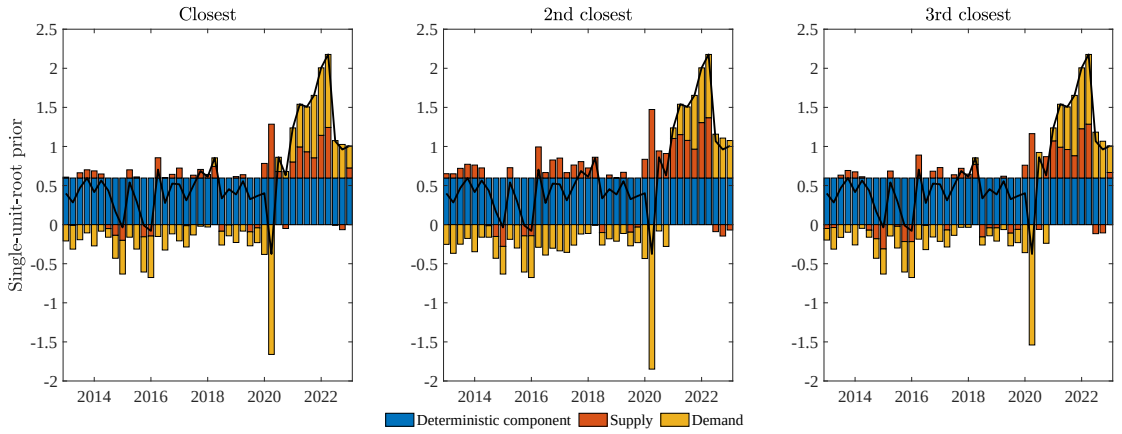
This section studies the drivers of the recent inflation surge in a number of countries, taking into account the dispersion in the deterministic components. We estimate a bivariate SVAR with real GDP and GDP deflator, both in log first differences, restricted with the single-unit-root prior, for the US, the euro area, and four other small open economies. We employ the same identifying sign restrictions used in section 2.2 to recover demand and supply shocks. We choose a Gamma prior density for δ with mode equal to 1 and

standard deviation equal to 1, and construct the posterior distribution for the parameters in a standard hierarchial fashion. In all cases, $\bar{\mathbf{Y}}_0$ is set to the sample mean.

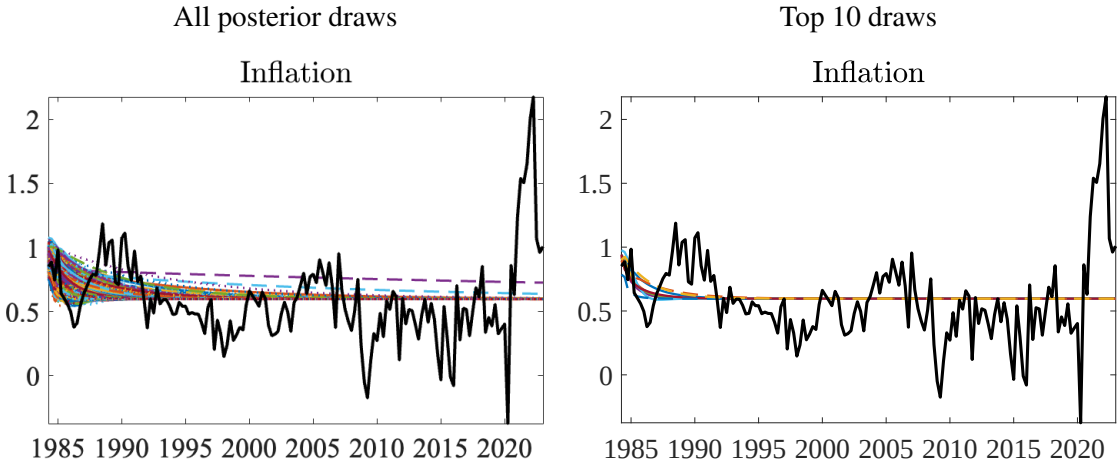
5.1 US INFLATION

Figure 6: US estimates using the single-unit-root prior.

(a) Historical decomposition of inflation for top 3 draws



(b) Deterministic components of inflation



Note: The black line in all the panels shows quarterly percentage change of the GDP deflator. Panel (a) displays the historical shock decompositions of inflation from the 3 draws closest to the pointwise median response for the IRFs. The colored lines in panel (b) display the deterministic components of inflation for draws from the posterior distribution.

Figure 6 documents the US results, using the 1983:Q1-2022:Q4 sample. Panel (a) reports selected historical shock decompositions for inflation while the deterministic components are plotted in panel (b).

The modal value of δ returned by the algorithm of Giannone et al. (2015) is very low (0.0001), and its posterior dispersion is tiny (see Figure A-2 in the online appendix). Put it differently, although the prior is centered around a high value of δ, US data favors a

specification where a-posteriori virtually all VAR models share the same unconditional mean value, as shown in panel b) of Figure 6. This highlights the marginal likelihood gains when δ is low, as discussed earlier.

With the modal value we obtain, model draws featuring similar impulse responses, now also imply similar historical shock decompositions. As an example, the three draws closest to the pointwise median impulse response now produce a coherent narrative about post-COVID inflation drivers. As shown in panel a) of Figure 6, supply factors were important in the initial phase of the inflation surge, but demand factors became the main drivers since 2021. They account for 56 percent of the inflation surge in 2021 and 77 percent in 2022. On average over the post-pandemic period, demand factors explain over 50 percent of the movements in inflation.

We have estimated numerous specifications to check the robustness of these results. Here we briefly discuss what we obtain in two such exercises. First, given the COVID outliers, we stopped estimation in 2019 and computed historical decompositions for the 2021-2022 period with the estimates we obtained. Without our prior, the deterministic components are very dispersed and the historical decomposition of inflation whimsical (see Figure A-3 in the appendix). The problem disappears when the single-unit-root restriction is imposed and the result agree with our full sample estimates. Thus, the extreme volatility of output growth due to COVID is not crucial for our conclusions. Second, we have estimated a five variable VAR as in section 2.3.4, and identified three demand shocks (residual demand, monetary, and investment), and two supply shocks (labor supply and productivity). With our prior in place, we calculated the combine share of fluctuations due to the three demand shocks in 2021 and 2022. We found that while in the first two quarters of 2021 supply shocks are important, demand shocks dominate afterward, see **??** in the appendix for more details. In particular, in line with conclusions of the two variable VAR, the three demand shocks explain over 60 percent of the variability of inflation from the middle of 2021 to the end of 2022.

A final word of caution is in order. While the single-unit-root prior eliminates a large portion of the uncertainty around the deterministic component, uncertainty about the point estimates of historical decompositions remains (see Read (2024)). However, this uncertainty is aligned with the one of impulse responses. Once the uncertainty around the deterministic component is reduced, a researcher who wants to reduce the residual uncertainty around historical decompositions should use the same methods employed to reduce uncertainty around impulse responses (see Antolín-Díaz and Rubio-Ramírez (2018), Arias et al. (2019), Baumeister and Hamilton (2015), and Kilian and Murphy (2012) among others).

5.2 Discussion

Our results are consistent with recent papers supporting the prevalent, if not dominant, role of demand shocks for the recent US inflation surge. This is the case when New Keynesian models with Phillips curve nonlinearities (Benigno and Eggertsson (2023), Harding et al. (2023) and Schmitt-Grohé and Uribe (2023)) and multiple industries models (Rubbo, 2023) are used. It is also the case when factor models (Eickmeier and Hofmann, 2022) and SVAR models similar to ours (Giannone and Primiceri, 2024) are estimated.

A few papers have tried to disentangle more granular drivers within the broad category

of demand shocks. The spending of excess savings accumulated during the pandemic is found to play a major role in Bardóczy et al. (2024) using a HANK model. Accommodative monetary policy is a major force in Giannone and Primiceri (2024), Gagliardone and Gertler (2023), Bocola et al. (2024), and Comin et al. (2023). A number of papers have also emphasized the role of fiscal policy in the experience. Bianchi et al. (2023) show that unfunded fiscal shocks caused a persistent inflation increase, Di Giovanni et al. (2023a) discover an important impact of fiscal policy on current inflation using a multi-sector model with a network structure, while Mori (2024) finds a major role for fiscal policy in driving the inflation surge in a SVAR model that disentangle the three demand shocks just reviewed.

The prevalent role of shocks to prices given wages highlighted by Bernanke and Blanchard (2024) is not necessarily in contrast with our outcomes. As an example, high oil prices (the most important shock to prices considered by Bernanke and Blanchard (2024)) can reflect supply disruptions but also the strong recovery in world demand. Likewise, fluctuations in labor market tightness, which are of second order in Bernanke and Blanchard (2024), but of crucial in Ball et al. (2022), Benigno and Eggertsson (2023) and Ascari et al. (2024b), are not necessarily driven only by demand factors.

While a consensus seems to emerge on the prevalence of demand shocks, it is important to stress that some papers find a more balanced split between demand and supply forces. Gagliardone and Gertler (2023) highlight the role of oil shocks (in combination with loose monetary policy) [4] Shapiro (2024) constructs sectoral-based decompositions and finds a slightly prevalent role for supply forces. This is the case also in the Phillips curve analysis of Crump et al. (2024). A strong support for supply chain disruptions is found in Bai et al. (2024). Finally, Beaudry et al. (2024) document the amplification of broad-based supply shocks in presence of deviations from rational expectations.

5.3 Inflation in other countries

Although the recent inflation surge is a global phenomenon, it is far from obvious that the role of demand and supply factors across countries should be similar [5] . For example, the fiscal stimulus during the COVID pandemic has been particularly pronounced in the US. In addition, European countries have been more exposed to the economic consequences of the Russian invasion of Ukraine. It is, therefore, of independent interest to estimate a SVAR model using data from other countries and we start with the euro area.

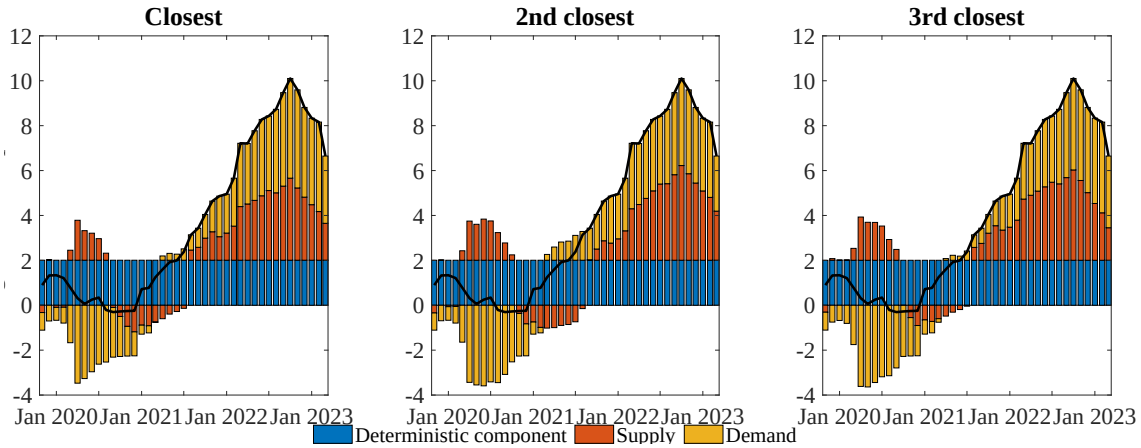
Because the sample for the euro area is shorter, we use data on industrial production (rather than on GDP) and HICP inflation at a monthly frequency in the exercise. The sample period goes from 2001:M1 to 2023:M3. As with US data, the model estimated with a diffuse prior features a large dispersion of deterministic components (see Appendix F). If anything, the problem is exacerbated. Yet, when the model is estimated with the single-unit-root prior, the uncertainty around the deterministic component is substantially reduced, as shown in Figure 7. Notably, the modal value of δ is as low as with US data.

[4] Giannone and Primiceri (2024) make the point that supply shocks cannot produce much inflation if the monetary policy authority is a strict inflation targeter. For inflation to climb, monetary policy must provide an unusually high degree of accommodation relative to the historical rule.

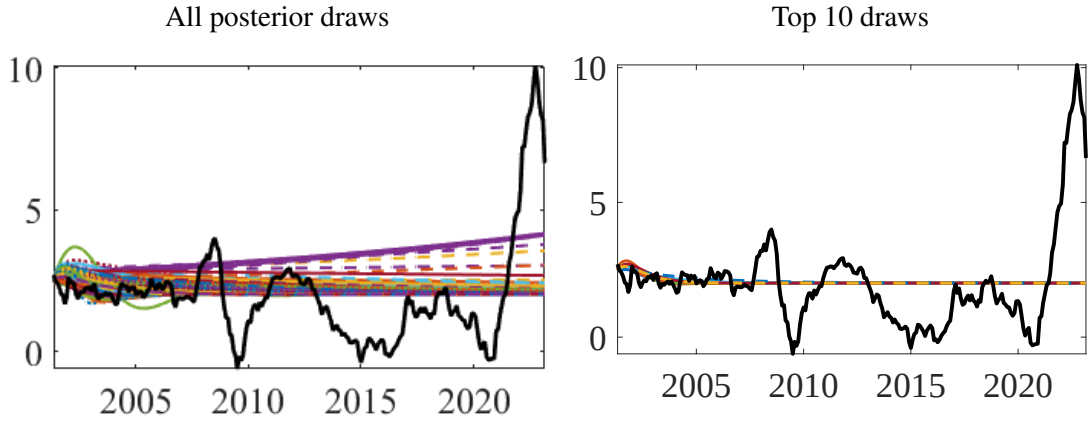
[5] The separate role of domestic and international shocks is studied in Aastveit et al. (2024), Di Giovanni et al. (2023b), Forbes et al. (2024) and Friis et al. (2025). In general, a relevant role for domestic demand forces emerges also once controlling for international factors.

Figure 7: Euro area estimates using the single-unit-root prior.

(a) Historical decomposition of inflation for top 3 draws



(b) Deterministic components of inflation



Note: The black line in all the panels shows year-on-year HICP inflation in the euro area. Panel (a) displays the historical shock decompositions of inflation from the 3 draws closest to the pointwise median response for the IRFs. The colored lines in panel (b) display the deterministic components of inflation for draws from the posterior distribution.

Demand and supply factors now contributed more or less equally to the recent inflation surge, with a more prevalent role for demand factors in 2022. Thus, while supply factors played a role only until mid 2022 in the US, they explain a substantial share of inflation in the euro area also in the latest part of the sample. This outcome is far from surprising, as the US are largely insulated from the consequences of the Ukraine war. Nonetheless, demand factors are prevalent also in the euro area, at least since the second quarter of 2022. This agrees with the results of Ascari et al. (2023). More recently, Giannone and Primiceri (2024) confirm our finding and highlight the role of monetary policy while Ascari et al. (2024a) show that fiscal policy also contributed to the surge. On the other hand, Bańbura et al. (2023) and De Santis (2024) find a prevalent role for supply shocks in empirical specifications that allow for two or more supply shocks.

To provide complementary evidence, we also estimate the same SVAR model using data for four small open economies: Norway, Sweden, Canada and Australia. In this case,

we use quarterly data on real GDP growth and year-on-year CPI inflation for the period 1993:Q1–2023:Q2. The single unit-root-prior shrinks the uncertainty around the deterministic component for all countries, as illustrated in Figure 8b and in all cases the modal value of δ is low. As shown in Figure 8a, the role of demand factors is everywhere prevalent, if not dominant. Thus, despite the large heterogeneity in the policy responses to the COVID 19 pandemic, the different exposure to supply bottlenecks and to the Ukraine war, demand factors seem to drive the recent inflation surge in all the countries we considered.

A broader international comparison is performed in Barro and Bianchi (2024). They use data for OECD countries (20 nations plus the euro area) and conclude that the fiscal expansion following the COVID pandemic has been a key driver of inflation rates. Our result that demand shocks are prevalent in all countries is very much consistent with their findings although the sample, the estimation procedure, and the approach employed in the empirical analysis are very different.

6 Solving excess dispersion: Alternative approaches

For those researchers who are reluctant to use the single-unit-root prior for inferential purposes, we offer two alternative pragmatic solutions that can go a long way in making the dispersion in deterministic components less relevant for practical purposes. The first is to demean the data prior to estimation. Such an approach partially works since poor estimates of the VAR constant contribute to make historical decompositions whimsical. The second is to construct a *median* historical decomposition taking parameter uncertainty into account.

6.1 Demeaning the data

As discussed in Section 2, the deterministic component consists of two terms:

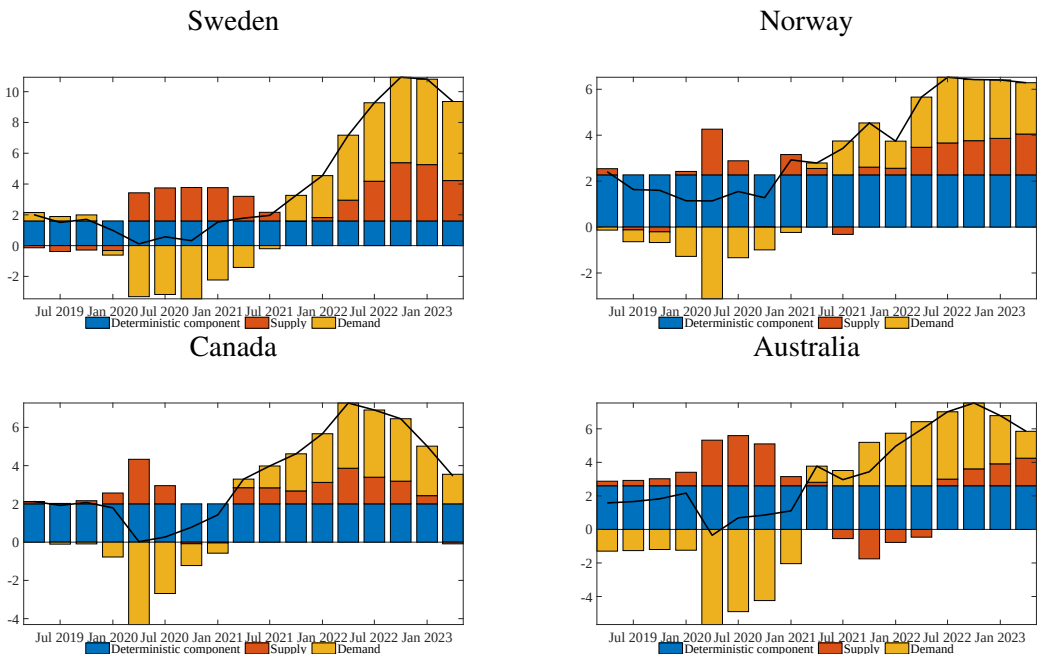
$$(\mathbf{I} + \mathbf{A} + \mathbf{A}^2 + \cdots + \mathbf{A}^{t-1})\mathbf{C} + \mathbf{A}^t\mathbf{Y_0}$$

If the VAR constant is zero, the first term disappears and, as t increases, the last term goes towards zero, as long as the data is stationary and ergodic. Thus, a simple way to restrict the historical decomposition across draws to be consistent with the impulse responses is to demean the data and estimate the VAR without a constant. This approach is equivalent to our single-unit-root prior setup when $\delta \to 0$ and $\bar{\mathbf{Y}}_0 = 0$.

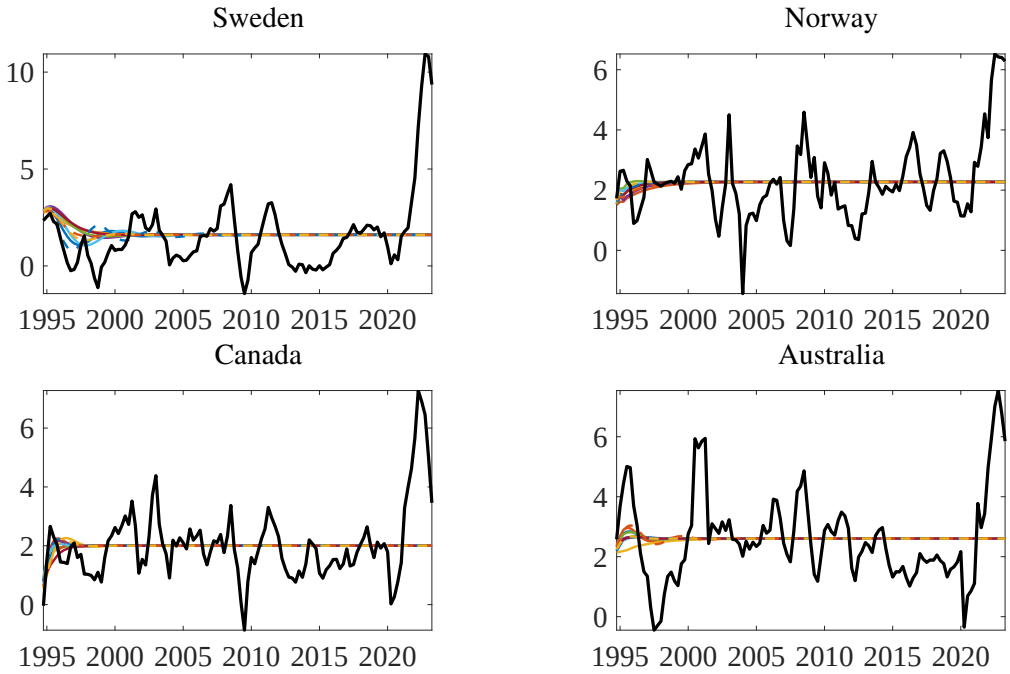
We present the historical decomposition of US inflation with demeaned data in Figure 9. In this specification, the role of demand shocks is still evident and slightly more prevalent than with a single-unit-root prior. Note, however, that the uncertainty in the deterministic component is larger than in Figure 6. By demeaning and estimating the VAR without a constant, the deterministic component is forced to be on the path described by $\mathbf{A}^t\mathbf{Y}_0$. Therefore, draws that imply different $\mathbf{A}$'s will force the estimated structural shocks to take different paths, potentially altering the narrative in a given historical episode.

Figure 8: Selected countries estimates using the single-unit-root prior.

(a) Historical decompositions of inflation



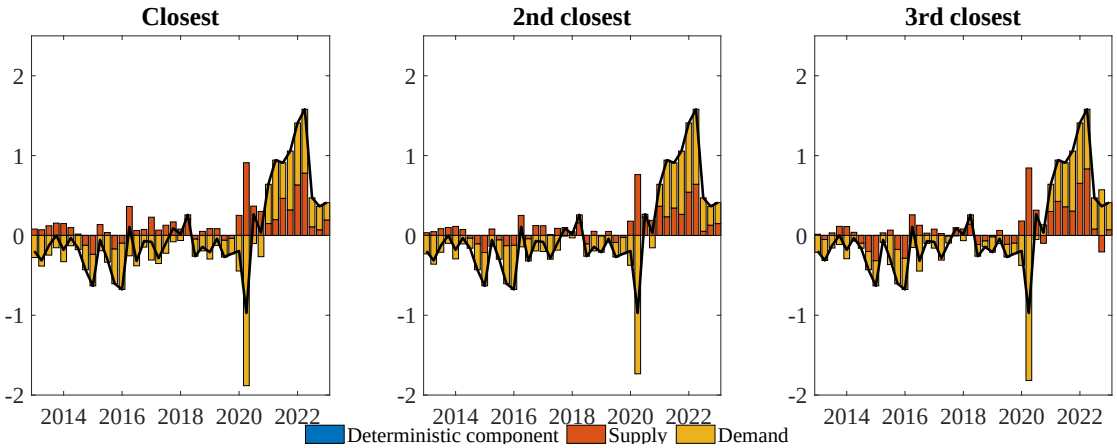
(b) Deterministic components of inflation. Top 10 draws.



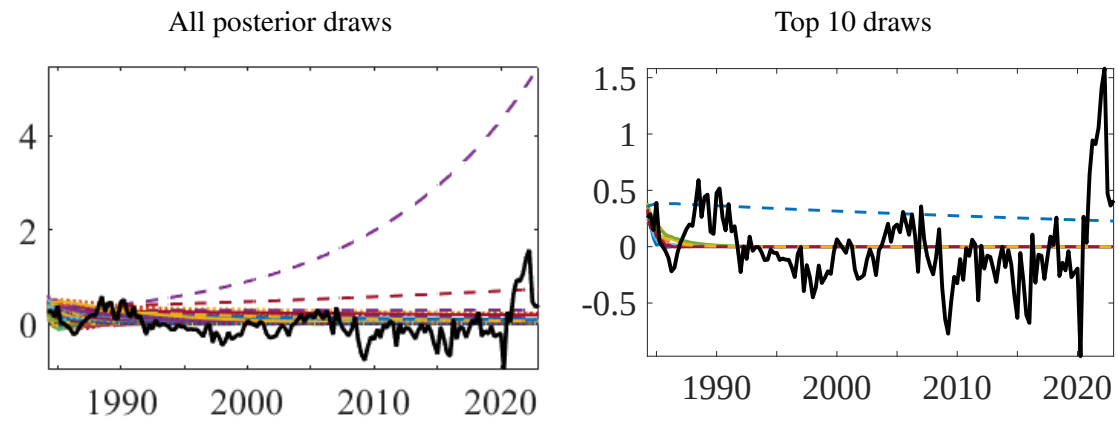
Note: The black line in all the panels shows year-on-year CPI inflation. Panel (a) displays the historical shock decompositions of inflation from the draw closest to the pointwise median response for the IRFs. The colored lines in panel (b) display the estimated deterministic components of inflation from the ten draws closest to the pointwise median response based on the IRFs.

Figure 9: US estimates without constant (demeaned data) and a diffuse prior.

(a) Historical decomposition of inflation for top 3 draws



(b) Deterministic components of inflation



Note: The black line in all the panels shows the de-meaned quarterly growth rate of the GDP deflator. Panel (a) displays the historical shock decompositions of inflation from the 3 draws closest to the pointwise median response for the IRFs. The colored lines in panel (b) display the deterministic components of inflation for draws from the posterior distribution.

6.2 Computing a median historical decomposition

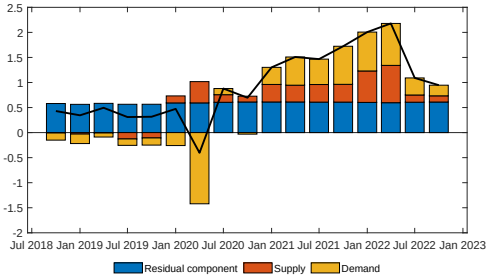
Another pragmatic approach consists in acknowledging that the estimated deterministic components are uncertain and take the information into account when reporting historical decompositions. To do this, one could compute a large number of historical decompositions, each associated with a draw for the VAR parameters and, depending on the loss function employed, construct either the pointwise mean or the pointwise median historical decomposition at each time. Here we prefer the second, even though it does not preserve additivity, since certain draws of the parameters imply explosive systems or very erratic historical decomposition, especially when a flat prior is used. Hence, at each quarter, the data is decomposed into the median contribution of demand shocks, the median contribution of supply shocks and a residual component that absorbs the difference between data and the two median stochastic components, see also Bergholt et al. (2025).

Figure 10 presents the median historical decomposition for the four prior distributions

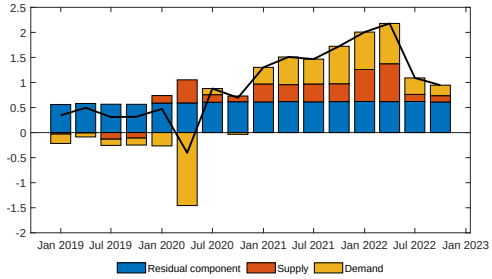
we considered using US data. Three facts are evident. First, the results are quite similar across specifications. Second, demand shocks have an important role during the current inflation surge in all cases. Third, by comparing the bottom-right panel of Figure 10 and panel a) in Figure 6 one can conclude that the pointwise median historical decomposition and the historical decomposition obtained with the single-unit-root prior provide a similar reading of the current inflation experience.

Figure 10: Historical decomposition of inflation, pointwise median.

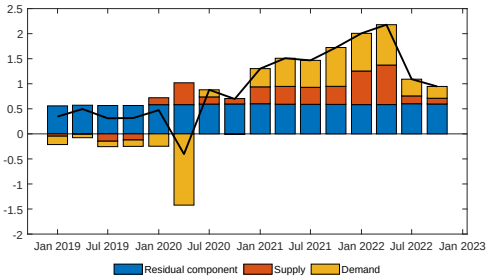
(a) Diffuse prior



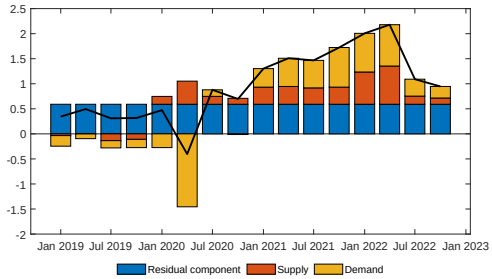
(b) Normal-Inverse Wishart prior



(c) Minnesota prior



(d) Single-unit-root prior



Note: This figure displays the pointwise median historical shock decompositions of inflation for different priors together with the quarterly growth rate of the GDP deflator (black line).

7 Conclusion

The paper shows that the uncertainty around the deterministic component in SVAR models is generally large in standard specifications. This means that draws producing similar impulse responses amy feature vastly different deterministic components. In turn, draws associated with extreme deterministic components exhibit a distorted historical decomposition because shocks have to match the gap between the variable of interest and a (highly uncertain) estimated deterministic component. Hence, there is no guarantee that "good" impulse response functions are associated with "good" historical decompositions.

We provided a solution to the problem based on the use of the single-unit-root prior. We also suggest two pragmatic alternatives: demeaning the data and using a pointwise median historical decomposition. Regardless of the approach one uses, we find that the recent US inflation surge is mainly driven by demand factors. The prevalence of demand

shocks is confirmed, and in some cases reinforced, when using data for Norway, Sweden, Canada and Australia. The contribution is generally more balanced in the euro area where supply factors are still relatively important at the end of our sample.

The problems we highlight also appear when one wants to separate permanent from transitory components. The excess dispersion problem becomes even more acute because variations in the deterministic components matter for the level of the permanent component. Thus, a large uncertainty in the former translates into a large uncertainty in the latter. In addition, since in small samples, the long run contribution of the shocks is imprecisely estimated, additional uncertainty may affect what is perceived as permanent and as transitory. This makes estimates of, e.g., trend inflation generally poor and inflation regimes hard to characterize. These issues are currently investigated in Canova and Fosso (2024).

REFERENCES

Aastveit, K. A., Bjørnland, H. C., Cross, J. L., and Kalstad, H. O. (2024). Unveiling inflation: Oil shocks, supply chain pressures, and expectations. Norges Bank Working Paper 12/2024.

Antolin-Diaz, J., Petrella, I., and Rubio-Ramírez, J. F. (2021). Dividend momentum and stock return predictability: A Bayesian approach. CEPR Discussion Paper 16613.

Antolín-Díaz, J. and Rubio-Ramírez, J. F. (2018). Narrative sign restrictions for SVARs. *American Economic Review*, 108(10):2802–29.

Arias, J. E., Caldara, D., and Rubio-Ramírez, J. F. (2019). The systematic component of monetary policy in svars: An agnostic identification procedure. *Journal of Monetary Economics*, 101:1–13.

Ascari, G., Bonam, D., Mori, L., and Smadu, A. (2024a). Fiscal policy and inflation in the euro area. CEPR Discussion Paper 19683.

Ascari, G., Bonomolo, P., Hoeberichts, M., and Trezzi, R. (2023). The euro area great inflation surge. *SUERF Policy Brief 548*.

Ascari, G., Grazzini, J., and Massaro, D. (2024b). Great layoff, great retirement and post-pandemic inflation. De Nederlandsche Bank Working Paper 812.

Bai, X., Fernández-Villaverde, J., Li, Y., and Zanetti, F. (2024). The causal effects of global supply chain disruptions on macroeconomic outcomes: evidence and theory. NBER Working Paper 32098.

Ball, L. M., Leigh, D., and Mishra, P. (2022). Understanding US inflation during the COVID era. *Brookings Papers on Economic Activity*, Fall 2022:1–54.

Bańbura, M., Bobeica, E., and Martínez Hernández, C. (2023). What drives core inflation? The role of supply shocks. ECB Working Paper 2875.

Bardóczy, B., Sim, J., and Tischbirek, A. (2024). The macroeconomic effects of excess savings. FEDS Working Paper 2024-62.

Barro, R. and Bianchi, F. (2024). Fiscal influences on inflation in OECD countries, 2020-2023. NBER Working Paper 31838.

Baumeister, C. and Hamilton, J. D. (2015). Sign restrictions, structural vector autoregressions, and useful prior information. *Econometrica*, 83(5):1963–1999.

Beaudry, P., Hou, C., and Portier, F. (2024). The dominant role of expectations and broad-based supply shocks in driving inflation. CEPR Discussion Paper 18963.

Benigno, P. and Eggertsson, G. B. (2023). It's baaack: The surge in inflation in the 2020s and the return of the non-linear Phillips curve. NBER Working Paper 31197.

Bergholt, D., Furlanetto, F., and Vaccaro-Grange, E. (2025). Did monetary policy kill the Phillips curve? Some simple arithmetics. *Review of Economics and Statistics, forthcoming*.

Bernanke, B. and Blanchard, O. (2024). What caused the US pandemic-era inflation? *American Economic Journal: Macroeconomics, forthcoming*.

Bianchi, F., Faccini, R., and Melosi, L. (2023). A fiscal theory of persistent inflation. *Quarterly Journal of Economics*, 138(4):2127–2179.

Blanchard, O. J. and Quah, D. (1989). The dynamic effects of aggregate demand and supply disturbances. *American Economic Review*, 79(4):655–673.

Bocola, L., Dovis, A., Jørgensen, K., and Kirpalani, R. (2024). Bond market views of the Fed. NBER Working Paper 32620.

Canova, F. and De Nicolò, G. (2002). Monetary disturbances matter for business cycle fluctuations in the G7. *Journal of Monetary Economics*, 46:1131–1159.

Canova, F. and Ferroni, F. (2022). Mind the gap! Stylized dynamic facts and structural models. *American Economic Journal: Macroeconomics*, 14(4):104–35.

Canova, F. and Fosso, L. (2024). Observing the stars to understand the gaps: An empirical laboratory. Manuscript.

Carriero, A., Clark, T. E., and Marcellino, M. (2015). Bayesian VARs: specification choices and forecast accuracy. *Journal of Applied Econometrics*, 30(1):46–73.

Chan, J. C. (2022). Asymmetric conjugate priors for large Bayesian VARs. *Quantitative Economics*, 13(3):1145–1169.

Comin, D. A., Johnson, R. C., and Jones, C. J. (2023). Supply chain constraints and inflation. NBER Working Paper 31179.

Crump, R. K., Eusepi, S., Giannoni, M., and Şahin, A. (2024). The unemployment–inflation trade-off revisited: The Phillips curve in COVID times. *Journal of Monetary Economics*, 145:103580.

De Santis, R. A. (2024). Supply chain disruption and energy supply shocks: Impact on euro-area output and prices. *International Journal of Central Banking*, 20(2):193–235.

Del Negro, M. and Schorfheide, F. (2011). Bayesian Macroeconometrics. In *The Oxford Handbook of Bayesian Econometrics*. Oxford University Press.

Di Giovanni, J., Kalemli-Ozcan, S., Silva, A., and Yildirim, M. A. (2023a). Quantifying the inflationary impact of fiscal stimulus under supply constraints. NBER Working Paper 30892.

Di Giovanni, J., Kalemli-Özcan, Silva, A., and Yildirim, M. A. (2023b). Pandemic-era inflation drivers and global spillovers. NBER Working Paper 31887.

Doan, T., Litterman, R., and Sims, C. (1984). Forecasting and conditional projection using realistic prior distributions. *Econometric Reviews*, 3(1):1–100.

Eickmeier, S. and Hofmann, B. (2022). What drives inflation? Disentangling demand and supply factors. Deutsche Bundesbank Discussion Paper 46/2022.

Forbes, K., Ha, J., and Kose, M. A. (2024). Rate cycles. CEPR Discussion Paper 19272.

Friis, I., Furlanetto, F., Matsen, K., and Robstad, Ø. (2025). Understanding the prevalence of demand shocks in the recent inflation surge: An international comparison. Norges Bank, manuscript.

Fry, R. and Pagan, A. (2011). Sign restrictions in structural vector autoregressions: A critical review. *Journal of Economic Literature*, 49(4):938–960.

Furlanetto, F., Lepetit, A., Robstad, Ø., Rubio Ramírez, J., and Ulvedal, P. (2025). Estimating hysteresis effects. *American Economic Journal: Macroeconomics*, 17(1):35–70.

Gagliardone, L. and Gertler, M. (2023). Oil prices, monetary policy and inflation surges. NBER Working Paper 31263.

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

Giannone, D., Lenza, M., and Primiceri, G. E. (2019). Priors for the long run. *Journal of the American Statistical Association*, 114(526):565–580.

Giannone, D. and Primiceri, G. (2024). The drivers of post-pandemic inflation. NBER Working Paper 32859.

Harding, M., Lindé, J., and Trabandt, M. (2023). Understanding post-covid inflation dynamics. *Journal of Monetary Economics*, 140:S101–S118.

Inoue, A. and Kilian, L. (2013). Inference on impulse response functions in structural VAR models. *Journal of Econometrics*, 177(1):1–13.

Jeffreys, H. (1946). An invariant form for the prior probability in estimation problems. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 186(1007):453–461.

Kabaca, S. and Tuzcuoglu, K. (2023). Supply drivers of US inflation since the pandemic. Bank of Canada Staff Working Paper 2023-19.

Karlsson, S. (2013). Forecasting with Bayesian autoregressions. *Handbook of Economic Forecasting*, (2):791–897.

Kilian, L. and Lütkepohl, H. (2017). *Structural Vector Autoregressive Analysis*. Cambridge University Press.

Kilian, L. and Murphy, D. P. (2012). Why agnostic sign restrictions are not enough: Understanding the dynamics of oil market VAR models. *Journal of the European Economic Association*, 10(5):1166–1188.

Koop, G. and Korobilis, D. (2013). Large time-varying parameter VARs. *Journal of Econometrics*, 177(2):185–198.

Miranda-Agrippino, S. and Ricco, G. (2019). Bayesian vector autoregressions: Estimation. In *Oxford Research Encyclopedia of Economics and Finance*.

Mori, L. (2024). Fiscal shocks and the surge in inflation. Manuscript.

Read, M. (2024). Sign restrictions and supply-demand decompositions of inflation. Reserve Bank of Australia Working Paper 2024-05.

Rubbo, E. (2023). What drives inflation? Lessons from disaggregated price data. Manuscript.

Schmitt-Grohé, S. and Uribe, M. (2023). Heterogeneous downward nominal wage rigidity: Foundations of a nonlinear Phillips curve. NBER Working Paper 30774.

Shapiro, A. H. (2024). Decomposing supply and demand driven inflation. *Journal of Money Credit and Banking, forthcoming*.

Sims, C. A. (1986). Are forecasting models usable for policy analysis? *Federal Reserve Bank of Minneapolis Quarterly Review*, 10:2–16.

Sims, C. A. (1993). A nine-variable probabilistic macroeconomic forecasting model. In *Business cycles, indicators, and forecasting*, pages 179–212. University of Chicago press.

Sims, C. A. (1996). Inference for multivariate time series models with trend. Manuscript.

Sims, C. A. (2000). Using a likelihood perspective to sharpen econometric discourse: Three examples. *Journal of Econometrics*, 95(2):443–462.

Sims, C. A. and Zha, T. (1998). Bayesian methods for dynamic multivariate models. *International Economic Review*, 39:949–968.

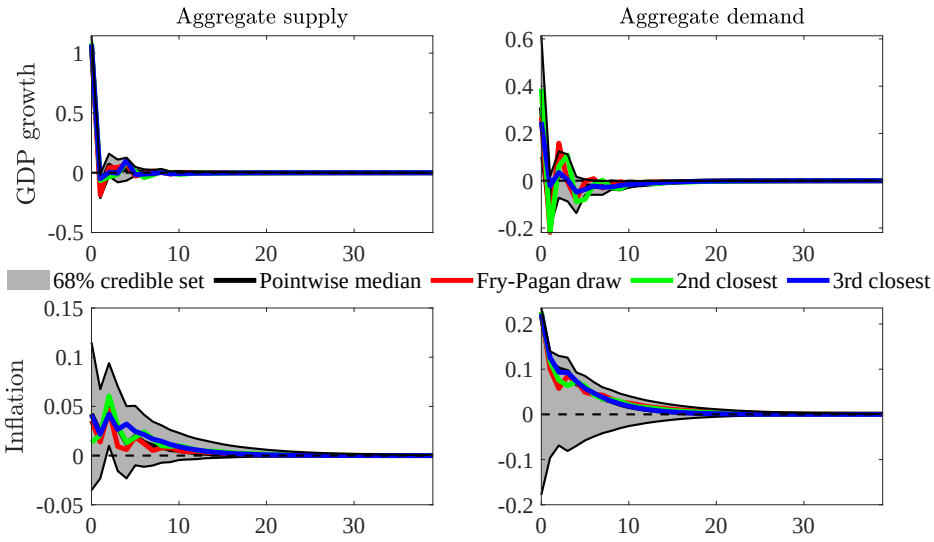
Villani, M. (2009). Steady state priors for vector autoregressions. *Journal of Applied Econometrics*, 24(4):630–650.

ONLINE APPENDIX

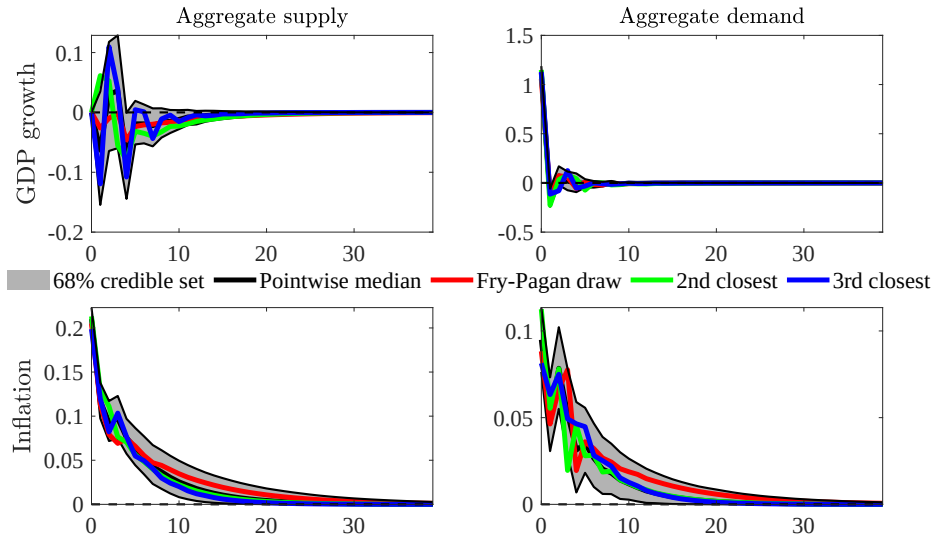
A ADDITIONAL RESULTS FOR THE US

Figure A-1: IRFs, different identification schemes

(a) Blanchard-Quah



(b) Cholesky



Note: The black line is the pointwise median and the shaded areas cover the 68 percent and min-max identified sets. The red line is the impulse response for the draw closest to the pointwise median; the green and blue lines are impulse responses for the 2nd and 3rd draws closest to the pointwise median, respectively.

Figure A-2: Posterior distribution of δ, US data

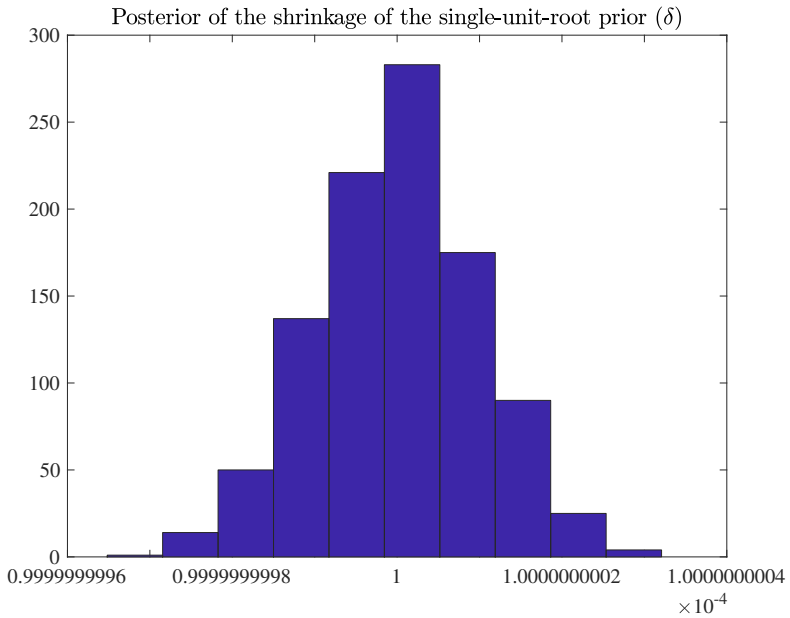
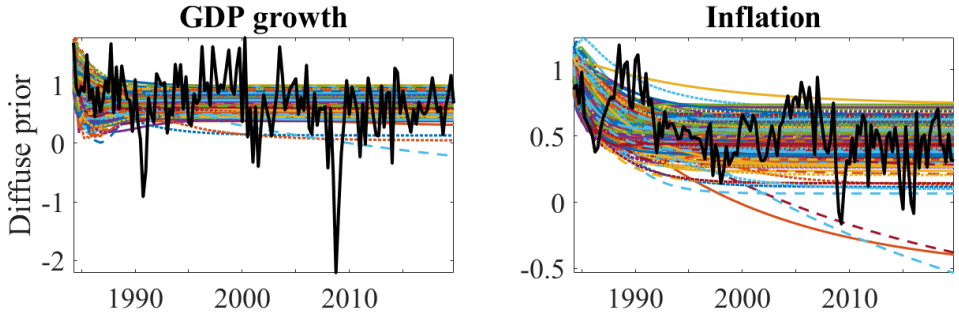


Figure A-3: Estimated deterministic components, diffuse prior, 1983-2019



Note: The figure reports the deterministic components for all draws from the posterior distribution.

B ANALYTICAL DERIVATIONS

This appendix provides the details of the analytical derivations of the quantities presented in Section 3.1. First, we derive an expression for the mean squared error of the constant term $\mathbf{C}$. Using the VAR model, the mean squared error of $\hat{\mathbf{C}}$ is:

$$\begin{aligned}
\text{MSE}(\hat{\mathbf{C}}) &= E\Big(\big(\hat{\mathbf{C}}-\mathbf{C}\big)\big(\hat{\mathbf{C}}-\mathbf{C}\big)'\Big) \\
&= E\Big(\big(\bar{\mathbf{u}}-(\hat{\mathbf{A}}-\mathbf{A})\bar{\mathbf{Y}}_{-1}\big)\big(\bar{\mathbf{u}}-(\hat{\mathbf{A}}-\mathbf{A})\bar{\mathbf{Y}}_{-1}\big)'\Big) \\
&= E\Big(\bar{\mathbf{u}}\bar{\mathbf{u}}'-\bar{\mathbf{u}}\bar{\mathbf{Y}}'_{-1}(\hat{\mathbf{A}}-\mathbf{A})'-(\hat{\mathbf{A}}-\mathbf{A})\bar{\mathbf{Y}}_{-1}\bar{\mathbf{u}}'+(\hat{\mathbf{A}}-\mathbf{A})\bar{\mathbf{Y}}_{-1}\bar{\mathbf{Y}}'_{-1}(\hat{\mathbf{A}}-\mathbf{A})'\Big) \\
&= E\big(\bar{\mathbf{u}}\bar{\mathbf{u}}'\big)-E\big(\bar{\mathbf{u}}\bar{\mathbf{Y}}'_{-1}(\hat{\mathbf{A}}-\mathbf{A})'\big)-E\big((\hat{\mathbf{A}}-\mathbf{A})\bar{\mathbf{Y}}_{-1}\bar{\mathbf{u}}'\big) \\
&+ E\big((\hat{\mathbf{A}}-\mathbf{A})\bar{\mathbf{Y}}_{-1}\bar{\mathbf{Y}}'_{-1}(\hat{\mathbf{A}}-\mathbf{A})'\big) \qquad \text{(B-1)}
\end{aligned}$$

Since $\mathbf{A}$ can be defined, with data in deviation from the sample mean, as $\hat{\mathbf{A}} = \mathbf{A} + \big(\sum_{t=1}^{T}\tilde{\mathbf{Y}}_{t-1}\tilde{\mathbf{Y}}'_{t-1}\big)^{-1}\big(\sum_{t=1}^{T}\tilde{\mathbf{Y}}_{t-1}\tilde{\mathbf{u}}'_t\big)$ we can write the mean squared error of $\hat{\mathbf{C}}$ as:

$$\begin{aligned}
\text{MSE}(\hat{\mathbf{C}}) &= E\big(\bar{\mathbf{u}}\bar{\mathbf{u}}'\big)-E\big(\bar{\mathbf{u}}\bar{\mathbf{Y}}'_{-1}\big)\text{Bias}(\hat{\mathbf{A}})'-\text{Bias}(\hat{\mathbf{A}})E\big(\bar{\mathbf{Y}}_{-1}\bar{\mathbf{u}}'\big)+\text{Bias}(\hat{\mathbf{A}})E\big(\bar{\mathbf{Y}}_{-1}\bar{\mathbf{Y}}'_{-1}\big)\text{Bias}(\hat{\mathbf{A}})' \\
&\approx \frac{1}{T}\boldsymbol{\Sigma}_{\mathbf{u}}+\text{Bias}(\hat{\mathbf{A}})E\big(\bar{\mathbf{Y}}_{-1}\bar{\mathbf{Y}}'_{-1}\big)\text{Bias}(\hat{\mathbf{A}})' \\
&\approx \frac{1}{T}\boldsymbol{\Sigma}_{\mathbf{u}}+\text{Bias}(\hat{\mathbf{A}})\Big(\frac{1}{T}\boldsymbol{\Sigma}_{\mathbf{Y}}+\boldsymbol{\mu}\boldsymbol{\mu}'\Big)\text{Bias}(\hat{\mathbf{A}})' \qquad \text{(B-2)}
\end{aligned}$$

where $\bar{\mathbf{u}} = \frac{1}{T}\sum_{t=1}^{T}\mathbf{u}_t$, $\tilde{\mathbf{Y}}_t = \mathbf{Y}_t - \bar{\mathbf{Y}}$, $\bar{\mathbf{Y}} = \frac{1}{T}\sum_{t=1}^{T}\mathbf{Y}_t$, $\boldsymbol{\mu} = (I-\mathbf{A})^{-1}\mathbf{C}$, $\boldsymbol{\Sigma}_{\mathbf{y}} = \text{var}(\mathbf{Y}_{t-1})$. The approximation comes from the fact that the terms $E\big(\bar{\mathbf{u}}\bar{\mathbf{Y}}'_{-1}\big)\text{Bias}(\hat{\mathbf{A}})'$ and $\text{Bias}(\hat{\mathbf{A}})E\big(\bar{\mathbf{Y}}_{-1}\bar{\mathbf{u}}'\big)$ disappear in standard samples as these contain the term $\frac{1}{T^3}$.

Recall the definition of the deterministic component:

$$\begin{aligned}
DC_t &= \big(\mathbf{I}+\mathbf{A}+\cdots+\mathbf{A}^{t-1}\big)\,\mathbf{C}+\mathbf{A}^t\mathbf{Y_0} \\
&= (\mathbf{I}-\mathbf{A})^{-1}\mathbf{C}+\mathbf{A}^t(\mathbf{Y_0}-(\mathbf{I}-\mathbf{A})^{-1}\mathbf{C}) \qquad \text{(B-3)}
\end{aligned}$$

The mean squared error of the estimated deterministic component is:

$$\text{MSE}(\hat{DC_t}) = E\Big(\big(\hat{DC_t}-DC_t\big)\big(\hat{DC_t}-DC_t\big)'\Big) \qquad \text{(B-4)}$$

Focusing on the term $\hat{DC_t}-DC_t$:

$$\begin{aligned}
\hat{DC_t}-DC_t &= (\mathbf{I}-\hat{\mathbf{A}})^{-1}\hat{\mathbf{C}}+\hat{\mathbf{A}}^t(\mathbf{Y_0}-(\mathbf{I}-\hat{\mathbf{A}})^{-1}\hat{\mathbf{C}})-(\mathbf{I}-\mathbf{A})^{-1}\mathbf{C}+\mathbf{A}^t(\mathbf{Y_0}-(\mathbf{I}-\mathbf{A})^{-1}\mathbf{C}) \\
&= \Gamma(\mathbf{A})(\hat{\mathbf{C}}-\mathbf{C})+(\Gamma(\hat{\mathbf{A}})-\Gamma(\mathbf{A}))\hat{\mathbf{C}}+(\hat{\mathbf{A}}^t-\mathbf{A}^t)\mathbf{Y_0} \qquad \text{(B-5)}
\end{aligned}$$

where $\Gamma(\mathbf{A}) = (\mathbf{I}-\mathbf{A}^t)(\mathbf{I}-\mathbf{A})^{-1}$. Substituting (B-5) into (B-4), we obtain:

$$\begin{aligned}
\text{MSE}(\hat{DC_t}) &= E\Big(\big(\hat{DC_t}-DC_t\big)\big(\hat{DC_t}-DC_t\big)'\Big) \\
&= E\Big(\big(\Gamma(\mathbf{A})(\hat{\mathbf{C}}-\mathbf{C})(\hat{\mathbf{C}}-\mathbf{C})'\Gamma(\mathbf{A})'+\Gamma(\mathbf{A})(\hat{\mathbf{C}}-\mathbf{C})\hat{\mathbf{C}}'(\Gamma(\hat{\mathbf{A}})-\Gamma(\mathbf{A}))'
\end{aligned}$$

$$
\begin{aligned}
&+ \Gamma(\mathbf{A})(\hat{\mathbf{C}}-\mathbf{C})\mathbf{Y}_0'(\hat{\mathbf{A}}^t-\mathbf{A}^t)' + (\Gamma(\hat{\mathbf{A}})-\Gamma(\mathbf{A}))\hat{\mathbf{C}}(\hat{\mathbf{C}}-\mathbf{C})'\Gamma(\mathbf{A})' \\
&+ (\Gamma(\hat{\mathbf{A}})-\Gamma(\mathbf{A}))\hat{\mathbf{C}}\hat{\mathbf{C}}'(\Gamma(\hat{\mathbf{A}})-\Gamma(\mathbf{A}))' + (\Gamma(\hat{\mathbf{A}})-\Gamma(\mathbf{A}))\hat{\mathbf{C}}\mathbf{Y}_0'(\hat{\mathbf{A}}^t-\mathbf{A}^t)' \\
&+ (\hat{\mathbf{A}}^t-\mathbf{A}^t)\mathbf{Y}_0(\hat{\mathbf{C}}-\mathbf{C})'\Gamma(\mathbf{A})' + (\hat{\mathbf{A}}^t-\mathbf{A}^t)\mathbf{Y}_0\hat{\mathbf{C}}'(\Gamma(\hat{\mathbf{A}})-\Gamma(\mathbf{A}))' \\
&+ (\hat{\mathbf{A}}^t-\mathbf{A}^t)\mathbf{Y}_0\mathbf{Y}_0'(\hat{\mathbf{A}}^t-\mathbf{A}^t)'\Big) \\
&= E\Big(\big(\Gamma(\mathbf{A})(\hat{\mathbf{C}}-\mathbf{C})(\hat{\mathbf{C}}-\mathbf{C})'\Gamma(\mathbf{A})'\big)\Big) + \text{other terms} \qquad \text{(B-6)} \\
&= \Gamma(\mathbf{A})\text{MSE}(\hat{\mathbf{C}})\Gamma(\mathbf{A})' + \text{other terms} \qquad \text{(B-7)}
\end{aligned}
$$

Here we focus on the term related to $\hat{\mathbf{C}}$. As long as the process is stationary and ergodic, and for the length of a standard samples, $\hat{\mathbf{A}}^t - \mathbf{A}^t$ converges to zero and the equation above reduces to:

$$
\begin{aligned}
\text{MSE}(\hat{DC_t}) &= E\left(\big(\hat{DC_t}-DC_t\big)\big(\hat{DC_t}-DC_t\big)'\right) \\
&\approx E\Big(\big(\Gamma(\mathbf{A})(\hat{\mathbf{C}}-\mathbf{C})(\hat{\mathbf{C}}-\mathbf{C})'\Gamma(\mathbf{A})' + \Gamma(\mathbf{A})(\hat{\mathbf{C}}-\mathbf{C})\hat{\mathbf{C}}'(\Gamma(\hat{\mathbf{A}})-\Gamma(\mathbf{A}))' \\
&+ (\Gamma(\hat{\mathbf{A}})-\Gamma(\mathbf{A}))\hat{\mathbf{C}}(\hat{\mathbf{C}}-\mathbf{C})'\Gamma(\mathbf{A})' + (\Gamma(\hat{\mathbf{A}})-\Gamma(\mathbf{A}))\hat{\mathbf{C}}\hat{\mathbf{C}}'(\Gamma(\hat{\mathbf{A}})-\Gamma(\mathbf{A}))'\Big)
\end{aligned}
\qquad \text{(B-8)}
$$

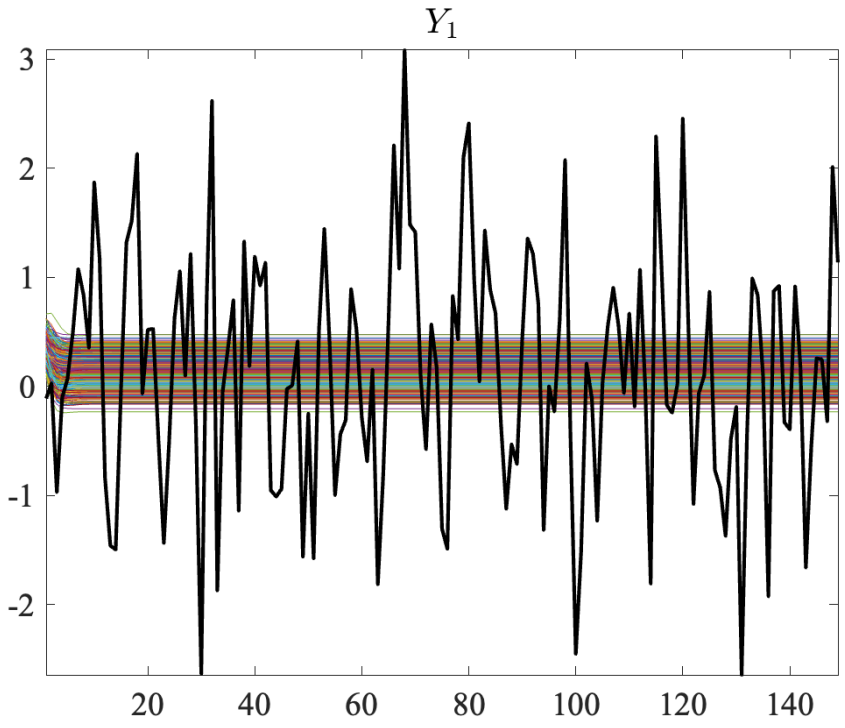
Notice that the terms $\Gamma(\mathbf{A})(\hat{\mathbf{C}}-\mathbf{C})\hat{\mathbf{C}}'(\Gamma(\hat{\mathbf{A}})-\Gamma(\mathbf{A}))'$ and $(\Gamma(\hat{\mathbf{A}})-\Gamma(\mathbf{A}))\hat{\mathbf{C}}(\hat{\mathbf{C}}-\mathbf{C})'\Gamma(\mathbf{A})'$ are negative semi-definite if the bias in $\hat{\mathbf{A}}$ is different from zero, thus reducing the mean squared error of the deterministic component $\text{MSE}(\hat{DC_t})$. This is because the bias in $\hat{\mathbf{C}}$ is of opposite sign with respect to the bias in $\hat{\mathbf{A}}$:

$$
\text{Bias}(\hat{\mathbf{C}}) = E(\hat{\mathbf{C}}-\mathbf{C}) = E(\bar{\mathbf{u}} - (\hat{\mathbf{A}}-\mathbf{A})\bar{\mathbf{Y}}_{-1}) = -\text{Bias}(\hat{\mathbf{A}})\boldsymbol{\mu} \qquad \text{(B-9)}
$$

In general, $\Gamma(\mathbf{A})(\hat{\mathbf{C}}-\mathbf{C})\hat{\mathbf{C}}'(\Gamma(\hat{\mathbf{A}})-\Gamma(\mathbf{A}))'+(\Gamma(\hat{\mathbf{A}})-\Gamma(\mathbf{A}))\hat{\mathbf{C}}(\hat{\mathbf{C}}-\mathbf{C})'\Gamma(\mathbf{A})'+(\Gamma(\hat{\mathbf{A}})-\Gamma(\mathbf{A}))\hat{\mathbf{C}}\hat{\mathbf{C}}'(\Gamma(\hat{\mathbf{A}})-\Gamma(\mathbf{A}))'$ is negative semi-definite. The relevant term for explaining excess dispersion thus is the first term, the mean squared error of $\hat{\mathbf{C}}$. The dispersion unambiguosly increases if the dispersion in the estimate of the constant term increases.

C ADDITIONAL SIMULATION RESULTS

Figure C-1: Deterministic components in simulated data: $a_{1,1} = 0.2, T = 150$.



Note: The figure reports the estimated deterministic components for $Y_{1,t}$ for all draws from the posterior distribution. The black line represents the simulated data.

Figure C-2: Posterior distributions of impulse responses, Minnesota and Minnesota and SUR priors, simulated data.

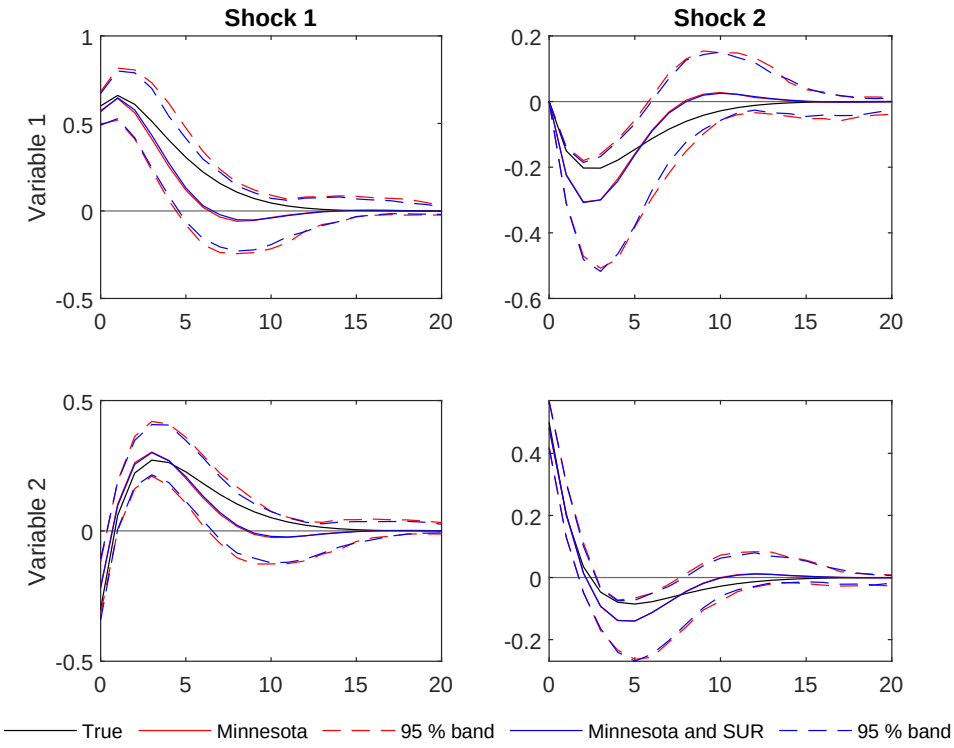
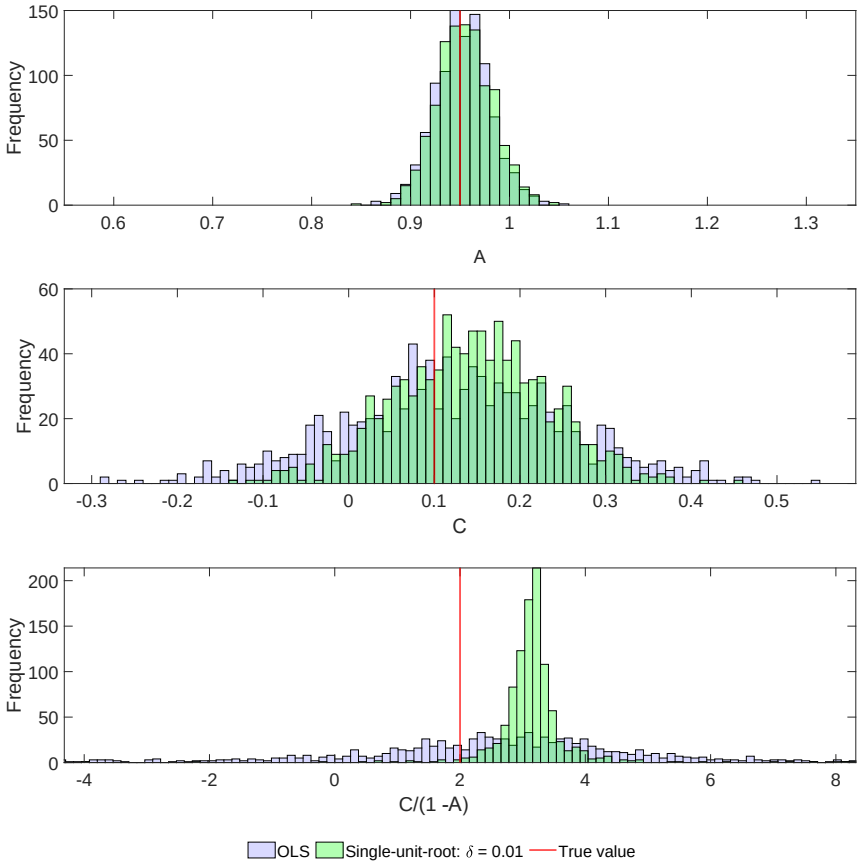


Table C-1: RMSE, Minnesota and Minnesota-SUR priors, simulated data.

h	Minnesota prior		Minnesota and SUR priors	
	Variable 1	Variable 2	Variable 1	Variable 2
1	0.62	0.59	0.61	0.58
4	1.20	0.67	1.20	0.68
8	1.48	0.78	1.47	0.78
20	1.46	0.79	1.46	0.80

*N*ote: Results are reported for the simulation exercise. Out-of-sample performance is calculated using 100 additional observations. Variable 1 is the more persistent (AR1=0.95) and variable 2 is the less persistent (AR1=0.6).

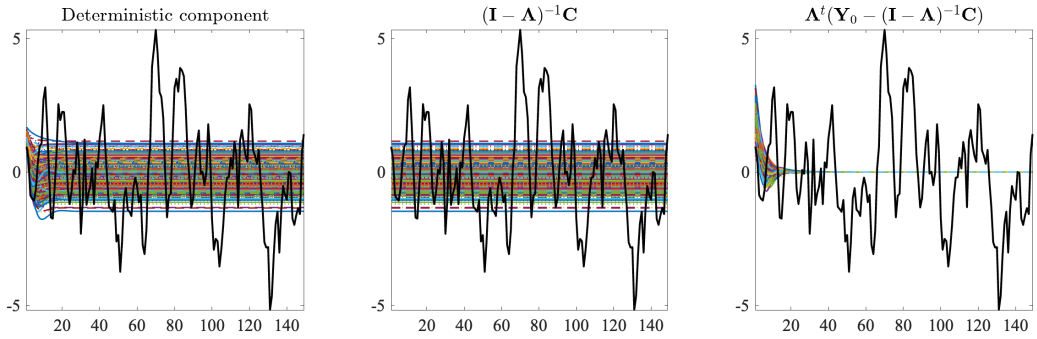
Figure C-3: Posterior distributions, Flat and SUR priors, simulated data.



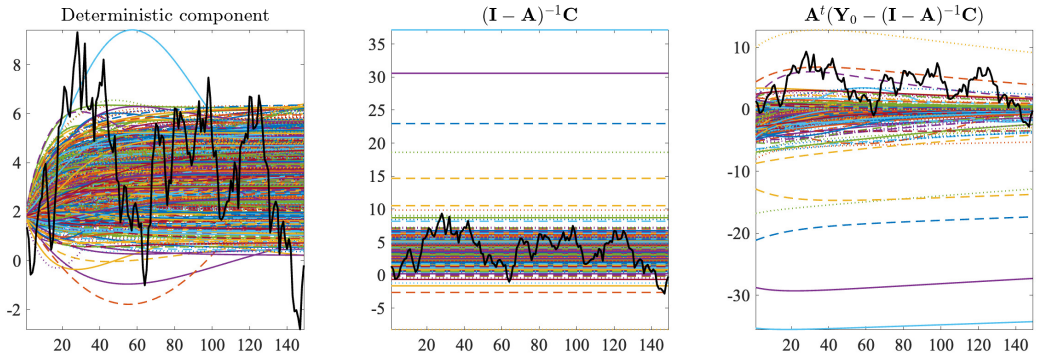
D ANOTHER LOOK AT DISPERSION VS. OVERFITTING IN SIMULATIONS

Figure D-1: Estimated deterministic component of $Y_{1,t}$, different DGP.

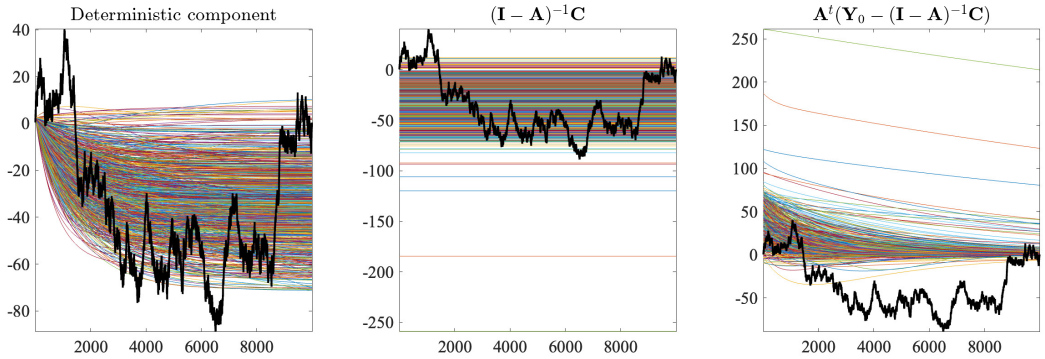
(a) $T = 150$ stationary



(b) $T = 150$ unit root



(c) $T = 10000$ unit root



Note: The figure reports the estimated deterministic components for $Y_{1,t}$ for all draws from the posterior distribution. The black line represents the simulated data.

Figure D-1 provides evidence that shed further light on the relationship between dispersion in the deterministic component and overfitting. We consider three cases. In each case we decompose the deterministic component DC_t (left column) into the steady state term $(\mathbf{I}-\mathbf{A})^{-1}\mathbf{C}$ (middle column) and the transitional dynamic term $\mathbf{A}^t(\mathbf{Y}_0-(\mathbf{I}-\mathbf{A})^{-1}\mathbf{C})$

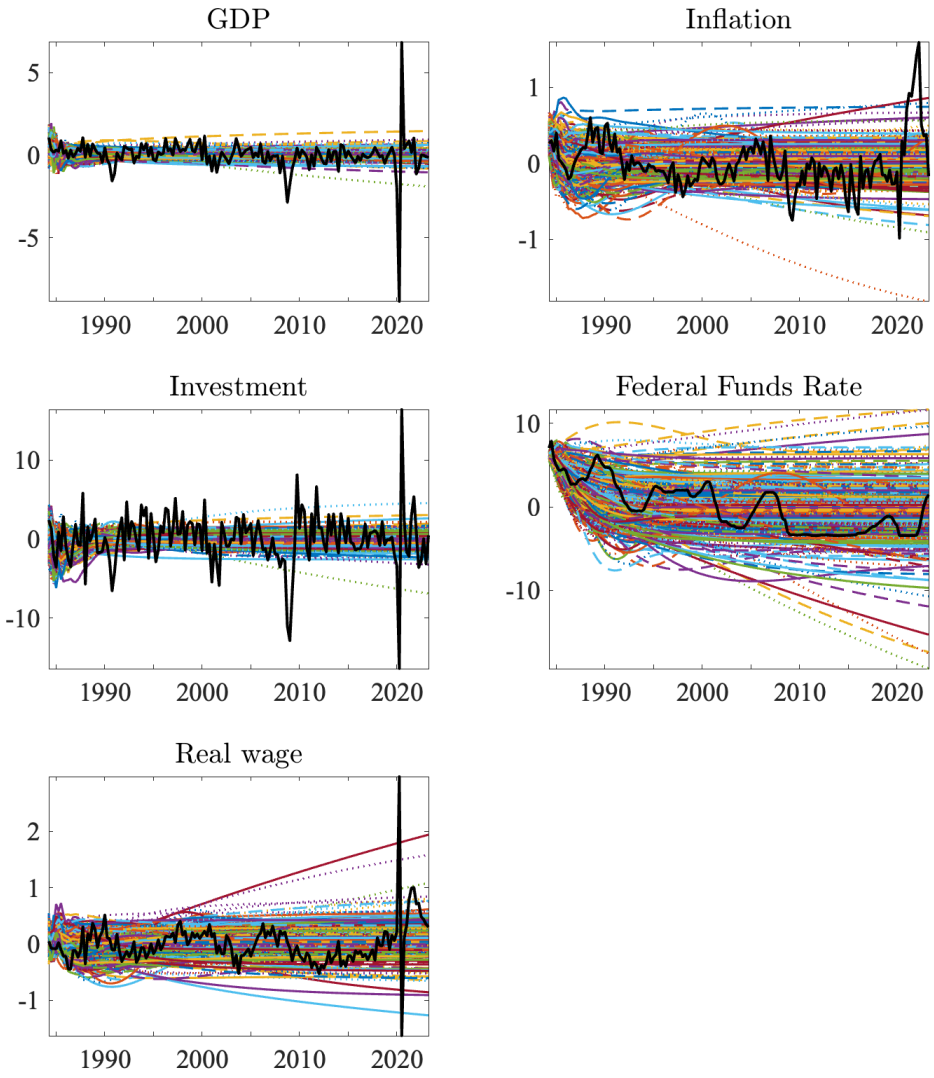
(right column). In the top row we consider the simulation exercise analyzed in the text with stationary data. Two observations stand out: almost all of the dispersion in the deterministic component arises from the steady state term; and the dispersion in the dynamic term is essentially zero after 20 observations. Second, there is no overfitting in this case , as the dynamics in the deterministic component by construction are transitory, and limited to the very beginning of the sample. Thus, the first row in Figure D-1 illustrates how the deterministic component could be excessively dispersed when overfitting is absent.

The second row illustrates a situation in which the true data generating process (DGP) is non-stationary (a random walk without drift). This is the case considered in the Appendix of Giannone et al. (2019). The problem of excess dispersion remains, as can be seen in the middle column. Moreover, the initial difference between $\mathbf{Y}_0$ and $(\mathbf{I}-\mathbf{A})^{-1}\mathbf{C}$ now also produces *transitional dynamics* throughout the sample, as some estimated eigenvalues of $\mathbf{A}$ are close to unity, see third column. Thus, in this setup, both excess dispersion and overfitting may be quantitatively important problems.

The final row in Figure D-1 still considers a random walk without drift but a sample of 10,000 observations is used to estimate the VAR. In this situation small sample issues are absent. Yet, even in this case we observe both excess dispersion, as well as long-lasting transitional dynamics when there are differences between the initial observation and the unconditional mean. Thus, for this data generating process,neither excess dispersion nor overfitting vanish when the sample size becomes very large. Giannone et al. (2019) proved this result for overfitting. Here, we show, via simulation, that it applies also to excess dispersion.

E A LARGER SCALE SVAR

Figure E-1: Deterministic components, diffuse prior, larger VAR



Note: The figure reports the estimated deterministic components for all draws from the posterior distribution for each data series. The black line represent the data series.

In this Appendix we present results for a VAR system with five variables. Under flat priors, the excess dispersion problem is exacerbated, in particular for inflation and interest rates, as shown in Figure E-1.

We then estimate the SVAR with the single-unit-root prior on the same variables. The identification scheme is presented in Table E-1. We identify three demand shocks and two supply shocks. The monetary shock is the only demand shock generating inverse impact

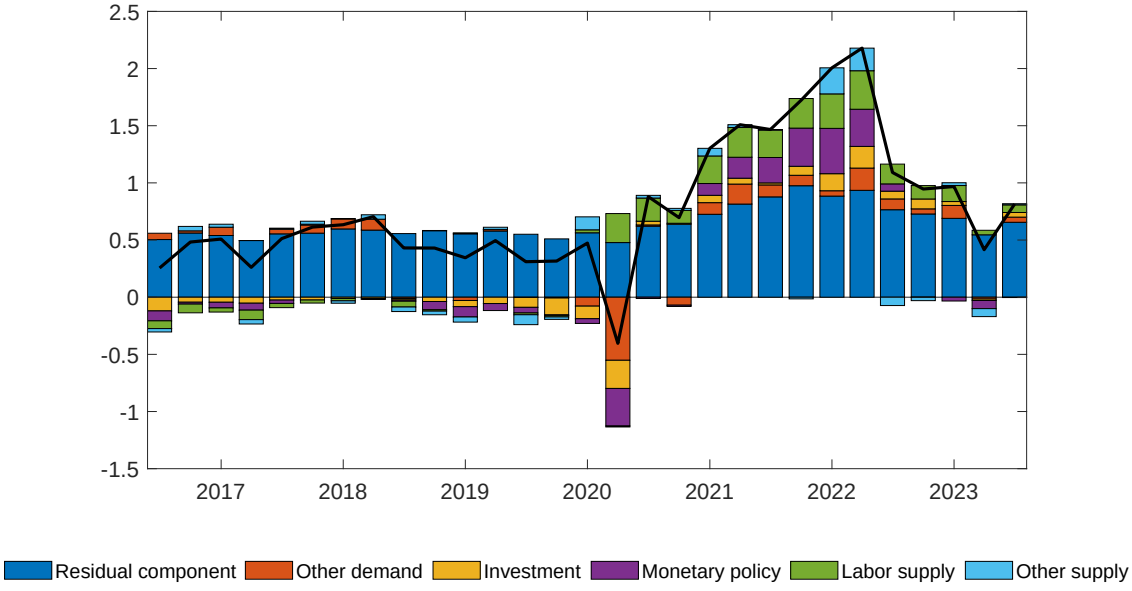
comovement between interest rates and inflation. The investment shock moves investment more than the other components of aggregate demand. The residual demand shock commingles government spending shocks, discount factor shocks and foreign shocks. On the supply side, a restriction on wages separates labor supply shocks from productivity shocks. This scheme is rather standard in the literature.

We present the historical decompositions based on the pointwise median and pointwise mean in Figure E-2 and E-3 respectively. The three demand shocks combined explain the bulk of the inflation surge in both cases. In agreement with Giannone and Primiceri (2024), Comin et al. (2023) and Gagliardone and Gertler (2023), monetary shocks play a large role: through the lenses of the VAR model, the response of policy to the inflation surge has been more muted than what it is implied by the historical rule.

Table E-1: Sign restrictions imposed on the larger SVAR

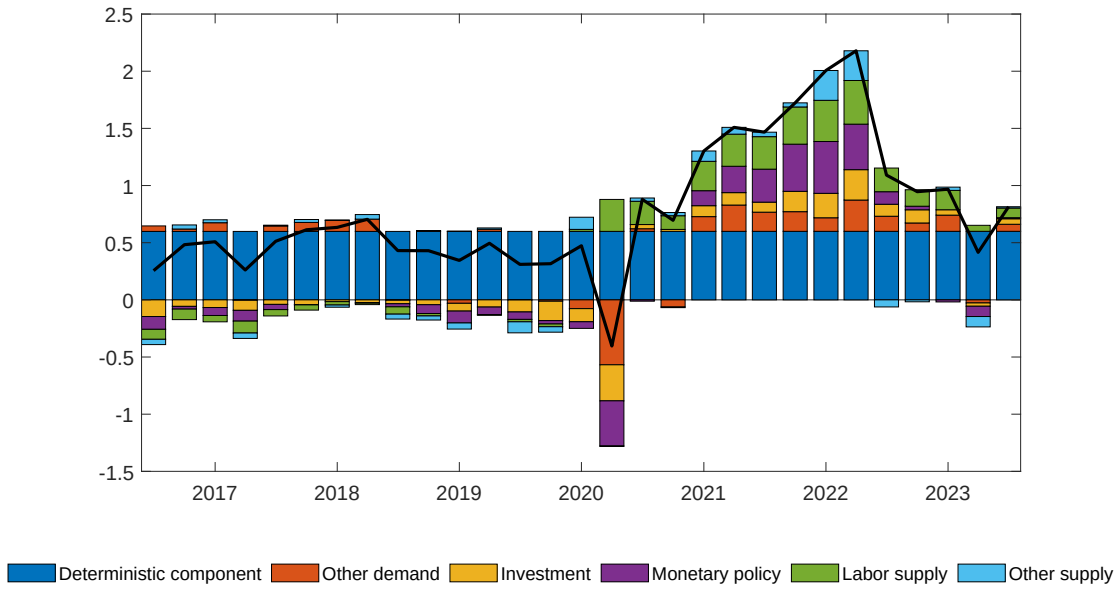
	Other demand	Investmant	Monetary policy	Labor supply	Other supply
GDP	+	+	+	+	+
Inflation	+	+	+	-	-
Investment	+	+	+		
Interest rate	+	+	-		
Real wage				-	+
Investment/GDP	-	+			

Figure E-2: Median Historical decomposition of inflation, single-unit-root prior, Larger SVAR



Note: The point-wise median contributions from the shocks and the deterministic component do not necessarily sum up to the level of inflation. The blue bars are the difference between inflation and the sum of the median contributions from the structural shocks.

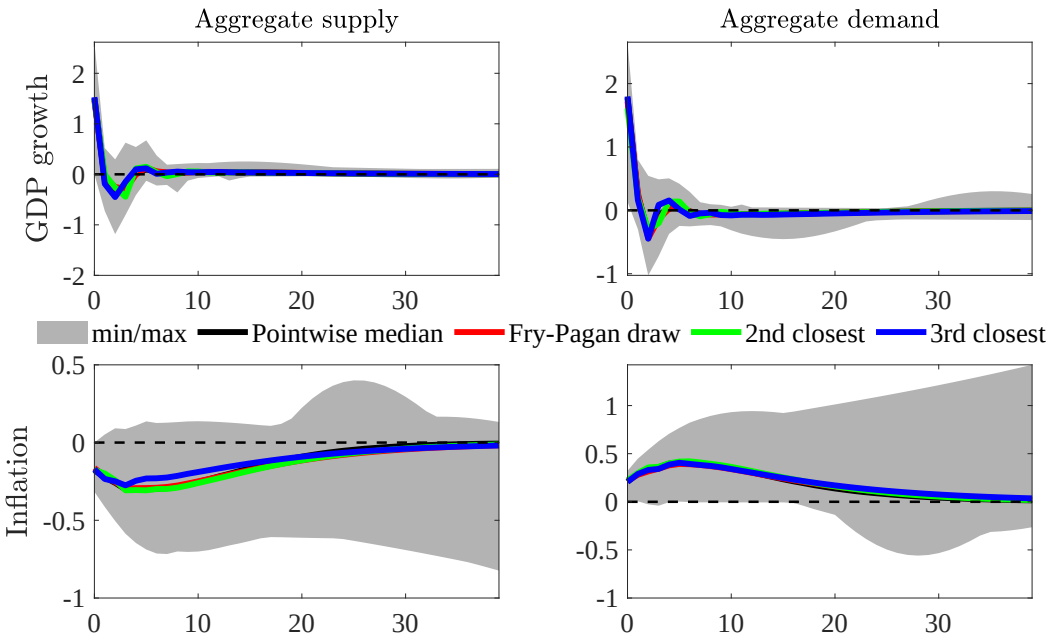
Figure E-3: Mean Historical decomposition of inflation, single-unit-root prior, Larger SVAR



F ADDITIONAL FIGURES FOR THE EURO AREA

Figure F-1: IRFs to identified shocks in the euro area

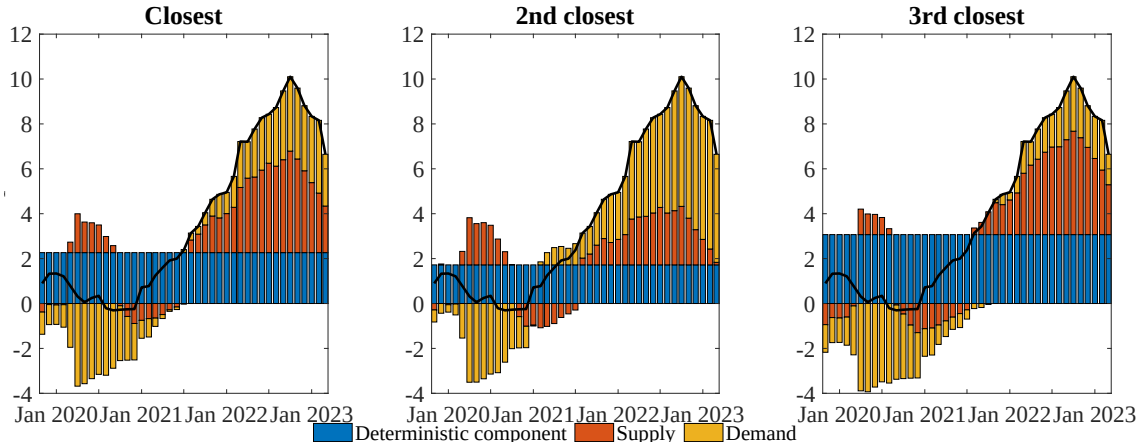
Pointwise median and three closest draws, diffuse prior



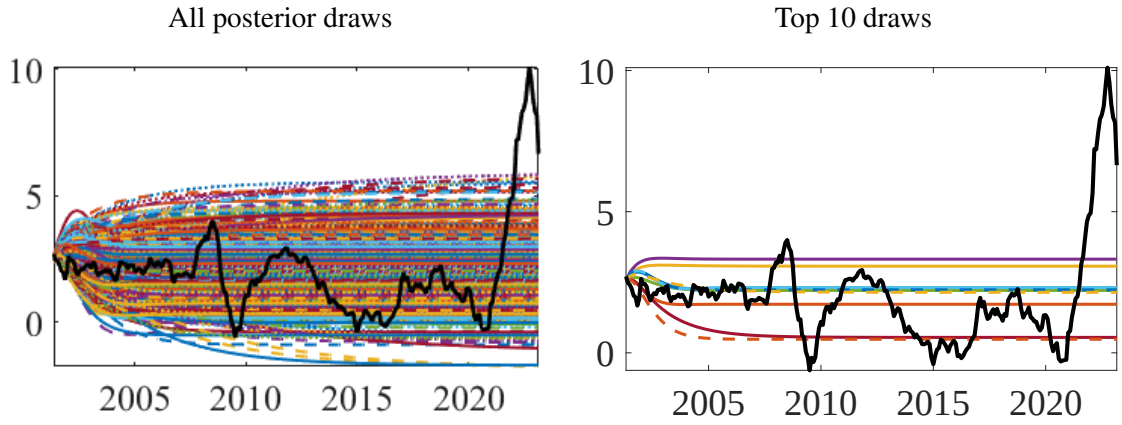
Note: The black line is the pointwise median and the shaded areas the min-max identified set. The red line is the IRF for the draw that is closest to the pointwise median; the green and blue lines are IRFS for 2nd and 3rd draws closest to the pointwise median, respectively.

Figure F-2: Historical decompositions and deterministic components of euro area inflation, diffuse prior

(a) Historical decomposition of inflation for the 3 draws closest to the pointwise median IRFs



(b) Deterministic components of inflation



Note: The black line in all the panels shows HICP inflation. Panel (a) displays the historical shock decompositions of inflation from the 3 draws closest to the pointwise median response for the IRFs. The colored lines in panel (b) display the deterministic components of inflation for draws from the posterior distribution.