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

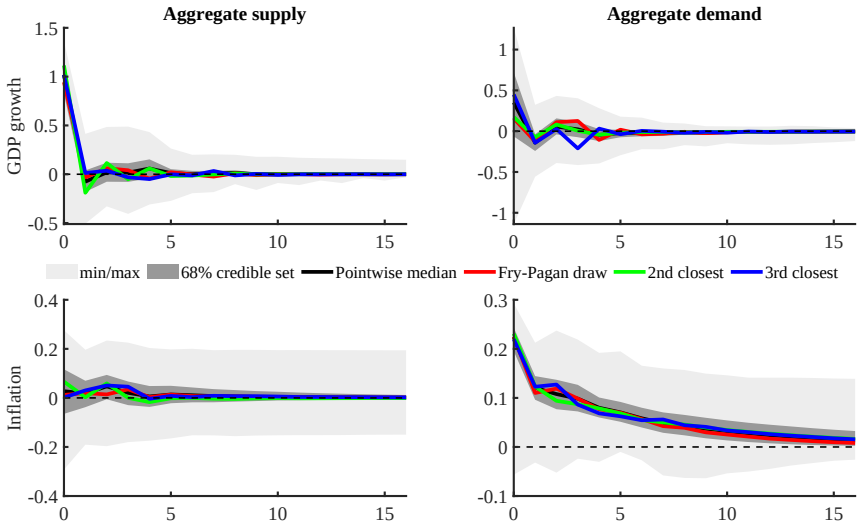
DRAGO BERGHOLT, FABIO CANOVA, FRANCESCO FURLANETTO, NICOLÒ MAFFEI-FACCIOLI AND PÅL ULVEDAL

SUPPLEMENTARY APPENDIX

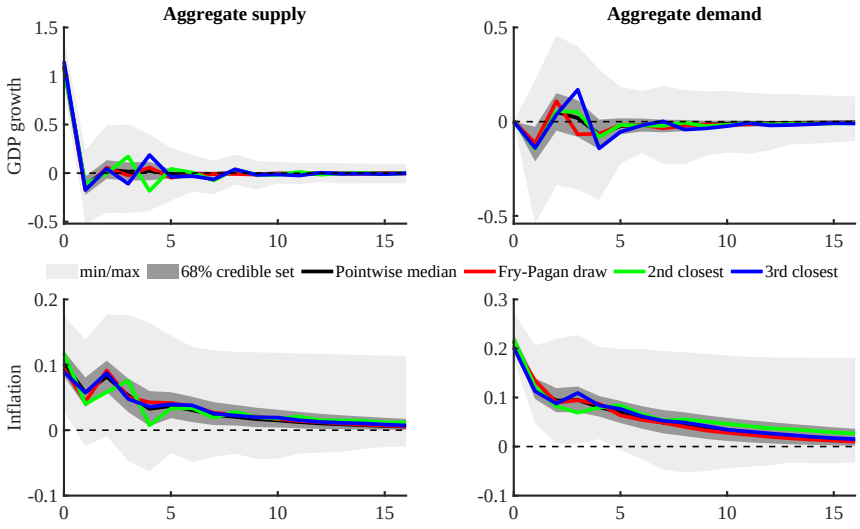
A ADDITIONAL RESULTS FOR THE US

Figure A-1: IRFs, different identification schemes

(a) Blanchard-Quah

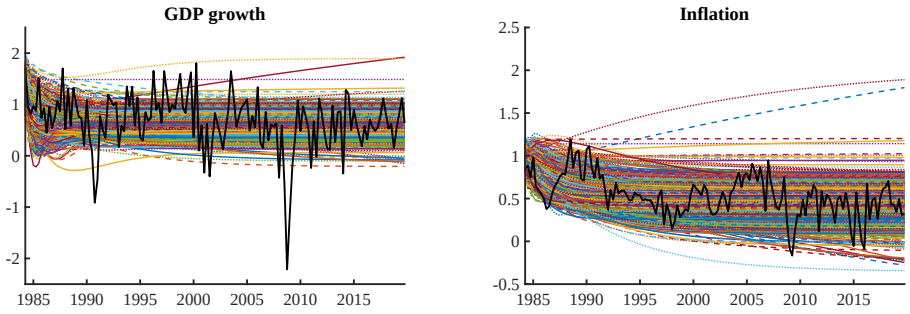


(b) Cholesky



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: 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:

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

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:

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

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

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:

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

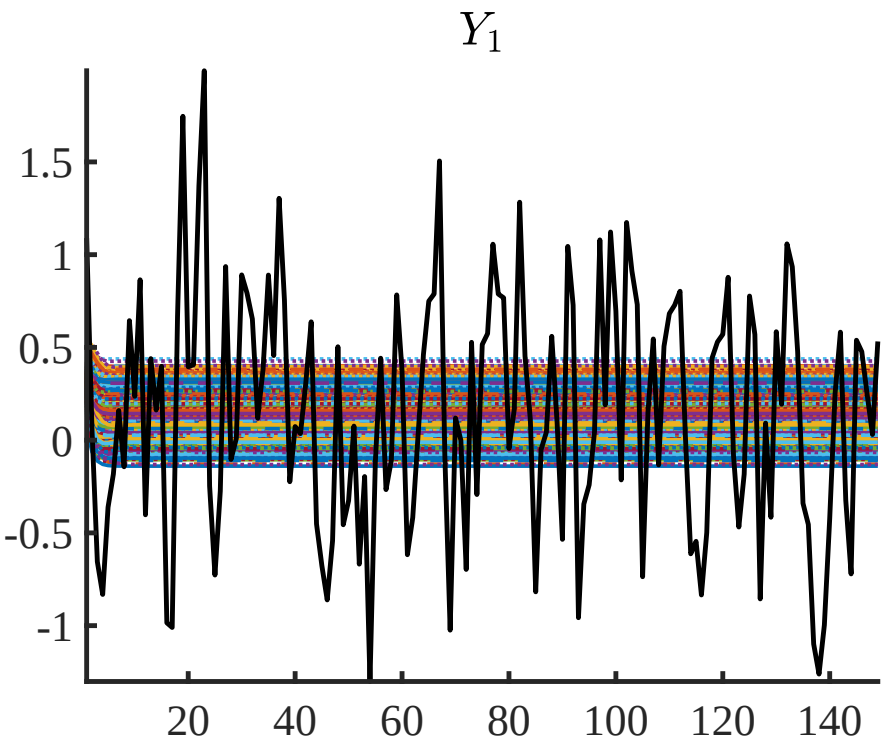
Notice that the terms $\Gamma(\mathbf{A})(\hat{\mathbf{C}}-\mathbf{C})\hat{\mathbf{C}}'(\Gamma(\hat{\mathbf{A}})-\Gamma(\mathbf{A}))'$ and $(\Gamma(\hat{\mathbf{A}})-\Gamma(\mathbf{A}))\hat{\mathbf{C}}(\hat{\mathbf{C}}-\mathbf{C})'\Gamma(\mathbf{A})'$ are negative semi-definite if the bias in $\hat{\mathbf{A}}$ 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.

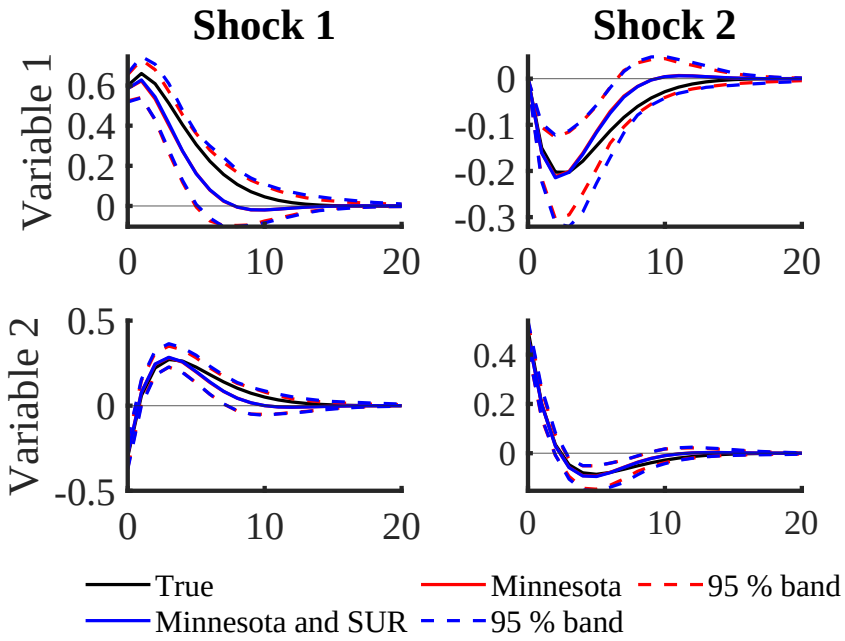
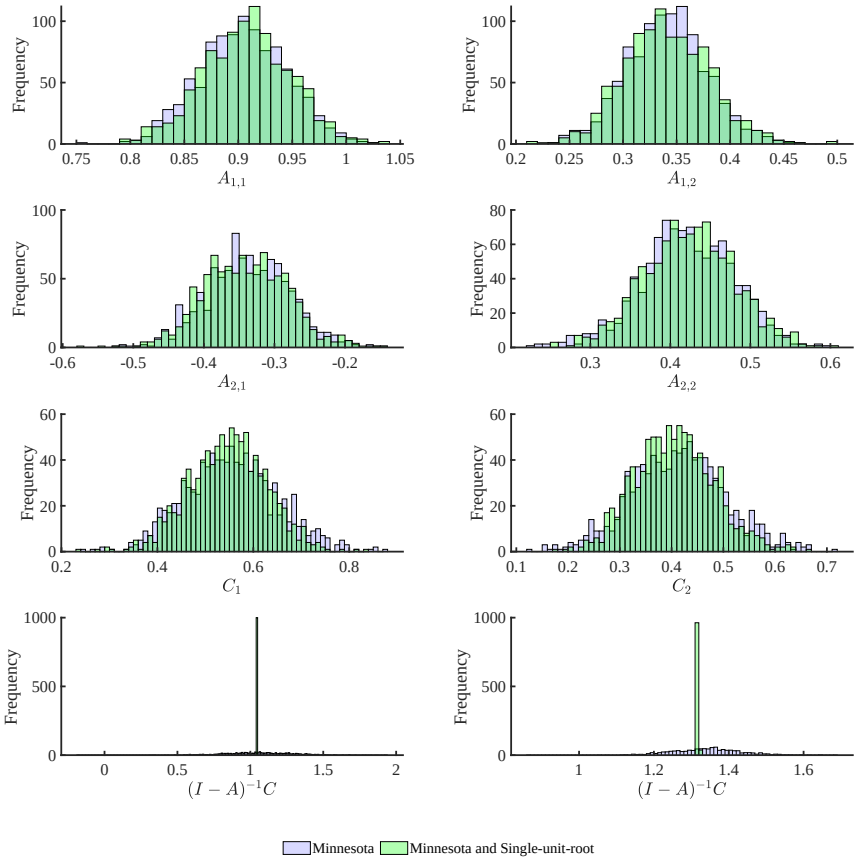


Table C-1: RMSE, Minnesota and Minnesota-SUR priors, simulated data.

h	Minnesota prior		Minnesota and SUR priors	
	Variable 1	Variable 2	Variable 1	Variable 2
1	0.62	0.59	0.62	0.58
4	1.20	0.67	1.21	0.68
8	1.47	0.78	1.47	0.78
20	1.55	0.80	1.55	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

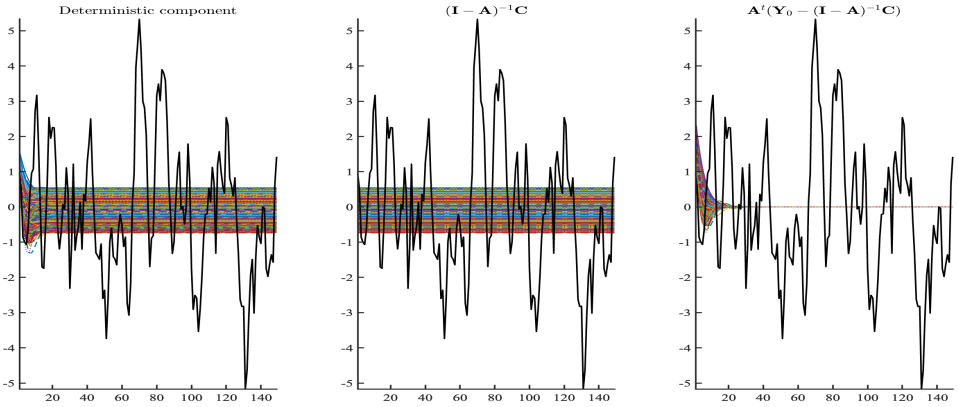
Figure D-1 provides evidence that shed further light on the relationship between dispersion in the deterministic component and overfitting. We consider three cases. In each case we decompose the deterministic component DC_t (left column) into the steady state term $(\mathbf{I}-\mathbf{A})^{-1}\mathbf{C}$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

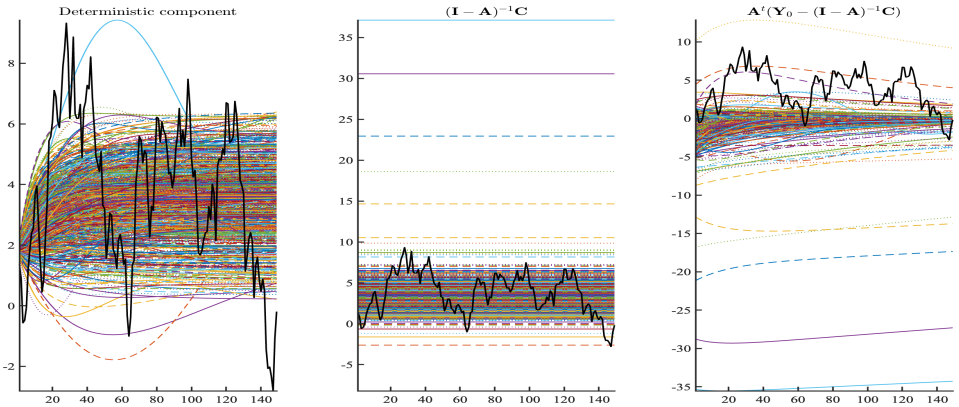
The final row in Figure D-1 still considers a random walk without drift but a sample of 10,000 observations is used to estimate the VAR. In this situation small sample issues are absent. Yet, even in this case we observe both excess dispersion, as well as long-lasting transitional dynamics when there 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.

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

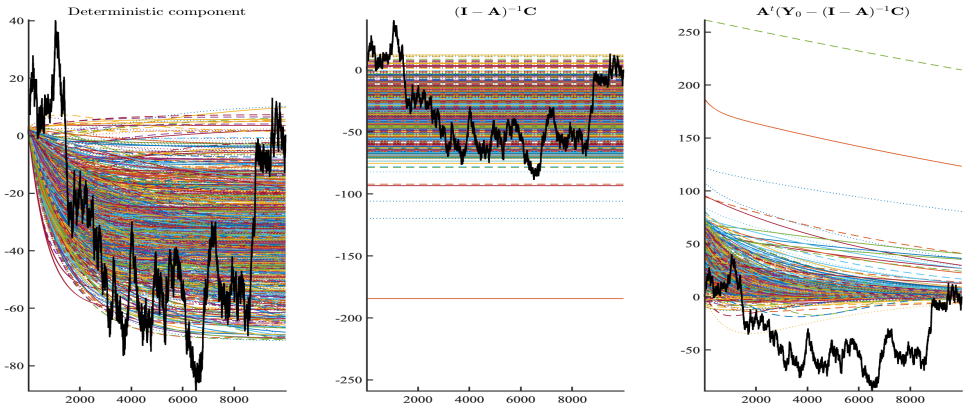
(a) $T = 150$ stationary



(b) $T = 150$ unit root



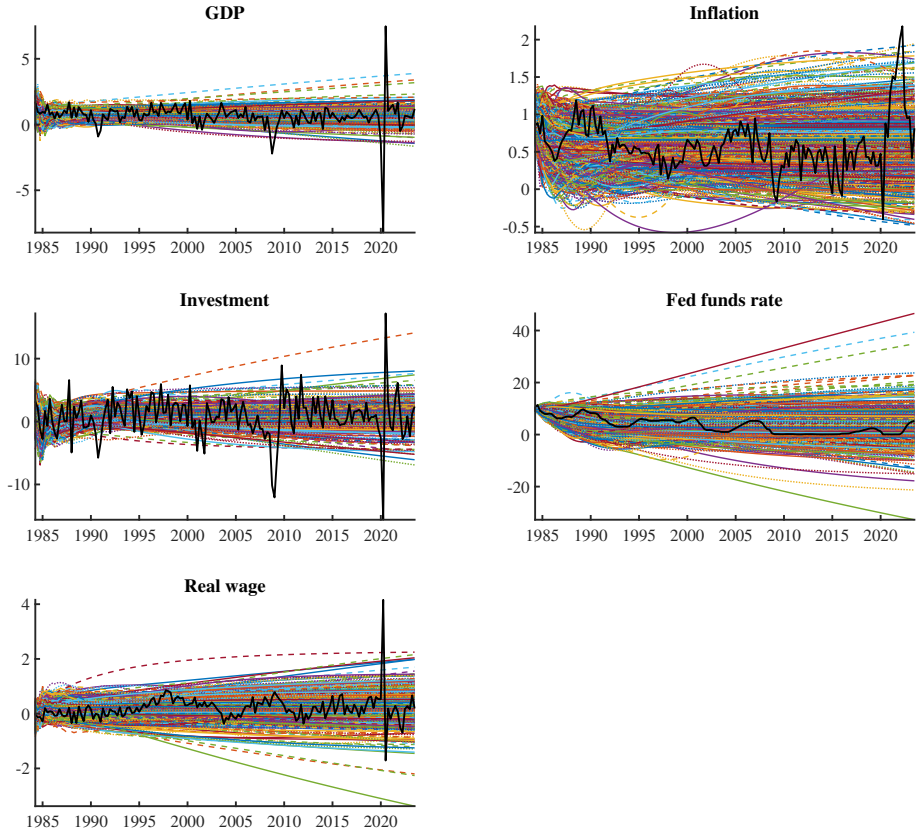
(c) $T = 10000$ unit root



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.

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.

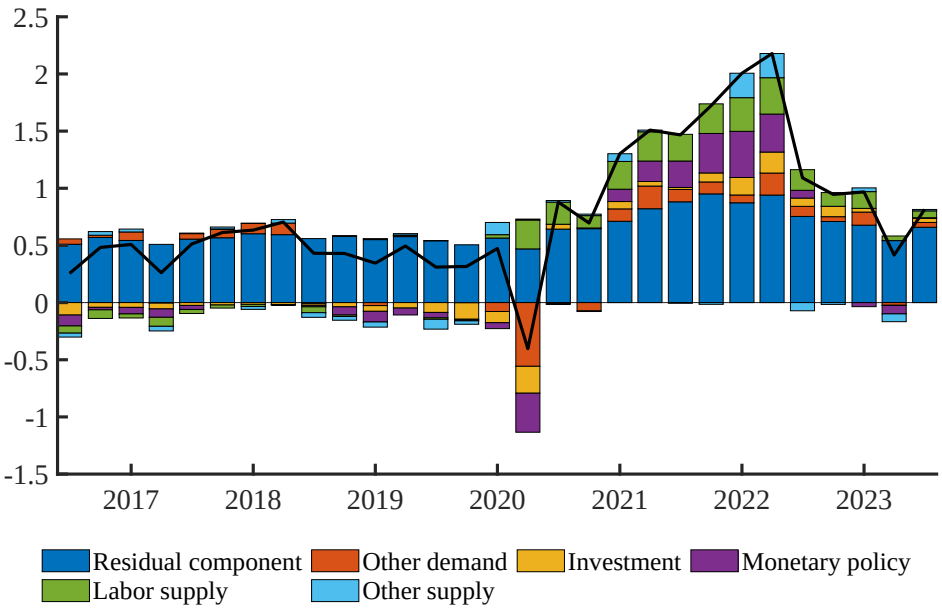
We present the historical decompositions based on the pointwise median and pointwise mean in Figure E-2 and E-3 respectively. The three demand shocks combined ex-

plain 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

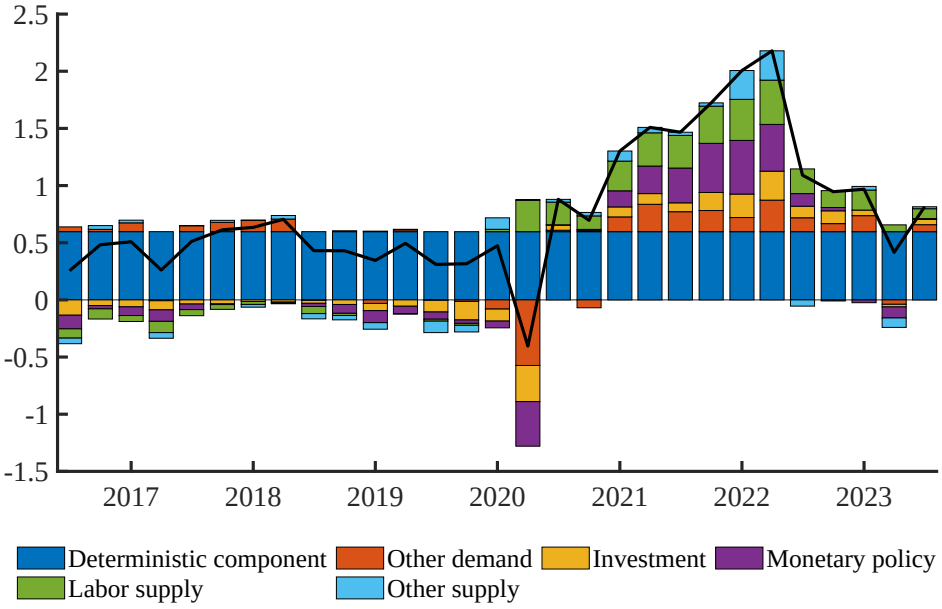
	Other demand	Investmant	Monetary policy	Labor supply	Other supply
GDP	+	+	+	+	+
Inflation	+	+	+	-	-
Investment	+	+	+		
Interest rate	+	+	-		
Real wage				-	+
Investment/GDP	-	+			

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.

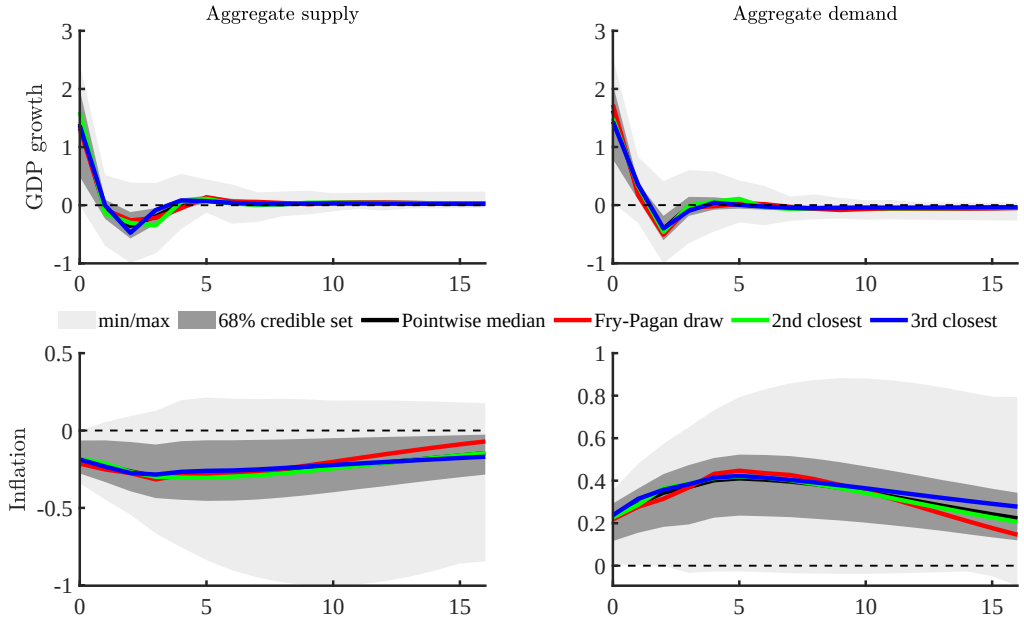
Figure E-3: Mean Historical decomposition of inflation, single-unit-root prior, Larger SVAR



F ADDITIONAL FIGURES FOR THE EURO AREA

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

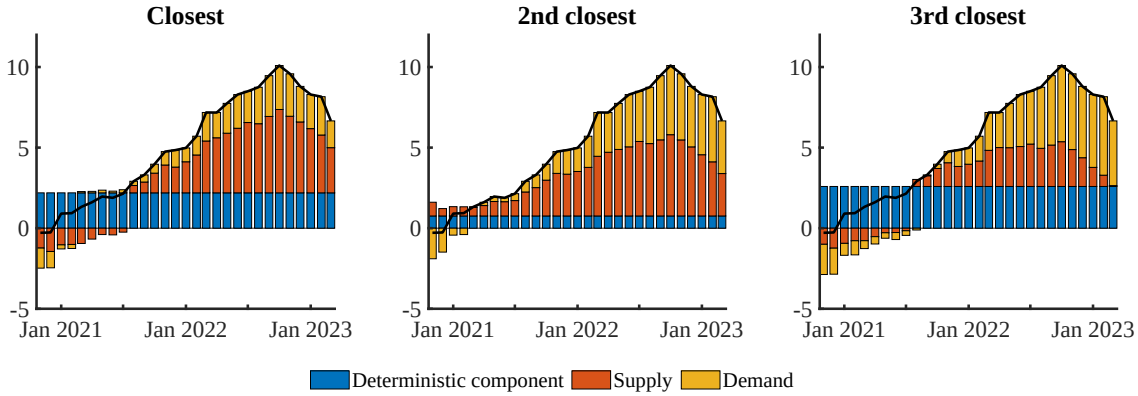
Pointwise median and three closest draws, diffuse prior



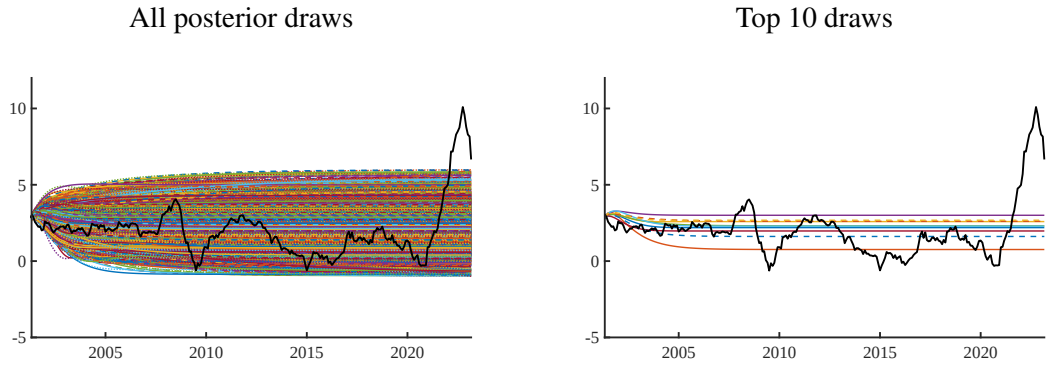
Note: The black line is the pointwise median and the shaded areas the min-max identified set. The red line is the IRF for the draw that is closest to the pointwise median; the green and blue lines are IRFS for 2nd and 3rd draws closest to the pointwise median, respectively.

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.