Evaluating Policy Counterfactuals: A VAR-Plus Approach[†]

Tomás E. Caravello — MIT
Alisdair McKay — FRB Minneapolis
Christian K. Wolf — MIT & NBER

September 15, 2024

Abstract: In a rich family of linearized structural macroeconomic models, the counterfactual evolution of the macro-economy under alternative policy rules is pinned down by just two objects: first, reduced-form projections with respect to a large information set; and second, the dynamic causal effects of policy shocks. In particular, no assumptions about the structural shocks affecting the economy are needed. We propose to recover these two sufficient statistics using a "VAR-Plus" approach, and apply it to evaluate several monetary policy counterfactuals.

Keywords: Lucas critique, policy counterfactuals, policy shocks, impulse response matching, invertibility. *JEL codes:* E32, E61.

[†]Email: `tomasec@mit.edu`, `alisdair.mckay@mpls.frb.org` and `ckwolf@mit.edu`. We received helpful comments from Marios Angeletos, Régis Barnichon, Marco Bassetto, Anmol Bhandari, Francesco Bianchi, Ricardo Caballero, Gabriel Chodorow-Reich, Martin Eichenbaum, Simon Gilchrist, Cosmin Ilut, Giuseppe Moscarini, Mikkel Plagborg-Møller, Giorgio Primiceri, Valerie Ramey, Matt Rognlie, Ben Schumann, Ludwig Straub, Iván Werning, and seminar participants at various venues. We also thank Seungki Hong, Klodiana Istrefi, and Diego Känzig for valuable discussions, and Valeria Morales Vasquez for superb research assistance. Wolf acknowledges that this material is based upon work supported by the NSF under Grant #2314736. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NSF, the Federal Reserve Bank of Minneapolis, or the Federal Reserve System.

1 Introduction

How would the economy have evolved if policy had been set differently? For example, how would a different monetary policy reaction function have shaped the average business cycle? And how would it have changed particular historical episodes?

We propose a new approach to this classic question of policy counterfactual evaluation—"VAR-Plus." The usual strategy is to build a dynamic stochastic general equilibrium (DSGE) model that can account for the entire history of macroeconomic fluctuations, change policy in that model, re-solve it, and then report counterfactual outcomes.[1] A commonly held concern with this approach is that the "shocks" that are added to the model to fit the data are more akin to statistical residuals than to true structural disturbances. As a result, some view these shocks as "dubiously structural" (e.g., see Chari et al., 2009), casting doubt on this method of analysis.[2] The principal appeal of our approach—which builds on and extends the recent "semi-structural" policy counterfactual literature (Barnichon and Mesters, 2023; McKay and Wolf, 2023)—is that it entirely sidesteps the need for the researcher to say anything about the underlying shocks driving the history of cyclical fluctuations.

Identification result. We are interested in the counterfactual evolution of the macroeconomy under alternative policy rules, both unconditional—i.e., how the "average" business cycle would unfold—and conditional on particular historical episodes. Extending prior results in McKay and Wolf (2023), we show that, across a large family of linearized macroeconomic models, these counterfactuals are pinned down by just two "sufficient statistics."

(i) *Reduced-form projections.* The first statistic is a set of reduced-form projections. For unconditional average business-cycle counterfactuals, those projections are impulse responses of macroeconomic aggregates to reduced-form ("Wold") innovations. For counterfactuals conditional on particular episodes, the projections are forecasts, from each date in the episode of interest. These projections need to be relative to an information set that spans the (unknown) shocks buffeting the macro-economy—i.e., we are maintaining the assumption of "invertibility", as typically done in the applied macroeconometrics literature (Sims, 1980; Fernández-Villaverde et al., 2007).

[1] Smets and Wouters (2007) is the exemplar of this approach. A notable recent example is Crump et al. (2023), who re-evaluate U.S. monetary policy during the post-covid inflationary episode.

[2] Commonly discussed examples of such "dubiously structural" shocks are price and wage mark-up shocks as well as innovations to household discount factors as consumer demand shocks.

(ii) *Policy causal effects.* The second statistic is the set of dynamic causal effects of changes in policy on current and future macroeconomic aggregates—i.e., the space of macroeconomic outcomes that is achievable through manipulation of the policy instrument(s), now and in the future. For example, for monetary policy, the researcher needs to know the causal effects of changes in the policy rate both today and at all future horizons.

The identification result reveals that any two structural models in the class we consider—no matter their detailed parametric structure, and in particular completely independently of the structural shocks they feature—that agree on these two sufficient statistics will also agree on the policy counterfactuals that they imply. The intuition has two parts. First, knowledge of policy causal effects ensures that we can correctly predict how any given reduced-form projection would be altered by a hypothetical change in policy. Second, given the assumption of invertibility, correctly predicting how reduced-form projections change is equivalent to correctly predicting the counterfactual propagation of the economy's true (though unknown) structural shocks, simply because those shocks are a 1-1 function of the Wold residuals.

We operationalize this identification result using our "VAR-Plus" method. The approach first constructs the required sufficient statistics, and with those evaluates the counterfactuals.

Getting the counterfactuals. In the first step of our methodology we leverage empirical time-series techniques—like Vector Autoregressions (VARs)—to recover the reduced-form projections (i) together with *some* policy causal effects (ii).[3]

(i) A reduced-form VAR is a convenient way of estimating the Wold representation of the data, and from here the required projections. This VAR should be specified with the invertibility requirement in mind; in practice, this requires including time series that are strong predictors of the variables whose counterfactual evolution is to be evaluated.

(ii) The standard semi-structural time series toolkit—like Structural VARs—can be used to estimate the causal effects of identified shocks to the policy instrument under consideration (e.g., see Ramey, 2016). In practice, since empirical evidence on policy shocks is limited, this will only *partially* pin down the required full space of policy causal effects. For example, for monetary policy, empirical analysis may deliver the causal effects of a transitory rate cut, but may be silent on the effects of persistent rate changes.

[3] We are using "VARs" as a shorthand for a wider menu of time series estimation techniques, including in particular also Local Projections (LPs). See, e.g., Plagborg-Møller and Wolf (2021) and Montiel Olea et al. (2024) for discussions of the trade-offs between the two estimation methods, which are not our focus here.

To complement the purely empirical evidence of the "VAR" step, we next rely on additional structural assumptions ("Plus") to complete part (ii). In this step we extrapolate from the empirical VAR evidence on *some* policy interventions to the full space of *all* possible interventions. Specifically, we consider one or more candidate models of policy transmission, and estimate them by requiring consistency with the policy shock evidence from the VAR step—i.e., model estimation via impulse-response matching (as in Christiano et al., 2005, 2010). This step yields a distribution over models of policy transmission and thus over the dynamic causal effects of *any possible* change in policy; importantly, that distribution is by design consistent with the available policy shock evidence, and then extrapolates beyond it—to predict the effects of other, unobserved policy changes—using model structure.

Leveraging our identification result, and with sufficient statistics (i) and (ii) in hand, we can finally evaluate the counterfactual of interest.

Why this approach? The principal appeal of our "VAR-Plus" approach is that it relies on weaker assumptions than the "quantitative DSGE" paradigm, and is thus less vulnerable to concerns of model mis-specification. To explain why, we review the pieces of information that our approach uses to answer a given policy counterfactual question. If the contemplated counterfactual involves a policy change that is spanned by the empirical VAR evidence on policy transmission—e.g., it only involves a transitory change in nominal rates—, then our method delivers an appealingly *semi-structural* counterfactual: all that is required are the theoretical identification result coupled with the VAR evidence. This will be the case, at least approximately, in two of our three applications. If instead the VAR evidence does not suffice, then the "Plus" step—i.e., the extrapolation of policy causal effects—starts to play a role. Importantly, however, this extrapolation is the *only* role for model structure; there is never any need to specify the shocks driving the cycle, thus sidestepping a key concern with the full-information, likelihood-based DSGE approach, as reviewed earlier.

Applications. We showcase our method with applications to monetary policy counterfactuals. We begin by constructing the required sufficient statistics.

(i) We first estimate a large-dimensional reduced-form VAR, with the specification closely following that of Angeletos et al. (2020). We collect the VAR-implied Wold innovation impulse responses as well as forecasts at each in-sample date.

(ii) We take the monetary shock series of Aruoba and Drechsel (2022), and use VAR methods to estimate its dynamic causal effects on macroeconomic aggregates. This evidence pins down the propagation of a *transitory* change in interest rates.

To extrapolate from transitory to more *persistent* rate changes, we consider a variety of quantitative business-cycle models, notably canonical RANK and HANK models (see Christiano et al., 2005; Kaplan et al., 2018), but also extended behavioral versions of these models with cognitive discounting (as in Gabaix, 2020). We find that all models can match the transitory monetary shock evidence quite closely; the RANK and HANK models furthermore largely agree on the extrapolation beyond the observed policy experiment to other horizons. On the other hand, models with cognitive discounting extrapolate quite differently; in particular, and as expected, they tend to imply much weaker effects of future policy on current outcomes.

With the sufficient statistics in hand, we study three monetary policy counterfactuals.

1. We ask whether U.S. monetary policy could have reduced the volatility of the aggregate output gap as well as inflation over a post-war sample period. Our analysis suggests that substantial volatility reductions would have been feasible, in particular for output.

2. We study how the Great Recession would have evolved in the absence of a binding lower bound on nominal rates. We find that a standard "dual mandate" central bank would have liked to reduce rates substantially into negative territory, suggesting that the implemented unconventional policy measures were insufficient.

These two counterfactuals are largely pinned down by the VAR step, with rather little role for the "Plus"-step extrapolation. The same is not true, however, for our third counterfactual.

3. We evaluate monetary policy options after the summer of 2021, when inflation had started to accelerate. The inflation spike is expected to be persistent, giving a larger role for far-ahead changes in policy. In the baseline HANK and RANK models, the policymaker can use forward guidance to steer inflation expectations, reducing current inflation at no cost to output in the short run. In our behavioral models, this strategy is much less effective. Given this disagreement across models and thus across the (relevant) policy causal effects, our method indicates large uncertainty on the counterfactual path of interest rates.

The large uncertainty in the third counterfactual reflects an important gap in our understanding of monetary policy transmission. The available empirical evidence only pins down the causal effects of transitory rate changes; some counterfactuals, however, depend crucially on the effects of persistent changes in policy. Standard HANK and RANK models extrapolate to such persistent monetary policy changes in quite similar ways, while less forward-looking behavioral models behave very differently. Discriminating between such models thus appears to be of chief importance—and in fact much more important than the incomplete-markets margin (i.e., RANK vs. HANK) that has received much attention recently.

Further literature. We contribute to a recent literature on policy shock impulse responses as "sufficient statistics" for policy counterfactuals (see McKay and Wolf, 2023; Barnichon and Mesters, 2023, 2024). Our analysis here differs in two key ways. First, to construct our counterfactuals of interest, we need to evaluate a counterfactual *system* for the propagation of a full set of reduced-form (Wold) innovations. We discuss when this reduced-form approach is valid, emphasize that it allows the researcher to remain silent on the primitive shocks driving the cycle, and connect our analysis with the critique of Chari et al. (2009). Second, we combine empirical VAR evidence with model-based policy causal effect extrapolation to evaluate our counterfactuals *exactly*, rather than just *approximately*. Our analysis reveals when the empirical evidence alone already suffices, when additional policy causal effect extrapolation is necessary, and how the dominant frameworks in the literature achieve this extrapolation. As such, our analysis also echoes the "sufficient statistics" results of the more recent trade and New Keynesian pricing literatures (e.g., as in Arkolakis et al., 2012; Auclert et al., 2022). Finally, our combination of direct empirical evidence and model-based policy causal effect extrapolation—plus our emphasis on invertibility and econometrician information sets—also distinguishes our analysis from Hebden and Winkler (2021), who rely exclusively on model-implied policy causal effects for policy evaluation.

We note that the estimand of our strategy is the effect of a systematic change in policy rule that is communicated to and understood by the private sector, as is typically assumed in the DSGE literature. This differs from the estimand of Sims and Zha (1995), who instead contemplate experiments in which the private sector is repeatedly surprised by policy shocks that implement the counterfactual, raising concerns related to the Lucas (1976) critique.

Outline. We begin in Section 2 with the identification result. We present our "VAR-Plus" methodology and discuss its theoretical properties in Section 3. Our applications to monetary policy counterfactuals follow in Sections 4 and 5. Section 6 concludes. Supplementary results follow in several appendices, and all replication codes are available online.[4]

[4]https://github.com/tcaravello/varplus

2 Identification result

This section presents our identification result. Sections 2.1 to 2.3 state and prove the main result, while Section 2.4 digs deeper into the key role played by the invertibility assumption. Our discussion here closely builds on but materially extends McKay and Wolf (2023).

2.1 General environment

Our identification result applies to a family of linearized infinite-horizon models with aggregate risk. The results are most easily stated using linearized perfect-foresight notation.[5]

Our description of the economic environment proceeds in two steps. First, we begin by introducing the structural vector moving-average (SVMA) representation of our economy. Second, we present the linearized perfect-foresight system whose transition paths equal the impulse responses collected in the SVMA coefficients.

Stochastic economy. We assume that our stochastic economy admits representation as a general SVMA(∞):

$$y_t = \sum_{\ell=0}^{\infty} \Theta_\ell \varepsilon_{t-\ell}. \tag{1}$$

y_t is a vector of macroeconomic aggregates, the shock vector ε_t is distributed as

$$\varepsilon_t \sim N(0, I),$$

and the $n_y \times n_\varepsilon$-dimensional matrices Θ_ℓ denote the impulse response of the vector of macroeconomic observables y_t at horizon ℓ to a date-t vector of shocks ε_t. We will throughout impose the high-level assumption that the matrices Θ_ℓ are absolutely summable across ℓ. Finally, in all of the following, the notation $\mathbb{E}_t[\bullet]$ will be reserved for expectations conditioning on the sequence of shocks $\{\varepsilon_{t-\ell}\}_{\ell=0}^{\infty}$ up to date t. Consistent with the classic Frisch (1933) impulse-propagation paradigm, the SVMA(∞) system (1) allows for an unrestricted dynamic linear transmission from shocks ε_t to outcomes y_t.

Impulse-response system. Leveraging the equivalence between linearized systems with aggregate risk and perfect-foresight transition paths, we obtain the impulse responses Θ_ℓ as

[5] By certainty equivalence, solutions to linearized perfect-foresight systems correspond to impulse response functions in linearized economies with aggregate risk (Fernández-Villaverde et al., 2016; Auclert et al., 2021).

solutions of a linear, perfect-foresight, infinite-horizon economy. Below boldface denotes time paths for $t = 0, 1, 2, \ldots$, and all variables are expressed in deviations from the deterministic steady state. The economy is summarized by the system

$$\mathcal{H}_w \boldsymbol{w} + \mathcal{H}_x \boldsymbol{x} + \mathcal{H}_z \boldsymbol{z} + \mathcal{H}_e e_0 = \mathbf{0}, \tag{2}$$

$$\mathcal{A}_x \boldsymbol{x} + \mathcal{A}_z \boldsymbol{z} + \mathcal{A}_v v_0 = \mathbf{0}. \tag{3}$$

Here x_t and w_t are n_x- and n_w-dimensional vectors of endogenous variables, respectively, z_t is an n_z-dimensional vector of policy instruments, e_t is an n_e-dimensional vector of exogenous structural shocks, v_t is an n_v-dimensional vector of policy shocks, and we write $y_t = (x_t', z_t')'$, $\varepsilon_t = (e_t', v_t')'$.[6] The distinction between w and x is that the variables in x are observable while those in w are not. Equation (2) summarizes the $n_x + n_w$-dimensional non-policy block of the model, with $\{\mathcal{H}_w, \mathcal{H}_x, \mathcal{H}_z, \mathcal{H}_e\}$ embedding private-sector relations. Equation (3) is the policy rule, with the instrument z set as a function of x and v.

Given the date-0 shocks $\{e_0, v_0\}$, an equilibrium is a set of bounded sequences $\{\boldsymbol{w}, \boldsymbol{x}, \boldsymbol{z}\}$ that solve (2) - (3). We will assume that the policy rule $\{\mathcal{A}_x, \mathcal{A}_z\}$ is such that an equilibrium exists and is unique. We write the implied mapping from shocks to outcomes as

$$y_\ell = \Theta_\ell \cdot \varepsilon_0.$$

Stacked together, those perfect-foresight mappings from date-0 shocks to date-ℓ outcomes deliver the SVMA(∞) representation (1).

The model (2) - (3) embeds an economically meaningful restriction—"instrument sufficiency", in the language of McKay and Wolf (2023). Policy is allowed to shape private-sector outcomes *only* through the current and expected future values of the policy instrument, i.e., via the path $\boldsymbol{z}$; whether or not that path is the result of the systematic component of policy (i.e., $\mathcal{A}_x$ and $\mathcal{A}_z$) or because of a policy shock (i.e., v) is entirely irrelevant.[7] As discussed in more detail by McKay and Wolf, many modern macro models satisfy this property, from representative-agent New Keynesian models (Christiano et al., 2005; Smets and Wouters, 2007), to heterogeneous-agent environments (Kaplan et al., 2018), and also including certain models with behavioral frictions (e.g., like Gabaix, 2020, see also Appendix A.1).

[6] The boldface vectors $\{\boldsymbol{w}, \boldsymbol{x}, \boldsymbol{z}\}$ stack time paths for all variables (e.g., $\boldsymbol{x} = (\boldsymbol{x}_1', \ldots, \boldsymbol{x}_{n_x}')'$). The maps $\{\mathcal{H}_w, \mathcal{H}_x, \mathcal{H}_z, \mathcal{H}_e\}$ and $\{\mathcal{A}_x, \mathcal{A}_z, \mathcal{A}_v\}$ are conformable and map bounded sequences into bounded sequences.

[7] Mathematically, this model property is reflected in the implicit assumption that the private-sector linear maps $\{\mathcal{H}_w, \mathcal{H}_x, \mathcal{H}_z, \mathcal{H}_e\}$ are invariant to the policy rule $\{\mathcal{A}_x, \mathcal{A}_z, \mathcal{A}_v\}$.

Some definitions. Under our assumptions on (1), the autocovariance function $\Gamma_y(\bullet)$ of macroeconomic observables y_t exists and by standard arguments is given as

$$\Gamma_y(\ell) = \sum_{m=0}^{\infty} \Theta_m \Theta'_{m+\ell}.$$

Next, the Wold representation of y_t is

$$y_t = \sum_{\ell=0}^{\infty} \Psi_\ell u_{t-\ell}, \tag{4}$$

where $u_t^\dagger \equiv y_t - \mathbb{E}\left(y_t \mid \{y_\tau\}_{-\infty<\tau\leq t-1}\right)$ denotes one-step-ahead forecast errors, $\text{Var}(u_t^\dagger) = \Sigma_u$, and $u_t \equiv \text{chol}(\Sigma_u)^{-1} u_t^\dagger$ are orthogonalized Wold innovations, with $\text{Var}(u_t) = I$ and $\text{chol}(\bullet)$ giving the lower-triangular Cholesky factor. Our assumptions on (1) ensure that this Wold representation exists, features no deterministic term, and that $\Psi(L)$ is square-summable.

2.2 Objects of interest

We wish to study the evolution of the economy if policy were set as

$$\tilde{\mathcal{A}}_x \boldsymbol{x} + \tilde{\mathcal{A}}_z \boldsymbol{z} = \boldsymbol{0} \tag{5}$$

rather than (3). The macroeconomic observables y_t under the counterfactual policy rule would then follow the counterfactual SVMA process

$$\tilde{y}_t = \sum_{\ell=0}^{\infty} \tilde{\Theta}_\ell \varepsilon_{t-\ell}, \tag{6}$$

with the convention that now $\varepsilon_t = e_t$, and where the shock impulse responses $\tilde{\Theta}_\ell$ are derived from the solution of the perfect-foresight system (2) together with (5).[8]

To define our counterfactuals of interest we need to tackle some subtleties on when precisely the counterfactual rule (5) is followed. In particular, the counterfactual SVMA (6) embeds the assumption that the counterfactual rule (5) is actually followed *forever*. In some of our "conditional" counterfactuals, however, we will assume that the policymaker instead

[8] In writing the counterfactual SVMA (6) we are implicitly assuming that the counterfactual policy rule (5) induces a unique equilibrium. Under multiplicity, our identification results will pin down counterfactual moments for one *possible* equilibrium, by the exact same arguments as in McKay and Wolf (2023).

unexpectedly changes to the alternative rule (5) at some date t^*, having followed the original rule (3) up to $t^* - 1$. In that case we will have

$$\tilde{y}_t = \underbrace{\sum_{\ell=0}^{t-t^*} \tilde{\Theta}_\ell \varepsilon_{t-\ell}}_{\text{new shocks after } t^*} + \underbrace{\tilde{y}_t^*}_{\text{initial conditions}} \tag{7}$$

The first term in (7) is straightforward: all newly arriving shocks ε_t propagate according to the new counterfactual impulse responses $\tilde{\Theta}_\ell$. The second term reflects initial conditions: at date t^*, the policymaker revises the planned policy path to ensure that current and expected future values of x and z are related according to (5). Letting $y_t^* = \mathbb{E}_{t^*-1}[y_t]$ denote date-t^*-1 expectations under the initially prevailing rule, the initial conditions term $\tilde{y}_t^*$ can thus be obtained by solving the system

$$\mathcal{H}_w(\tilde{\boldsymbol{w}}^* - \boldsymbol{w}^*) + \mathcal{H}_x(\tilde{\boldsymbol{x}}^* - \boldsymbol{x}^*) + \mathcal{H}_z(\tilde{\boldsymbol{z}}^* - \boldsymbol{z}^*) = \mathbf{0}, \tag{8}$$

$$\tilde{\mathcal{A}}_x \tilde{\boldsymbol{x}}^* + \mathcal{A}_z \tilde{\boldsymbol{z}}^* = \mathbf{0}, \tag{9}$$

i.e., a system written in terms of forecast revisions.[9] We can now state our counterfactuals of interest.

1. **Unconditional business cycles.** We seek the counterfactual second moments of y_t, given as

$$\tilde{\Gamma}_y(\ell) = \sum_{m=0}^{\infty} \tilde{\Theta}_m \tilde{\Theta}'_{m+\ell}.$$

   Unconditional "average" counterfactuals of this sort have attracted interest in prior work; examples include Rotemberg and Woodford (1997) or Del Negro and Schorfheide (2004).

2. **Conditional episodes.** We distinguish two kinds of conditional counterfactuals—forecasts, and full historical episodes.

   (i) *Conditional forecasts.* Consider some date t^*, and suppose the policymaker from t^* commits to the new rule (5). We may ask how, from that point onward, the economy

[9] Here, boldface denotes sequences from t^* onwards. (9) says the new, counterfactual policy rule holds. By (8), the revised forecasts remain consistent with all private-sector relationships. Finally, under our assumptions on equilibrium existence and uniqueness, it follows that (8) - (9) has a unique solution.

would be *predicted* to evolve; i.e., we would like to recover the expectation

$$\mathbb{E}_{t^*}\left[\tilde{y}_{t^*+h}\right] = \tilde{\Theta}_h \varepsilon_{t^*} + \tilde{y}^*_{t^*+h}.$$

Such conditional forecasts are key inputs for central banks (see Svensson, 1997) and have been studied widely in the academic literature (e.g., Antolin-Diaz et al., 2021).

(ii) *Historical evolution.* Consider a particular episode, $t \in [t_1, t_1+1, \ldots, t_2]$. We may ask how the economy would have evolved over that time window if the policymaker had followed the rule (5) from date t_1 onward; i.e., we seek to recover

$$\tilde{y}_t = \sum_{\ell=0}^{t-t_1} \tilde{\Theta}_\ell \varepsilon_{t-\ell} + \tilde{y}^1_t, \quad \forall t \in [t_1, t_1+1, \ldots, t_2]$$

where $\tilde{y}^1_t \equiv \mathbb{E}_{t_1-1}\left[\tilde{y}_t\right]$ reflects initial conditions as of date t_1. Policy counterfactuals for particular historical episodes have also been the subject of much prior work; e.g., see Eberly et al. (2020) for a recent example.

While McKay and Wolf (2023) mainly focus on policy counterfactuals for particular structural shocks (e.g., an oil shock), our counterfactuals here necessarily involve *all* shocks hitting the macro-economy. As a result, the assumption of invertibility will take center stage when we turn to identification in Section 2.3.

2.3 Identification result

This section states and proves the main identification result. We first introduce some additional notation on how policy can affect the macro-economy.

Policy causal effects. Recall that the policy rule (3) is subject to the n_v-dimensional vector of policy shocks v_t. To state our identification result, we will instead consider a full menu of policy shocks that perturb the policy rule *at each possible horizon*; that is, we have

$$\mathcal{A}_x \boldsymbol{x} + \mathcal{A}_z \boldsymbol{z} + \boldsymbol{\nu} = \boldsymbol{0} \tag{3'}$$

where the policy shock vector $\boldsymbol{\nu}$ is now unrestricted—i.e., we allow for arbitrarily flexible wedges in the rule at each date $t = 0, 1, 2, \ldots$. Analogously to the discussion in Section 2.1, the solution of the system (2) - (3') given an arbitrary policy shock vector $\boldsymbol{\nu}$ alone yields

$$\boldsymbol{y} \;=\; \Theta_\nu \;\cdot\; \boldsymbol{\nu}.$$

Θ_ν is the *space* of allocations implementable through policy shocks—i.e., the paths of macroeconomic aggregates corresponding to any possible *time path* of the policy instrument.

The identification result. We can now state our main identification result. We note that the first part of Proposition 1—the identification argument for unconditional business cycles—is already contained in McKay and Wolf (Appendix A.5, 2023), while the arguments for the conditional episodes are new to the present paper.

Proposition 1. *Suppose that the SVMA(∞) process (1) is invertible; i.e., that*

$$\varepsilon_t \in \text{span}(\{y_\tau\}_{-\infty<\tau\leq t}) \tag{10}$$

Then knowledge of: (i) the Wold representation y_t (i.e., the history of innovations $\{u_{t-\ell}\}_{\ell=0}^{\infty}$ together with $\Psi(L)$); and (ii) policy causal effects Θ_ν suffices to construct all policy counterfactuals of interest—$\tilde{\Gamma}_y(\ell)$, $\tilde{y}_t$, and $\mathbb{E}_t\left[\tilde{y}_{t+h}\right]$.

Before providing the formal proof, we find it useful to first discuss the high-level intuition underlying this identification result. For this we will proceed in two steps, first assuming that the researcher could actually observe all of the primitive structural shocks ε_t (rather than just y_t and the associated Wold representation). In that case, it is immediate that she could recover the impulse responses to those shocks under the baseline policy rule, $\Theta(L)$. She could then leverage the results of McKay and Wolf (2023): since she knows how any possible time path of the policy instrument z affects macroeconomic outcomes (i.e., Θ_ν), she can predict how those observed shocks ε_t would have counterfactually propagated under the alternative rule (5)—i.e., she has obtained $\tilde{\Theta}(L)$, thereby the counterfactual SVMA representation (6), and thus the desired policy counterfactuals.

The identification result then goes one step further and states that, under the additional assumption of invertibility (i.e., under condition (10)), directly observing the true shocks ε_t is actually not necessary—it suffices to just observe the reduced-form Wold innovations u_t. The reason for this is simply that our counterfactuals of interest are just forecasts—and by invertibility, forecasts with respect to the econometrician's information set equal forecasts with respect to the full history of structural shocks ε_t. Since (counterfactual) forecasts based on a one-to-one function of the true shocks equal forecasts based on the shocks themselves, the researcher is able to recover the correct counterfactuals.[10]

Proof. Consider using the policy transmission map Θ_ν to predict the counterfactual propagation of the Wold innovations u_t under the counterfactual policy rule (5), proceeding as in McKay and Wolf (2023, Proposition 1). Formally, for $j \in \{1, \dots, n_y\}$, let $\Psi_{\bullet,j}$ be the impulse response of y_t to the j-th Wold innovation $u_{j,t}$, and then construct the counterfactual impulse responses $\tilde{\Psi}_{\bullet,j}$ as

$$\tilde{\Psi}_{\bullet,j} = \Psi_{\bullet,j} + \Theta_\nu \tilde{\boldsymbol{\nu}}_j,$$

where the artificial policy shocks $\tilde{\boldsymbol{\nu}}_j$ solve the system of equations

$$\tilde{A}_x \left(\Psi_{\bullet,x,j} + \Theta_{x,\nu}\tilde{\boldsymbol{\nu}}_j\right) + \tilde{A}_z \left(\Psi_{\bullet,z,j} + \Theta_{z,\nu}\tilde{\boldsymbol{\nu}}_j\right) = \mathbf{0}. \tag{11}$$

Combining the $\tilde{\Psi}_{\bullet,j}$'s for all j, we get the counterfactual process

$$\tilde{y}_t = \sum_{\ell=0}^{\infty} \tilde{\Psi}_\ell u_{t-\ell}. \tag{12}$$

Under invertibility, the Wold innovations u_t and true structural shocks ε_t are related as

$$u_t = P\varepsilon_t,$$

where P is an orthogonal matrix. It then again follows from McKay and Wolf (2023) that the counterfactual Wold lag polynomial $\tilde{\Psi}(L)$ satisfies

$$\tilde{\Psi}(L) = \tilde{\Theta}(L)P'. \tag{13}$$

We now recover each of the desired counterfactuals.

1. Consider using the counterfactual process (12) to recover the desired counterfactual second-moment properties. Its implied autocovariance function is

$$\sum_{m=0}^{\infty} \tilde{\Psi}_m \tilde{\Psi}'_{m+\ell} = \sum_{m=0}^{\infty} \tilde{\Theta}_m P'P\tilde{\Theta}'_{m+\ell} = \sum_{m=0}^{\infty} \tilde{\Theta}_m \tilde{\Theta}'_{m+\ell} = \tilde{\Gamma}_y(\ell),$$

[10]Note that these are rational, full-information forecasts. As discussed in Appendix A.1, our identification results *are* consistent with various kinds of frictions in expectation formation; however, the forecasts that the econometrician leveraging Proposition 1 constructs still always need to be full-information forecasts.

where the first equality uses (13), and the second follows since P is an orthogonal matrix.

2. Applying Proposition 1 of McKay and Wolf (2023) to the system (8) - (9) that defines initial conditions $\tilde{\boldsymbol{y}}^*$, we see that we can recover initial conditions at t^* as

$$\tilde{\boldsymbol{y}}^* = \boldsymbol{y}^* + \Theta_\nu \tilde{\boldsymbol{\nu}}^*, \tag{14}$$

where the artificial policy shocks $\tilde{\boldsymbol{\nu}}^*$ now solve

$$\tilde{A}_x \left(\boldsymbol{x}^* + \Theta_{x,\nu}\tilde{\boldsymbol{\nu}}^*\right) + \tilde{A}_z \left(\boldsymbol{z}^* + \Theta_{z,\nu}\tilde{\boldsymbol{\nu}}^*\right) = \mathbf{0}. \tag{15}$$

Note that our informational requirements (i) - (ii) suffice to construct $\tilde{\boldsymbol{\nu}}^*$ and thus allow us to also evaluate the initial conditions term $\tilde{\boldsymbol{y}}^*$. In particular, invertibility here is crucial to ensure that $\boldsymbol{x}^*$ and $\boldsymbol{z}^*$ are equal to date-$t^* - 1$ forecasts based on the Wold representation (4), given as $y^*_{t^*+h} = \sum_{\ell=1}^{\infty} \Psi_{h+\ell} u_{t^*-\ell}$. We can now recover the two counterfactuals.

(i) Consider using (12) and (14) to recover the conditional forecast $\mathbb{E}_t \left[\tilde{y}_{t+h}\right]$. We have

$$\tilde{\Psi}_h u_{t^*} + \tilde{y}^*_{t^*+h} = \underbrace{\tilde{\Psi}_h P}_{=\tilde{\Theta}_h} \varepsilon_{t^*} + \tilde{y}^*_{t^*+h} = \mathbb{E}_{t^*}\left[\tilde{y}_{t^*+h}\right].$$

(ii) Consider using (12) and (14) to recover the historical counterfactual $\tilde{y}_t$. We have

$$\sum_{\ell=0}^{t-t_1} \tilde{\Psi}_\ell u_{t-\ell} + \tilde{y}^1_t = \sum_{\ell=0}^{t-t_1} \underbrace{\tilde{\Psi}_\ell P}_{=\tilde{\Theta}_\ell} \varepsilon_{t-\ell} + \tilde{y}^1_t = \tilde{y}_t.$$

□

2.4 More on the invertibility assumption

Our discussion following Proposition 1 explained the role of invertibility in terms of forecasting. In this section we will use model-based simulations to show that, even when invertibility fails, we can still obtain accurate counterfactuals as long as we have forecasts that are reasonably accurate vis-à-vis the full-information benchmark.

Experiment. We consider a structural model—the medium-scale DSGE model of Smets and Wouters (2007)—as an artificial laboratory. In this environment we seek to recover the counterfactual second moments of output, inflation, and interest rates under an alternative monetary policy rule that puts a larger weight on output stabilization. To do so, we leverage Proposition 1, using the true matrix of policy shock causal effects Θ_ν, but then relying on information sets $\{y_{t-\ell}\}_{\ell=0}^{\infty}$ that are (potentially) insufficient to deliver invertibility.[11] When invertibility fails, then this procedure will not recover the true counterfactual; our question is just how inaccurate those predictions will end up being. For our explorations we will consider four information sets: interest rates, output, and inflation alone ("baseline"); the baseline plus hours worked; the baseline plus investment and consumption; and finally the baseline plus hours worked, wages, investment as well as consumption. Among those four information sets, only the fourth one satisfies invertibility.

The remainder of this section presents the main results of this exercise. Further implementation details are relegated to Appendix A.3.

Results. Our key finding—revealed in Figure 1—is that even small information sets can deliver predicted policy counterfactuals that are almost indistinguishable from the true ones. The top panel summarizes second moments via spectral densities over business-cycle frequencies, with the pink lines corresponding to the prevailing monetary policy rule, while the other lines indicate the true (solid) and predicted (dashed) counterfactual spectral densities. As expected, the predicted counterfactual based on the full information set is equal to the truth. More importantly, for *all* information sets, the predicted counterfactuals are close to each other, and so to the truth. In other words, though invertibility fails, implementing the steps of Proposition 1 yields predictions close to the true counterfactual.

The explanation for this finding lies in our discussion of the role of invertibility above—its key purpose being to deliver full-information forecasts. The bottom panel of Figure 1 makes this point by showing residual forecast uncertainty for interest rates, output, and inflation at different horizons (x-axis), and for our different information sets (different lines); for that panel, we have normalized the residual forecast uncertainty under the full information set to 1 at each horizon h. As $h \to \infty$, the residual forecast variances for all information sets of course limit to the same number—the unconditional variance. For intermediate h, forecasting uncertainty is instead strictly larger for smaller information sets. The differences, however, are moderate, with forecast variances that are only at most around 10 per cent larger

[11] To be precise, given Θ_ν and the Wold representation of the observables y_t, we *mechanically* follow the same steps outlined in the constructive proof of Proposition 1. We then report the second moments implied by the constructed counterfactual Wold representation $\tilde{\Psi}(L)$.

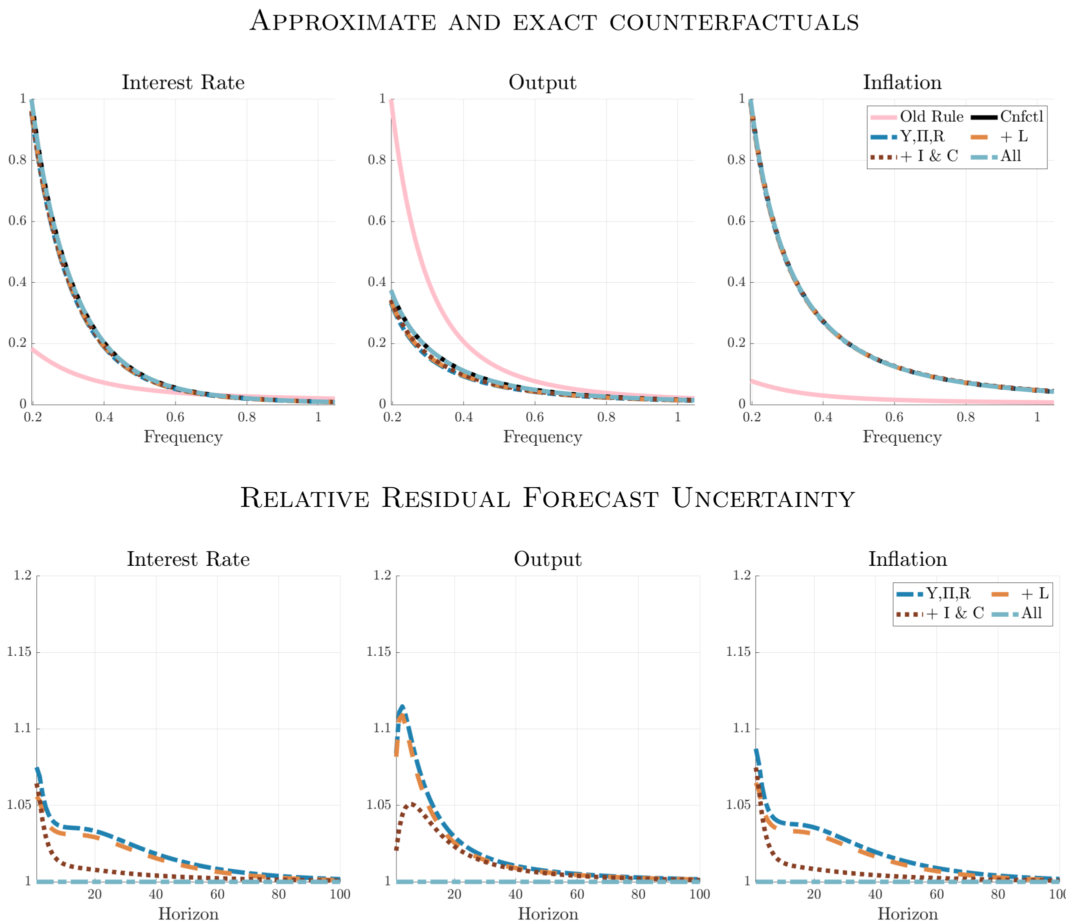

Figure 1: Top panel: business-cycle spectral densities for interest rates, output, and inflation under the old rule (solid pink) and under the counterfactual rule, true (solid dark blue) and predicted using Proposition 1 for different information sets (solid-dashed blue, dashed orange, dotted red, solid-dashed cyan). We normalize the peak spectral density to 1. Bottom panel: residual forecast variances for the same variables and for the same information sets, as a function of the forecast horizon (x-axis), and relative to the forecast variance for the full information set.

than with the full information set. Even the small information sets thus deliver accurate forecasts, and thus accurate counterfactuals.[12] In other words, the key requirement is to have forecasts of output, inflation, and nominal interest rates that approach the accuracy of the full-information benchmark—and for that, the past history of those three series evidently suffices, at least in the model of Smets and Wouters.

[12] This discussion suggests that, analogous to the SVAR literature, our results are continuous in the *degree* of invertibility (e.g., see Sims, 2012). Appendix A.2 confirms this, with the relevant measure now the R^2 in a population regression of full-information forecasts on the information set $\{y_{t-\ell}\}_{\ell=0}^{\infty}$. We also note that invertibility is testable when at least *some* shocks are observed; see Plagborg-Møller and Wolf (2022).

3 Policy counterfactuals via VAR-Plus

By Section 2, constructing policy counterfactuals requires two inputs: (i) reduced-form projections with respect to a large information set, and (ii) policy shock dynamic causal effects. We now introduce our VAR-Plus approach for constructing those inputs. We present the method in Sections 3.1 to 3.3, and then discuss its properties in Section 3.4.

3.1 Reduced-form projections

As the first step in our methodology, the researcher selects a set of macroeconomic observables y_t and then estimates their Wold representation (4)—i.e., she recovers the orthogonalized Wold innovations u_t and the lag polynomial $\Psi(L)$. In principle there are many ways of doing so. One particularly simple and convenient alternative is to estimate a VAR(p) in y_t:

$$y_t = \sum_{\ell=1}^{p} A_\ell y_{t-\ell} + u_t^\dagger \tag{16}$$

The orthogonalized Wold innovations u_t are then equal to $\text{chol}(\Sigma_u)^{-1}u_t^\dagger$ (where $\Sigma_u = \text{Var}(u_t^\dagger)$), and the Wold lag polynomial $\Psi(L)$ is given as $(I - A(L))^{-1}\,\text{chol}(\Sigma_u)$. This is the first half of the VAR step of our proposed methodology.

In practice, the vector of observables y_t should be chosen to be large enough—and contain enough forward-looking variables—so that the invertibility assumption is plausible. More practically, by the discussion in Section 2.4, this means that, if the researcher is interested in counterfactuals for some set of macroeconomic outcomes $y_t^\dagger$, then the vector of observables y_t should include $y_t^\dagger$ together with other series that are useful for forecasting $y_t^\dagger$.

3.2 Policy causal effects

We next require the policy causal effects Θ_ν. For this we will proceed in two sub-steps. First, we use semi-structural time-series methods to get empirical evidence at least on *parts* of Θ_ν, completing the VAR step of our methodology. Second, we use additional structure—in the form of one or multiple models of policy transmission—to first match and then extrapolate beyond that evidence, giving the rest of Θ_ν, i.e., the Plus part of our approach.

EMPIRICAL EVIDENCE ON Θ_ν. In the first sub-step, the researcher uses the standard time-series toolkit—typically in the form of a Structural Vector Autoregression (SVAR) or Local

Projection (LP)—to estimate the causal effects of a list of n_ν distinct policy shocks. For example, for monetary policy applications, she may estimate the causal effects of short-lived and more persistent innovations to the federal funds rate, following identification strategies as in Romer and Romer (2004) or Gertler and Karadi (2015). We then stack those estimated impulse responses of n_m targeted outcome variables over H impulse response horizons to the n_ν identified shocks in the $n_\nu \times n_m \times H$ vector $\hat{\theta}_\nu$.

Under standard asymptotic sampling theory, the asymptotic distribution of the policy shock causal effect vector $\hat{\theta}_\nu$ satisfies (e.g., see Christiano et al., 2010)

$$\hat{\theta}_\nu \overset{a}{\sim} N(\theta_\nu, V_{\theta_\nu}).$$

Our methodology requires the researcher to have an at least approximately consistent estimator of the asymptotic covariance matrix V_{θ_ν}. We discuss standard options and our preferred approach for doing so in Appendix B.2.

Taken together, these reduced-form projection and policy shock estimation exercises conclude the VAR step of the proposed methodology. As we will discuss in Section 3.4 and then showcase through our applications in Section 5, this step often already suffices to largely pin down the policy counterfactual of interest.

Impulse response extrapolation. The Plus step begins with the researcher writing down a list $\mathcal{M}$ of structural models of policy transmission, denoted by $\mathcal{M}_j$ for $j = 1, 2, \ldots, M$. In the notation of Section 2.1, a "model" is a tuple $\{\mathcal{H}_w, \mathcal{H}_x, \mathcal{H}_z\}$—i.e., a set of private-sector relations, but *without* structural shocks to those relations $\{\mathcal{H}_e e_0\}$, nor a policy rule $\{\mathcal{A}_x, \mathcal{A}_z\}$. Each model then has a parameter vector ψ_j mapping into $\{\mathcal{H}_w, \mathcal{H}_x, \mathcal{H}_z\}$, a prior distribution $p(\psi_j \mid \mathcal{M}_j)$ for the model parameters, and a prior probability $p(\mathcal{M}_j)$. We write $\theta_\nu(\psi_j, \mathcal{M}_j)$ as the model-implied analogue of the empirically observed policy shock causal effect vector; briefly, this object is defined as the impulse responses to a change in policy that comes as close as possible to the empirical targets. A discussion of how to construct this object for any given structural model is provided in Appendix B.3.[13]

Each model $\mathcal{M}_j$ among the list of contemplated models is estimated through standard impulse-response matching techniques (Rotemberg and Woodford, 1997; Christiano et al., 2005, 2010). Cast as a standard limited-information Bayesian estimation strategy, we can

[13] In the standard impulse response-matching literature, the researcher writes down a policy rule, and then restricts attention to contemporaneous shocks to that rule. We instead find the best fit to the empirical targets within the overall *space* implementable by policy, allowing us to not need to commit to any rule.

define an approximate likelihood of the "data," $\hat{\theta}_\nu$, as a function of ψ_j given $\mathcal{M}_j$:

$$p(\hat{\theta}_\nu \mid \psi_j, \mathcal{M}_j) \propto \exp\left[-0.5\left(\hat{\theta}_\nu - \theta_\nu(\psi_j, \mathcal{M}_j)\right)' V_{\theta_\nu}^{-1}\left(\hat{\theta}_\nu - \theta_\nu(\psi_j, \mathcal{M}_j)\right)\right]. \tag{17}$$

Combining the prior together with the likelihood (17), we obtain the posterior for ψ_j conditional on model $\mathcal{M}_j$ and given the policy shock causal effect data $\hat{\theta}_\nu$:

$$p(\psi_j \mid \hat{\theta}_\nu, \mathcal{M}_j) = \frac{p(\hat{\theta}_\nu \mid \psi_j, \mathcal{M}_j)p(\psi_j \mid \mathcal{M}_j)}{p(\hat{\theta}_\nu \mid \mathcal{M}_j)},$$

and where

$$p(\hat{\theta}_\nu \mid \mathcal{M}_j) = \int p(\hat{\theta}_\nu \mid \psi_j, \mathcal{M}_j)p(\psi_j \mid \mathcal{M}_j)d\psi_j$$

is the marginal density of $\hat{\theta}_\nu$ given model $\mathcal{M}_j$. Computation of these objects is standard, relying on the usual random walk Metropolis-Hastings algorithm both to draw from the posterior distribution and to compute the marginal likelihood. See Appendix B.3.

The final step is to recover posterior model probabilities—i.e., the posterior distribution across the model space $\mathcal{M}$. We have

$$p(\mathcal{M}_j \mid \hat{\theta}_\nu) = \frac{p(\hat{\theta}_\nu \mid \mathcal{M}_j)p(\mathcal{M}_j)}{\sum_{i=1}^{M} p(\hat{\theta}_\nu \mid \mathcal{M}_i)p(\mathcal{M}_i)}. \tag{18}$$

The researcher has thus arrived at a posterior distribution over models and parameter vectors, $p(\psi_j, \mathcal{M}_j \mid \hat{\theta}_\nu)$. Each parameterized model implies a policy transmission map

$$\Theta_\nu \quad = \quad \Theta_\nu(\psi_j, \mathcal{M}_j).$$

We have thus arrived at a posterior distribution over the causal effects of policy on macroeconomic aggregates Θ_ν, $p(\Theta_\nu)$, concluding the Plus step. By construction, this distribution is consistent with the empirical evidence from the VAR step, and then extrapolates beyond it according to the structure embedded in the contemplated structural models.[14]

3.3 Constructing policy counterfactuals

It now remains to put together the estimated inputs—i.e., reduced-form Wold innovations and projection coefficients $\{u_t, \Psi(L)\}$ as well as policy shock causal effects Θ_ν—to construct

[14] Strictly speaking, these statements presuppose that at least one of the contemplated models can match the targeted empirical evidence $\hat{\theta}_\nu$ well. In our empirical applications we will make sure that this is the case.

the various policy counterfactuals of interest. The formulas for mapping $\{u_t, \Psi(L)\}$ together with Θ_ν into our three desired counterfactuals are provided in the proof of the identification result in Proposition 1; see Section 2.3.

An important practical implementation challenge is how to take into account estimation uncertainty for the inputs $\{u_t, \Psi(L), \Theta_\nu\}$. For this we propose to proceed as follows. First, for the reduced-form inputs $\{u_t, \Psi(L)\}$, we simply look at point estimates. This is in keeping with standard practice in the policy counterfactual literature, which tends to take as given point estimates for the baseline second moments and forecasts (e.g., Rotemberg and Woodford, 1997; Eberly et al., 2020). Second, given those point estimates, we construct the policy counterfactuals by drawing Θ_ν from the posterior distribution estimated in our second step. Given the point estimates of $\{u_t, \Psi(L)\}$, the posterior distribution over Θ_ν thus maps into a posterior distribution over the counterfactuals $\{\tilde{\Gamma}(\ell), \mathbb{E}_t\left[\tilde{y}_{t+h}\right], \tilde{y}_t\}$.

3.4 Why this approach?

Our approach contributes to the recent literature on policy evaluation with less reliance on explicit model structure (Barnichon and Mesters, 2023; McKay and Wolf, 2023). As such, its chief appeal relative to the still dominant "quantitative DSGE" approach (e.g., as in Smets and Wouters, 2007; Justiniano et al., 2010) lies in its greater robustness to plausible model mis-specification. These robustness gains are best understood in two steps.

1. *The VAR step may well suffice.* To evaluate any given policy counterfactual question, the VAR step—which delivers Wold impulse responses as well as the causal effects $\hat{\theta}_\nu$ of *some* policy shocks—may already suffice. Intuitively, a given counterfactual of interest may only involve a transitory (say) change in policy, and such transitory changes may well already be spanned by the empirical evidence.[15] If that is so, then the output of our approach is actually robust to *arbitrary* forms of model mis-specification within the general linearized environment (2) - (3) that underlies our theoretical identification result. Our applications in Section 5 will reveal that this case is actually quite plausible in practice.

2. *Assumptions on primitive underlying shocks are never needed.* If additional model structure is needed—i.e., the "Plus" step—then it is exclusively for the purpose of extrapolating policy causal effects; in particular, the researcher need never say anything about the

[15]More formally, using just the empirically estimated policy shock causal effects $\hat{\theta}_\nu$ (rather than the full matrix Θ_ν), it may already be possible to enforce the counterfactual policy rule of interest to a high degree of accuracy, i.e., *nearly* solve equation (11) or (15).

stochastic shocks driving the macro-economy. This means that our approach sidesteps one of the literature's central concerns with the standard DSGE paradigm: likely mis-specification of the model's shock processes. In general, in large-scale quantitative models, it is not clear where exactly the driving shocks should enter, how many there should be, and what stochastic processes they should follow. This potential mis-specification is problematic for at least two reasons: first, mis-specified shocks are not plausibly structural, and so they may well invalidate policy counterfactual analysis, e.g., as argued forcefully by Chari et al. (2009); and second, mis-specification in the exogenous driving forces of the model will also necessarily threaten inference on the endogenous propagation mechanism (e.g., see Christiano et al., 2010; Fernández-Villaverde et al., 2016). In our approach all of these concerns are entirely moot, for a simple reason: for counterfactual policy evaluation, all that matters are the reduced-form projections that are generated by the unknown shocks, and those projections can actually be taken directly from data.[16]

We provide some further discussion—that furthermore touches on some related concerns about weak identification—in Appendix B.4.

4 Inputs for monetary policy counterfactuals

We now showcase our methodology with several monetary policy applications. We begin in this section with the required inputs, following the general discussion in Sections 3.1 and 3.2.

4.1 Reduced-form projections

In Section 5 we will seek to evaluate the counterfactual evolution of the output gap, inflation, and nominal interest rates under alternative assumptions on the systematic conduct of monetary policy. Consistent with our discussion in Section 3.1, we here begin by estimating a relatively large-dimensional reduced-form VAR that contains these three core observables, as well as other aggregates that are useful to predict them. We only sketch the procedure here, with implementation details relegated to Appendix C.1.

For our estimated reduced-form VAR we consider a set of 10 macroeconomic variables, as in Angeletos et al. (2020). Differently from those authors, however, we will transform all

[16] A secondary, related benefit of our proposed approach is that the researcher does not need to specify the prevailing policy rule. In practice, it is unlikely that historical policy conduct can actually be reduced to any simple rule, and so the policy rule block of DSGE models is also likely to be mis-specified. Misspecification of policy rules has however attracted less attention in the literature, so our discussion emphasizes it less.

of the included variables to stationarity (if necessary), as in Hamilton (2018); in particular, we treat the detrended real output series as a measure of the output gap. Our sample period stretches from 1960:Q1 – 2019:Q4—a long post-war, pre-covid sample. For our covid inflation counterfactual application (in Section 5.4) we then further extend that sample to the chosen forecast date (2021:Q2). In Appendix C.1 we document that the forecasts implied by our reduced-form VAR are competitive with other forecasting strategies.

4.2 Monetary policy shock evidence

We continue with the second part of the VAR step—evidence on the effects of monetary policy shocks on our three main outcomes of interest. Our construction of the empirical targets $\hat{\theta}_\nu$ of monetary policy transmission follows recent advances in the empirical time-series monetary policy shock literature. Specifically, we identify a monetary policy shock series following Aruoba and Drechsel (2022), who extend the classic Romer and Romer (2004) analysis to allow for a larger policymaker information set. We then study the dynamic causal effects of this monetary policy shock series on the output gap as well as on inflation and on the federal funds rate. The vector $\hat{\theta}_\nu$ stacks impulse responses of these three variables for the first five years after the policy shock. We only provide a visual display of the main results here, with implementation details provided in Appendix C.2.

Results are reported as the grey lines (medians) and light grey areas (confidence bands) in Figure 2 (on p.26). The estimated effects of monetary shocks look largely as in prior work: a transitory rate hike leads to a gradual and moderately persistent fall in output, as well as a more delayed fall in prices. Together, the results of Sections 4.1 and 4.2 complete the VAR step of our approach: we have now recovered (i) the required reduced-form projections and (ii) empirical evidence on the propagation of one *type* of monetary shock—a transitory nominal rate change. Those inputs will turn out to largely govern the results in our first two applications in Sections 5.2 and 5.3; the remainder of this section discusses the Plus step, which will be key for the third application, in Section 5.4.

4.3 Models of monetary policy transmission

For our Plus step—matching and extrapolating beyond the evidence $\hat{\theta}_\nu$ of Section 4.2—we consider several standard models of monetary policy transmission. Our first model is a standard representative-agent model with nominal rigidities ("RANK"), augmented with several other frictions to allow a quantitative fit to our empirical VAR evidence, following Christiano et al. (2005). Our second model then enriches the consumer block to feature heterogeneous households with uninsurable income risk, e.g., as in Kaplan et al. (2018) ("HANK"). Those first two models arguably capture the dominant approaches to quantitative business-cycle modeling of the past two decades. Finally, we will also consider behavioral variants of these models, with price- and wage-setters forming expectations with cognitive discounting, as in Gabaix (2020). We will see that such behavioral frictions can matter greatly for policy dynamic causal effect extrapolation at intermediate and long horizons, and that this can have material implications for policy counterfactual evaluation. Zooming out, in the notation of our general framework of Section 2.1, this section introduces the model relations that deliver the private-sector—or, more accurately, non-monetary-policy—block $\{\mathcal{H}_w, \mathcal{H}_x, \mathcal{H}_z\}$.

The remainder of this section proceeds as follows. We will first sketch the representative-agent, rational-expectations model. We then explain how the heterogeneous-agent and behavioral models depart from this benchmark. Throughout we will only provide brief verbal descriptions, with details in Appendix C.3. We do so because all the models we consider are standard; our contribution is instead in how we *use* these models for the Plus step.

Baseline RANK model. Our first model is a standard quantitative business-cycle model, as familiar from the "medium-scale DSGE" tradition, and in particular rich enough to allow us to match the empirical monetary policy shock evidence documented in Section 4.2. Following Christiano et al., the model features capital accumulation subject to investment adjustment costs and with variable capacity utilization, nominal rigidities (with indexation) in price- and wage-setting, and habit formation in consumer preferences. We now provide a brief sketch of each of the constituent model blocks.

- *Households & unions.* The economy is populated by a representative household with separable preferences for consumption and leisure, and allowing for habit formation. This agent chooses paths for consumption and assets to maximize lifetime utility. Labor supply is intermediated by labor unions (as in Erceg et al., 2000), with households taking hours worked as given when solving their consumption-savings problem. The unions face Calvo-style frictions in adjusting their wages, with full indexation to lagged price inflation (as in Christiano et al., 2005).

- *Production.* There is a unit continuum of perfectly competitive intermediate goods producers. They produce using capital and labor, and with a variable rate of capital capacity utilization; they sell their good to retailers who costlessly differentiate it, and set prices subject to Calvo frictions. Prices that are not re-optimized are fully indexed to lagged inflation (as in Christiano et al., 2005). The intermediate goods producers purchase capital goods from competitive capital goods producers. Those capital goods producers purchase the final good, turn it into the capital good subject to adjustment costs on their level of investment, and sell the capital good.

- *Policy.* There is a monetary and a fiscal authority. The fiscal authority issues nominal bonds with exponential maturity structure, spends a constant amount in real terms, and then adjusts labor taxes gradually to maintain long-run budget balance. In the representative-agent economy described here, this fiscal rule has real effects through the distortionary effects of taxes on labor supply. In the heterogeneous-agent economy sketched below, it also matters by affecting the timing of household income.

  The monetary authority sets nominal interest rates. Importantly, for our purposes—i.e., estimating the model to match $\hat{\theta}_\nu$ and then extrapolating to all of Θ_ν—we need not specify any particular monetary policy rule, as discussed above.

Stacking all model blocks except the behavior of the monetary authority, we obtain $\{\mathcal{H}_w, \mathcal{H}_x, \mathcal{H}_z\}$—i.e., the first model block (2).[17] Solving the system for every possible path of monetary policy shocks and thus nominal rates, we obtain the linear map of monetary policy causal effects, Θ_ν. We estimate the model to ensure consistency between Θ_ν and the empirical monetary policy shock targets $\hat{\theta}_\nu$; to do so we rely on a standard Bayesian implementation of impulse response-matching, with details on the set of estimated parameters and on the choice of priors provided in Appendix C.4.

HANK MODEL. Our second model is a heterogeneous-agent ("HANK") model. It differs from the representative-agent baseline in that the representative consumer is replaced by a unit continuum of households subject to uninsurable idiosyncratic income risk and borrowing constraints (e.g., Kaplan et al., 2018), delivering elevated average marginal propensities to consume (MPCs). To ensure consistency with the empirically observed gradual response of output to changes in monetary policy, we furthermore assume that households are inattentive to macroeconomic conditions, as in Auclert et al. (2020). Unlike habits, this modeling choice delivers sluggish responses to changing aggregates while still maintaining large MPCs out of transitory income changes. The remainder of the model is unchanged.

[17]This in particular means that the fiscal rule is contained in the "non-policy" block (2). Our counterfactuals thus keep the fiscal rule fixed, and only change assumptions on monetary policy conduct.

Cognitive discounting. Standard New Keynesian models, including the representative- and heterogeneous-agent variants presented so far, imply that inflation is strongly forward-looking. This model feature implies that small changes in future monetary policy can have large and immediate effects on inflation. The literature on the forward guidance puzzle (e.g., Del Negro et al., 2023) has questioned this feature of standard models.

For our final set of model variants, we will consider versions of our representative- and heterogeneous-agent baselines in which price- and wage-setting becomes less forward-looking; specifically, we follow Gabaix (2020) in assuming that agents engage in cognitive discounting. According to this view, agents do not trust that they understand the structure of the economy and thus shrink their expectation of future outcomes towards the economy's steady state. In particular, an innovation occurring s periods in the future is down-weighed by a factor m^s, where $m \in [0, 1]$ controls the strength of cognitive discounting, and with $m = 1$ corresponding to the rational-expectations benchmark. Our behavioral models will feature $m = 0.65$, at the lower end of the range considered by Gabaix; we make this choice (rather than estimating m) because, as we will see, our targeted monetary policy shock causal effects are only weakly informative about m.

4.4 Estimation results and policy extrapolation

We now use the empirical targets of Section 4.2 together with the four models of Section 4.3 to implement our Plus step, assuming a uniform prior across models. We first present results of the impulse response matching exercise—i.e., the ability of our models to match the targets, and the resulting posterior distribution across models. We then discuss how the different models extrapolate from the matched transitory policy shock causal effects to more persistent policy changes.

Model estimation via impulse response matching. We begin in Figure 2 by showing the monetary policy shock estimation targets (grey) and as well as the matched impulses at the posterior mode for each of our four models (beige and red, solid and dashed); the full posterior distributions for the estimated parameters are presented in Appendix C.4. We see that all of the models are able to match the empirical targets quite well: a transitory hike in nominal interest rates leads to a hump-shaped decline in output as well as a delayed decline in inflation. This similar fit suggests that the data do not strongly distinguish between our four models; this visual impression is confirmed in Table 4.1, which shows posterior probabilities across models, computed following (18). We see that, for all four models, posterior

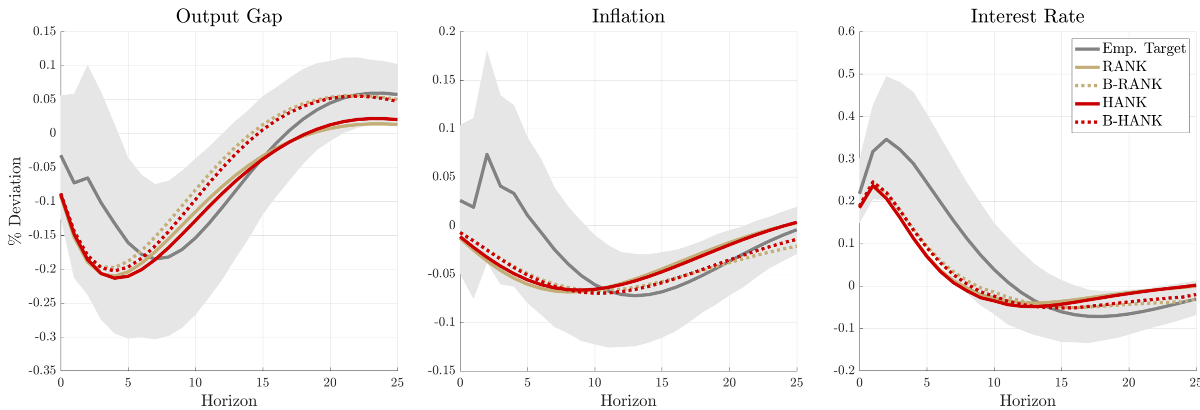

Figure 2: The grey line and shaded areas indicate the estimated impulse responses for a monetary policy shock (see Section 4.2), and their respective 16th and 84th percentile confidence bands. The remaining lines indicate the model-implied impulse responses at the estimated posterior modes. Beige: representative-agent consumer block. Red: heterogeneous-agent consumer block. Solid: no cognitive discounting. Dotted: cognitive discounting with $m = 0.65$.

probabilities have not moved particularly far from the uniform-prior starting point.[18]

Model	Baseline	Cognitive discounting	Total
RANK	0.4189	0.2435	0.6624
HANK	0.2183	0.1192	0.3376
Total	0.6372	0.3628	1.0000

Table 4.1: Posterior probabilities across the four models, assuming a uniform prior. The posterior model probabilities are computed as in (18).

Monetary policy shock extrapolation. Having estimated our models to match the empirical targets $\hat{\theta}_\nu$, we can now finally ask how the various estimated models extrapolate beyond the evidence on transitory interest rate changes to complete the full set of policy causal effects Θ_ν. We will proceed in two steps: first comparing the baseline representative- and heterogeneous-agent models, and then asking how cognitive discounting changes things. Results here are displayed in Figures 3 and 4, which show impulse responses to forward

[18]The analysis in Bundick and Smith (2020) as well as our results in Figures 3 and 4 suggest that empirical evidence on relatively near-term forward guidance would similarly be insufficient to discriminate between the models considered here. Furthermore, in results not reported here, we have also found that the fit across models is similar for the impulse responses of consumption and investment.

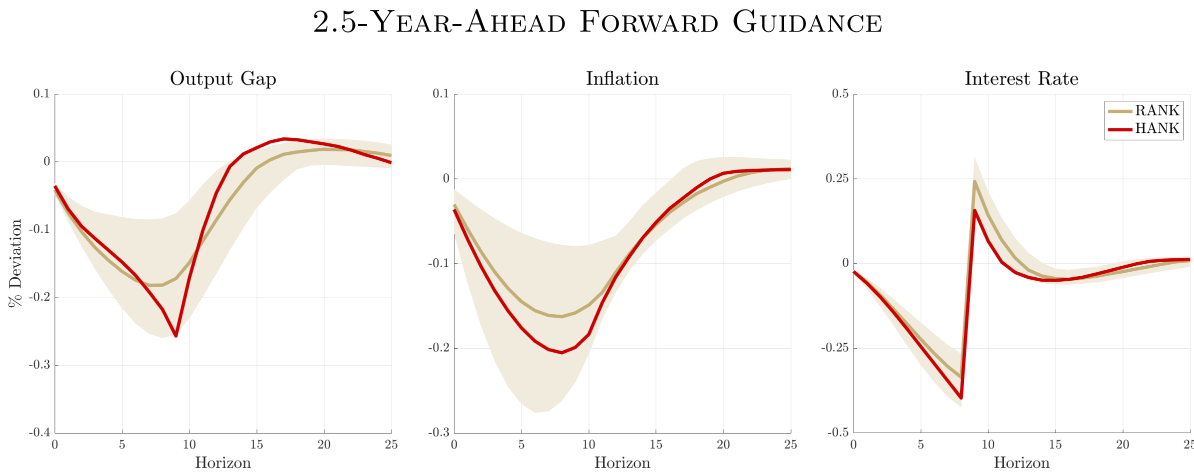

Figure 3: Policy causal effect extrapolation in our estimated RANK and HANK models. The figure shows output and inflation impulse responses (left and middle) to a forward guidance policy shock that leads to the interest rate movements depicted on the right. Beige: RANK model (solid = median, shaded = 16th and 84th percentile confidence bands). Red: HANK model (median).

guidance shocks—i.e., nominal interest rate movements that are (much) more delayed than our transitory targets $\hat{\theta}_\nu$.[19]

- *RANK vs. HANK.* We consider a far-ahead (forward guidance) monetary policy intervention: deviations from a standard monetary policy rule announced at $t = 0$ and occurring ten quarters later. The right panel shows the response of nominal interest rates, while the left and middle panels display the causal effects on output and inflation in RANK (beige) and HANK (red) associated with those interest rate paths.

  The main takeaway is that the two models, which closely agree on the effects of the targeted interest rate path by construction, also approximately agree on the dynamic causal effects of this much more delayed monetary intervention. In fact this is not just the case for the particular rate paths shown in Figure 3; it holds robustly across different possible paths, i.e., across the entirety of Θ_ν.

- *Baseline vs. cognitive discounting.* Figure 4 repeats the same exercise, but now comparing the baseline (beige) and behavioral (blue) representative-agent models.[20]

  We see that now there are more material differences across the two models. In par-

[19] Details on how we construct those particular delayed interest rate paths are provided in Appendix C.4.

[20] We use the term "behavioral" as shorthand for "with cognitive discounting." The baseline HANK model also has elements that could be considered behavioral (i.e., sticky information among consumers).

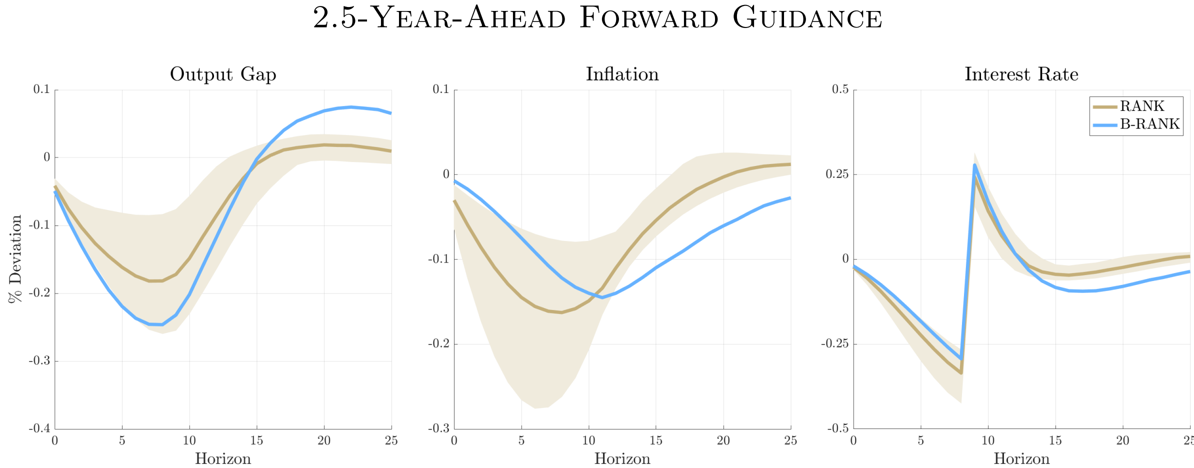

Figure 4: Policy causal effect extrapolation in our estimated RANK and B-RANK models. The figure shows output and inflation impulse responses (left and middle) to a forward guidance policy shock that leads to the interest rate movements depicted on the right. Beige: RANK model (solid = median, shaded = 16th and 84th percentile confidence bands). Blue: B-RANK model (median).

ticular, and as expected given the additional cognitive discounting, we see that the inflation responses in the behavioral model are more delayed. While Figure 4 shows this for the representative-agent models, we note that the same is also true for the heterogeneous-agent models.

Takeaways. In this section we have completed our Plus step—i.e., we have used structural models to match and then extrapolate beyond the empirical VAR evidence on monetary policy transmission, delivering a distribution over policy causal effects, Θ_ν. With an eye towards our applications in Section 5 we emphasize the following three takeaways.

1. By construction, the overall policy shock causal effects Θ_ν agree with the targeted empirical VAR evidence $\hat{\theta}_\nu$. Thus, as discussed in Section 3, counterfactuals that can (approximately) be enforced through those policy shock causal effects alone will be semi-structural; the Plus step discussed here is not needed in those cases.

2. Representative- and heterogeneous-agent models, once estimated to match the same empirical evidence on the effects of transitory interest rate changes, extrapolate very similarly to persistent rate changes. Thus, by our results in Section 2.3, it follows immediately that our two estimated RANK and HANK models of monetary policy transmission will imply very similar monetary policy counterfactuals in *all possible* applications.

3. Rational and behavioral models extrapolate quite differently to more persistent interest rate changes, with the inflation response in the former much more forward-looking. Thus, for counterfactuals involving such persistent monetary interventions, we expect the "Plus" step to matter greatly, and our method to indicate large posterior uncertainty.

The first takeaway will loom large in our first two applications in Sections 5.2 and 5.3, while the second and third play an important role for the last one, in Section 5.4.

5 Applications to monetary policy counterfactuals

We begin in Section 5.1 by describing our counterfactual assumptions on monetary policy design. In Sections 5.2 to 5.4 we evaluate how such alternative policy design would have shaped the average business cycle as well as two particular historical episodes.

5.1 The counterfactual policy experiment

For all three applications we will consider as our counterfactual monetary policy rule the one that minimizes the following standard "dual mandate" objective:[21]

$$\mathcal{L} = \mathbb{E}_0 \left[\sum_{t=0}^{\infty} \beta^t \left\{ \lambda_\pi \pi_t^2 + \lambda_y y_t^2 + \lambda_i (i_t - i_{t-1})^2 \right\} \right]. \tag{19}$$

This counterfactual is supposed to give us a reasonable approximation to the strategy of flexible inflation targeting, as discussed in Woodford (2003). Consistent with the discussion in the Federal Reserve Tealbook (2016), we consider an equal-weights parameterization, with $\lambda_\pi = \lambda_y = \lambda_i = 1$ (and no discounting, i.e., $\beta = 1$).

5.2 Average business cycle

For our first application we ask how the average U.S. post-war business cycle would have differed had the Federal Reserve (always) followed the flexible inflation targeting monetary policy rule (19). We communicate our main results by reporting the counterfactual volatilities of the output gap, inflation, and nominal interest rates.

[21] That is, given any matrix of monetary policy causal effects Θ_ν, we set the counterfactual policy rule coefficients $\{\tilde{\mathcal{A}}_x, \tilde{\mathcal{A}}_z\}$ exactly as in Proposition 2 of McKay and Wolf (2023).

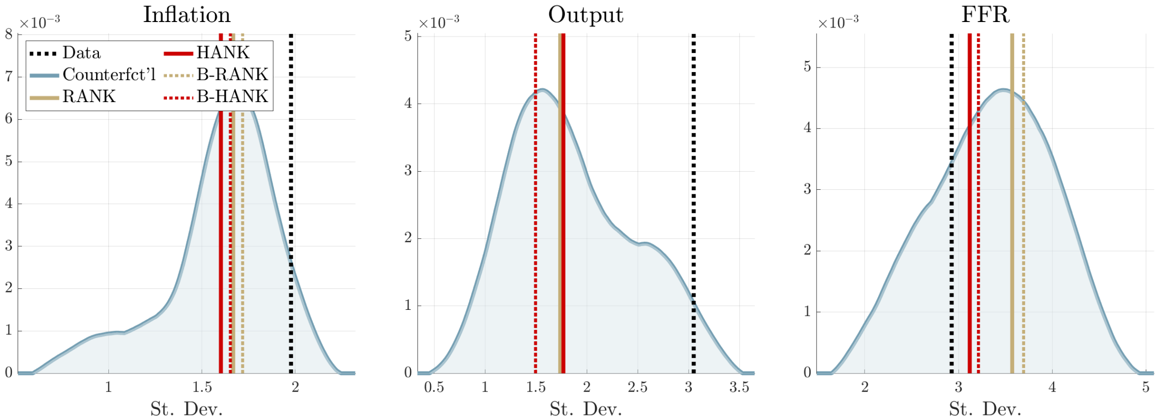

Figure 5: Counterfactual unconditional volatilities of inflation, output, and the federal funds rate, under the policy rule that minimizes (19). Black dashed: data point estimate under observed policy. Blue: posterior Kernel density of counterfactual volatilities drawing from posterior across all models and parameters. Beige: posterior mode of counterfactual using RANK models (baseline and behavioral). Red: posterior mode of counterfactual using HANK models (baseline and behavioral).

Main results. Results are reported in Figure 5. That figure shows actual (black) as well as counterfactual (colored) volatilities of inflation, the output gap, and the federal funds rate. The black-dashed lines are constructed by translating the estimated reduced-form VAR into its implied Wold representation, and then from there computing the three volatilities. To construct the counterfactual, we draw from the posterior over Θ_ν. For each draw, we then compute the volatilities under the counterfactual policy, following Proposition 1; finally, we construct a smoothed Kernel density estimate of the resulting posterior distribution. The colored vertical lines show the counterfactuals at the posterior modes of our four models.

The key finding is that the counterfactual policy could have actually achieved materially lower output gap volatilities and slightly more stable inflation, at the cost of only moderately more volatile nominal interest rates.[22] This finding is furthermore chiefly driven by our VAR evidence, and not by the causal effect extrapolation in the Plus step. To show this, we in Appendix D.1 repeat our analysis using only the empirically estimated policy causal effects $\hat{\theta}_\nu$ (rather than the entire model-implied matrix Θ_ν). That exercise yields the same qualitative patterns, with output gap volatility still dropping materially (if less than documented here), while inflation and interest rate volatilities do not change much. Another way of seeing the same result is to note that, in Figure 5, the beige and red, solid and dashed lines—i.e., the counterfactuals at the posterior modes of our four models—are all very close. The displayed

[22]Note that this volatility reduction does not just reflect infeasible rate cuts during the period of a binding lower bound on nominal interest rates; see Appendix D.1 for pre-2007 results.

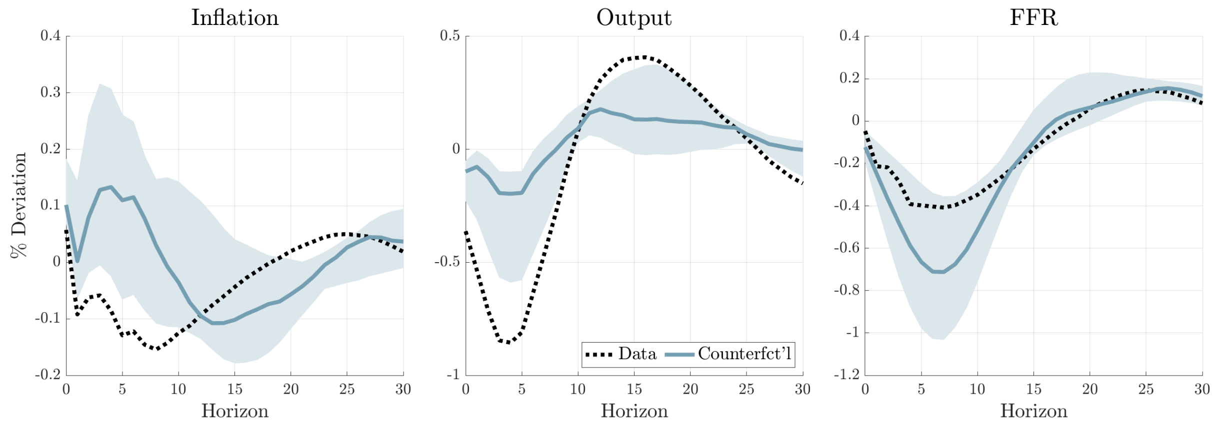

Figure 6: Counterfactual impulse responses of inflation, output, and the federal funds rate to the main business-cycle shock (see Angeletos et al., 2020), under the policy rule that minimizes (19). Black dashed: data point estimate under observed policy. Blue: posterior median (solid) and 16th and 84th percentile bands (shaded).

posterior uncertainty in Figure 5 (in light blue) thus chiefly reflects uncertainty about the causal effects of transitory interest rate changes, rather than any across-model uncertainty in how to extrapolate those policy causal effects.

Inspecting the mechanism. To see more clearly where those results are coming from, we study the counterfactual propagation of the combination of Wold residuals that "explains" the largest share of unemployment volatility at business-cycle frequencies—i.e., the "main business-cycle shock" of Angeletos et al. (2020). The black-dashed lines in Figure 6 show the propagation of this shock under the in-sample monetary reaction: inflation drops just a little, output drops materially, and monetary policy somewhat leans against this contraction. Our contemplated counterfactual policy leans against this shock much more, stabilizing output at the cost of moderately higher inflation and larger nominal interest rate movements.

The takeaway from Figure 6 is thus that our headline conclusions in Figure 5 are essentially driven by two moments of the data: first, that large output movements are associated with moderate inflation movements and partial interest rate offset; and second, that interest rate cuts boost output, with only little effect on inflation. Combining those two empirical moments—which are already well-documented from much prior work (notably Ramey, 2016; Angeletos et al., 2020)—with our identification results in Section 2 immediately delivers the conclusions in Figure 5; in other words, and as claimed, the key for everything reported here is the VAR step, with little incremental role for the Plus extrapolation.

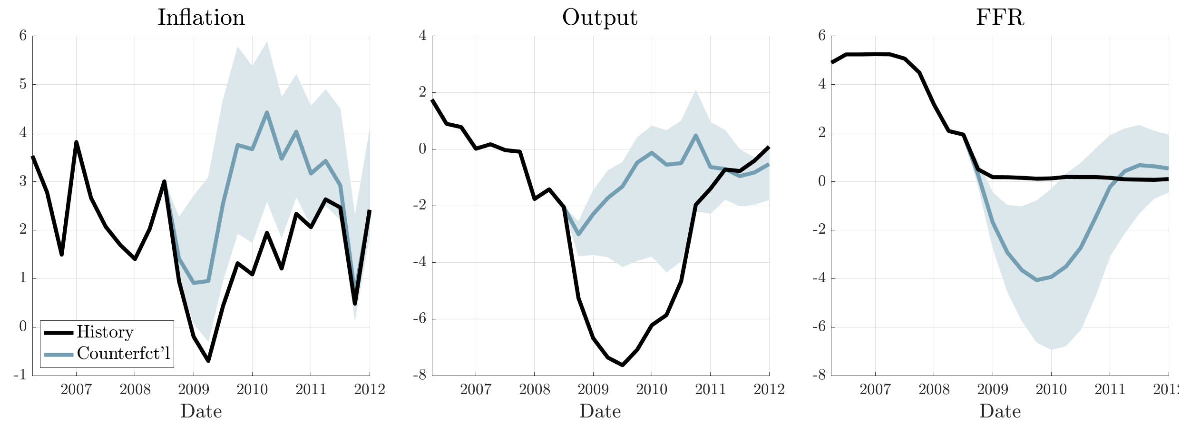

Figure 7: Counterfactual evolution of inflation, output, and the federal funds rate in the Great Recession, under the policy rule that minimizes (19) without any effective lower bound on rates. Black: data. Blue: posterior median (solid) and 16th and 84th percentile bands (shaded).

5.3 Great Recession

For our second application we evaluate how the economy would have evolved during the Great Recession if monetary policy had followed the inflation targeting framework described above, and *without* any effective lower bound on nominal interest rates.[23] Specifically, we will assume that the central bank follows this alternative, unconstrained rule from 2008:Q4 onwards, and does so throughout 2012:Q1—an example of a "historical evolution" counterfactual. A counterfactual of this sort is informative about the plausible costs of a binding lower bound constraint. Results are reported in Figure 7.

We see that, absent any effective binding lower bound on nominal interest rates, a policy that follows the rule of minimizing (19) would have involved a very aggressive rate cut, down to around -4 per cent. Such (infeasible) rate cuts would have materially reduced the output gap, at the cost of moderately elevated inflation. As in Section 5.2, and for the same reasons as discussed there, this counterfactual is almost exclusively governed by our VAR evidence (and not by the Plus extrapolation, see Appendix D.2); in particular, and again as before, the uncertainty displayed in Figure 7 chiefly reflects uncertainty about the causal effects of transitory rate changes, and not extrapolation uncertainty across models.

Our results are informative about the broader policy response to the Great Recession. Given constraints on nominal interest rates, policymakers attempted to substitute through

[23]We note that our methodology remains applicable to model environments with a linear non-policy block (2) and a non-linear policy rule, allowing for a binding lower bound on nominal interest rates. The argument is analogous to that in Appendix A.9 of McKay and Wolf (2023).

other stimulative measures, notably unconventional monetary policy as well as fiscal stimulus. If we interpret (19) as the objective for monetary policy, our counterfactual suggests that the unconventional monetary policy response was insufficient—in nominal interest rate space, additional stimulus of around 400 basis points would have been necessary.

5.4 Post-covid inflation

As the third and final application of our methodology, we evaluate monetary policy options at the height of the post-covid inflation. Using our estimated reduced-form VAR, we construct forecasts of inflation, the output gap, as well as the federal funds rate from 2021:Q2 onwards. We then apply our methodology to construct the forecast under the policy that minimizes the dual-mandate inflation targeting objective (19).

MAIN RESULTS. Figure 8 shows the historical evolution of inflation, the output gap, and the federal funds rate (black), their baseline forecasts from 2021:Q2 onwards (grey-dashed), and their counterfactual forecasts (blue). The panels distinguish the models used to extrapolate policy causal effects: the two rational-expectations models (top), the two behavioral models (middle), and the full set (bottom). We can see that, under the baseline forecast, inflation is expected to be persistently elevated, the output gap is slightly negative, and nominal interest rates rise sharply. Our focus is now on how the inflation targeting monetary policy moves the economy away from these baseline forecasts.

Consider first the top panel of Figure 8, which shows counterfactuals constructed with policy extrapolation via our rational-expectations models. We see that policy succeeds in reducing inflation sharply with only a small reduction in output. Furthermore, and counterintuitively, a *lower* interest rate path achieves this disinflation. This result reflects the extremely forward-looking nature of the model: the policymaker achieves low inflation in the short-run via negative output gaps in the far-future, implemented through future increases in real rates; in fact, small future output gaps move current inflation so much that *lower* short-term real rates can actually be used to stabilize output in the short run. The beige lines in Figure 9 illustrate this intuition: real rates initially decline and only later rise (left panel); short-run inflation is much more sensitive to real rates in the far-future than to real rates today (middle panel); combining the two, it follows that near-term disinflation can be achieved through moderate medium- and long-term real rate hikes (right panel).

Consider next the middle panel of Figure 8, which constructs our policy counterfactuals using instead the behavioral models to extrapolate policy causal effects. The counterfactual

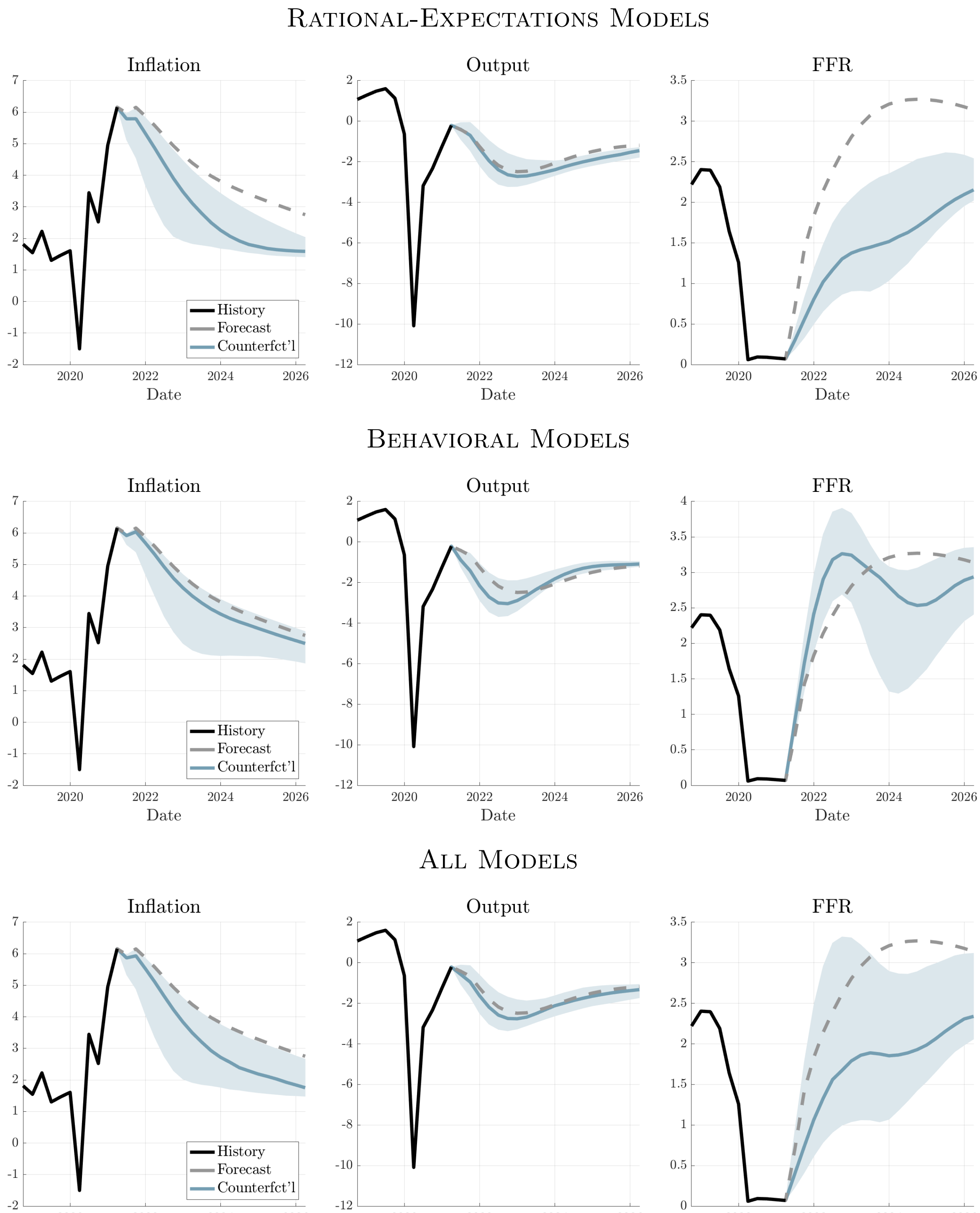

Figure 8: Counterfactual projections of inflation, output, and the federal funds rate in the post-covid inflationary episode (from 2021:Q2), under the policy rule that minimizes (19). Policy causal effects from rational-expectations models (top), behavioral models (middle), and all models jointly (bottom). Black: data. Grey: actual (VAR-implied) forecast. Blue: posterior median (solid) and 16th and 84th percentile bands (shaded).

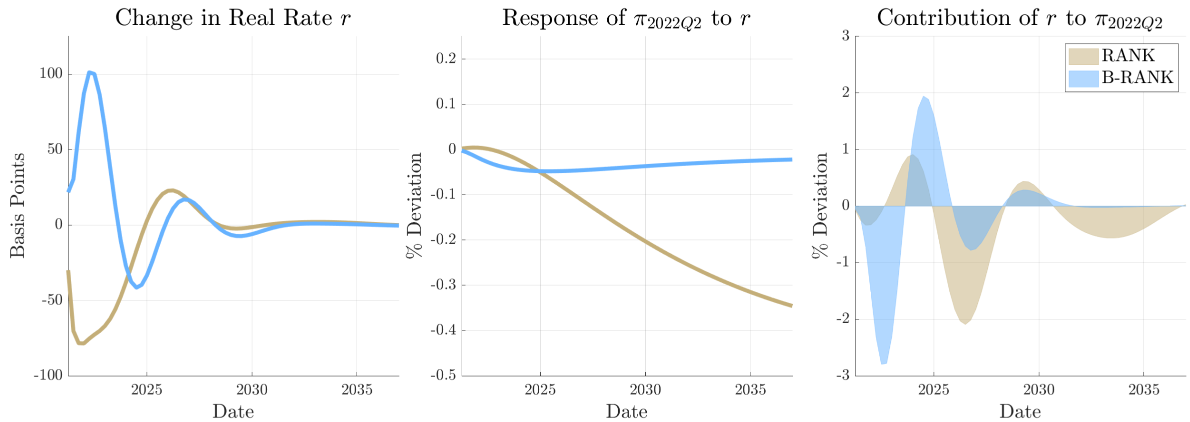

Figure 9: Behavior of real interest rates under the counterfactual monetary policy. Left panel: difference between the counterfactual and the forecast real interest rate path. Middle panel: change in inflation at 2022:Q2 in response to a policy-induced real interest rate change by horizon of the rate change, as indicated on the x-axis. Right panel: contribution of real interest rate changes at different horizons to the change in inflation at 2022:Q2. Beige: posterior mode of the baseline RANK model. Blue: posterior mode of the behavioral RANK model.

looks quite different: the federal funds rate is hiked somewhat *more* aggressively than in the baseline forecast, thus bringing inflation down slightly, though at the cost of moderately lower output. Intuitively, the policymaker now cannot rely on very far-ahead real interest rate movements to stabilize inflation in the short run. Instead, she faces an undesirable short-run trade-off between output and inflation, and chooses to respond to it through *higher* short-term real rates. A visual illustration is provided with the blue lines in Figure 9: real interest rates rise immediately (left panel); short-run inflation is not nearly as sensitive to long-run real rate fluctuations (middle panel); as a result, the short-term inflation reduction largely reflects short-term real rate movements.

Finally, the bottom panel of Figure 8 puts everything together, showing counterfactuals with policy causal effects extrapolated from our full set of four models. The main message of the figure is the very large uncertainty about the response of interest rates. The four models are similar in their abilities to fit our estimation targets—i.e., the effects of a transitory monetary shock—yet they differ in their predictions for the effects of far-ahead rate changes. As this counterfactual features persistent interest rate changes, the disagreement across models translates to considerable uncertainty regarding the desired path of output and inflation.

Discussion. Figure 8 connects closely with the "forward guidance puzzle" literature (e.g., see Del Negro et al., 2023; Carlstrom et al., 2015; McKay et al., 2016). Our results demonstrate that the model pathologies uncovered in that literature matter even away from any binding lower bound—they will apply whenever the economy is persistently away from its steady state. As shown in Section 4.4, the available empirical evidence on transitory changes in monetary policy does not discriminate between models with dramatically different causal effects of persistent policy changes; as a result, for monetary policy counterfactuals involving persistent interest rate changes, our method will invariably indicate large posterior policy counterfactual uncertainty.

This discussion suggests that, for the purposes of counterfactual policy evaluation, there will be high returns to future empirical work on the causal effects of persistent changes in monetary policy. Failing that, relying on other pieces of evidence to discriminate across the kinds of models considered here should prove useful. In particular, our results indicate that researchers should aim to distinguish between models with and without behavioral frictions; on the other hand, market incompleteness—the focus of the recent "HANK" literature—seems less central for the purpose of monetary policy causal effect extrapolation.

6 Conclusions

The main contribution of this paper is to propose a new "VAR-Plus" approach to evaluating the counterfactual evolution of the macro-economy under alternative assumptions on policy design. Leveraging a theoretical identification result, our method relies on reduced-form or semi-structural empirical evidence as much as possible (VAR), and then complements that evidence with additional structural assumptions whenever necessary (Plus). This approach is widely applicable, yet appealingly robust: if empirical evidence already suffices to pin down the counterfactual, then our approach is in fact semi-structural; and even if the evidence does not suffice, the only role of structure is to extrapolate the effects of policy shocks, with no need to develop a complete account of the origins of business-cycle fluctuations. As a result, our approach can sidestep many of the literature's concerns with the dominant "quantitative DSGE" approach to policy evaluation (as, e.g., articulated in Chari et al., 2009).

Our analysis suggests at least three avenues for future work. First, empirically, there are very high returns to analyses identifying the causal effects of persistent changes in monetary policy—i.e., the causal effects of forward guidance-type policies. Second, theoretically, it would be useful to gain a more complete understanding of the range of models that can be consistent with the available evidence on monetary policy propagation, and then of how they differ in extrapolating beyond that evidence. Third, it would be interesting to go beyond monetary policy and instead apply our insights to fiscal questions.

References

ANGELETOS, G.-M., F. COLLARD, AND H. DELLAS (2020): "Business-cycle anatomy," *American Economic Review*, 110, 3030–70.

ANTOLIN-DIAZ, J., I. PETRELLA, AND J. F. RUBIO-RAMIREZ (2021): "Structural scenario analysis with SVARs," *Journal of Monetary Economics*, 117, 798–815.

ARIAS, J. E., J. F. RUBIO-RAMÍREZ, AND D. F. WAGGONER (2018): "Inference based on structural vector autoregressions identified with sign and zero restrictions: Theory and applications," *Econometrica*, 86, 685–720.

ARKOLAKIS, C., A. COSTINOT, AND A. RODRÍGUEZ-CLARE (2012): "New trade models, same old gains?" *American Economic Review*, 102, 94–130.

ARUOBA, S. B. AND T. DRECHSEL (2022): "Identifying Monetary Policy Shocks: A Natural Language Approach," CEPR Discussion Paper No. DP17133.

AUCLERT, A., B. BARDÓCZY, M. ROGNLIE, AND L. STRAUB (2021): "Using the sequence-space Jacobian to solve and estimate heterogeneous-agent models," *Econometrica*, 89, 2375–2408.

AUCLERT, A., R. D. RIGATO, M. ROGNLIE, AND L. STRAUB (2022): "New Pricing Models, Same Old Phillips Curves?" Tech. rep., National Bureau of Economic Research.

AUCLERT, A., M. ROGNLIE, AND L. STRAUB (2020): "Micro jumps, macro humps: Monetary policy and business cycles in an estimated HANK model," Tech. rep., National Bureau of Economic Research.

BARNICHON, R. AND G. MESTERS (2023): "A Sufficient Statistics Approach for Macro Policy," *American Economic Review*, 113, 2809–2845.

——— (2024): "Evaluating policy institutions-150 years of us monetary policy," Woking Paper.

BUNDICK, B. AND A. L. SMITH (2020): "The dynamic effects of forward guidance shocks," *Review of Economics and Statistics*, 102, 946–965.

Carlstrom, C. T., T. S. Fuerst, and M. Paustian (2015): "Inflation and output in New Keynesian models with a transient interest rate peg," *Journal of Monetary Economics*, 76, 230–243.

Chari, V. V., P. J. Kehoe, and E. R. McGrattan (2009): "New Keynesian models: Not yet useful for policy analysis," *American Economic Journal: Macroeconomics*, 1, 242–266.

Christiano, L. J., M. Eichenbaum, and C. L. Evans (2005): "Nominal rigidities and the dynamic effects of a shock to monetary policy," *Journal of Political Economy*, 113, 1–45.

Christiano, L. J., M. Trabandt, and K. Walentin (2010): "DSGE models for monetary policy analysis," in *Handbook of monetary economics*, Elsevier, vol. 3, 285–367.

Crump, R. K., M. Del Negro, K. Dogra, P. Gundam, D. Lee, R. Nallamotu, and B. Pacula (2023): "The New York Fed DSGE Model Perspective on the Lagged Effect of Monetary Policy," Tech. rep., Federal Reserve Bank of New York.

Del Negro, M., M. P. Giannoni, and C. Patterson (2023): "The forward guidance puzzle," *Journal of Political Economy Macroeconomics*, 1, 43–79.

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

Eberly, J. C., J. H. Stock, and J. H. Wright (2020): "The Federal Reserve's Current Framework for Monetary Policy: A Review and Assessment," *International Journal of Central Banking*, 16, 5–71.

Erceg, C. J., D. W. Henderson, and A. T. Levin (2000): "Optimal monetary policy with staggered wage and price contracts," *Journal of Monetary Economics*, 46, 281–313.

Federal Reserve Tealbook (2016): "Report to the FOMC on Economic Conditions and Monetary Policy," .

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

Fernández-Villaverde, J., J. F. Rubio-Ramírez, and F. Schorfheide (2016): "Solution and estimation methods for DSGE models," in *Handbook of macroeconomics*, Elsevier, vol. 2, 527–724.

Frisch, R. (1933): "Propagation problems and impulse problems in dynamic economics," *Essays in Honor of Gustav Cassel*.

Gabaix, X. (2020): "A behavioral New Keynesian model," *American Economic Review*, 110, 2271–2327.

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

Geweke, J. (1999): "Using simulation methods for bayesian econometric models: inference, development,and communication," *Econometric Reviews*, 18, 1–73.

Hamilton, J. D. (2018): "Why You Should Never Use the Hodrick-Prescott Filter," *The Review of Economics and Statistics*, 100, 831–843.

Hebden, J. and F. Winkler (2021): "Impulse-Based Computation of Policy Counterfactuals," FEDS Working Paper.

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

Känzig, D. R. (2021): "The macroeconomic effects of oil supply news: Evidence from OPEC announcements," *American Economic Review*, 111, 1092–1125.

Kaplan, G., B. Moll, and G. L. Violante (2018): "Monetary policy according to HANK," *American Economic Review*, 108, 697–743.

Kilian, L. and H. Lütkepohl (2017): *Structural vector autoregressive analysis*, Cambridge University Press.

Li, D., M. Plagborg-Møller, and C. K. Wolf (2023): "Local projections vs. VARs: Lessons from Thousands of DGPs," Working Paper.

Lucas, Robert E., J. (1972): "Expectations and the Neutrality of Money," *Journal of Economic Theory*, 4, 103–124.

Lucas, R. (1976): "Econometric policy evaluation: A critique," in *Carnegie-Rochester Conference Series on Public Policy*, Elsevier, vol. 1, 19–46.

McKay, A., E. Nakamura, and J. Steinsson (2016): "The power of forward guidance revisited," *American Economic Review*, 106, 3133–3158.

McKay, A. and C. Wolf (2022): "Optimal policy rules in hank," Tech. rep., Working Paper, FRB Minneapolis.

McKay, A. and C. K. Wolf (2023): "What Can Time-Series Regressions Tell Us About Policy Counterfactuals?" Working Paper.

Molavi, P. (2019): "Macroeconomics with learning and misspecification: A general theory and applications," *Unpublished manuscript*.

Montiel Olea, J. L., M. Plagborg-Møller, E. Qian, and C. K. Wolf (2024): "Double Robustness of Local Projections and Some Unpleasant VARithmetic," Tech. rep., National Bureau of Economic Research.

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

——— (2022): "Instrumental variable identification of dynamic variance decompositions," *Journal of Political Economy*, 130, 2164–2202.

Ramey, V. A. (2016): "Macroeconomic shocks and their propagation," *Handbook of macroeconomics*, 2, 71–162.

Romer, C. D. and D. H. Romer (2004): "A new measure of monetary shocks: Derivation and implications," *American Economic Review*, 94, 1055–1084.

Rotemberg, J. J. and M. Woodford (1997): "An optimization-based econometric framework for the evaluation of monetary policy," *NBER macroeconomics annual*, 12, 297–346.

Sims, C. A. (1980): "Macroeconomics and reality," *Econometrica: journal of the Econometric Society*, 1–48.

Sims, C. A. and T. Zha (1995): "Does monetary policy generate recessions?" Working Paper.

Sims, E. R. (2012): "News, non-invertibility, and structural VARs," in *DSGE Models in Macroeconomics: Estimation, Evaluation, and New Developments*, Emerald Group Publishing Limited, 81–135.

Smets, F. and R. Wouters (2007): "Shocks and frictions in US business cycles: A Bayesian DSGE approach," *American economic review*, 97, 586–606.

Stock, J. H. and M. W. Watson (2016): "Dynamic factor models, factor-augmented vector autoregressions, and structural vector autoregressions in macroeconomics," in *Handbook of macroeconomics*, Elsevier, vol. 2, 415–525.

Svensson, L. E. (1997): "Inflation forecast targeting: Implementing and monitoring inflation targets," *European economic review*, 41, 1111–1146.

Wolf, C. K. (2020): "Svar (mis) identification and the real effects of monetary policy shocks," *American Economic Journal: Macroeconomics*, 12, 1–32.

——— (2023): "The missing intercept: A demand equivalence approach," Tech. rep., National Bureau of Economic Research.

Woodford, M. (2003): *Interest and prices: Foundations of a theory of monetary policy*, Princeton University Press.

Online Appendix for: Evaluating Policy Counterfactuals: A VAR-Plus Approach

This appendix contains supplemental material for the article "Evaluating Policy Counterfactuals: A "VAR-Plus" Approach." We here provide: (i) some additional theoretical results to complement the discussion in Section 2; (ii) practical implementation details for our method as described in Section 3; (iii) further information on our "VAR-Plus" inputs for the monetary policy applications as discussed in Section 4; and (iv) supplementary application results to complement Section 5.

Any references to equations, figures, tables, assumptions, propositions, lemmas, or sections that are not preceded by "A."—"D." refer to the main article.

Contents

A Supplementary theoretical results **44**

A.1 Behavioral models . . . 44

A.2 More on the role of invertibility . . . 45

A.3 Counterfactual analysis in Smets and Wouters (2007) . . . 47

B Further details for our method **48**

B.1 Reduced-form projections . . . 48

B.2 Impulse response target estimation . . . 48

B.3 Model estimation . . . 49

B.4 Vulnerabilities of the quantitative DSGE approach . . . 52

C Monetary policy “VAR-Plus” inputs **55**

C.1 Reduced-form projections . . . 55

C.2 Empirical evidence on monetary shock propagation . . . 57

C.3 Models of monetary policy transmission . . . 58

C.4 Model calibration and estimation . . . 68

D Supplementary details for empirical applications **72**

D.1 Average business cycle . . . 72

D.2 Great Recession . . . 72

D.3 Post-covid inflation . . . 74

A Supplementary theoretical results

This appendix provides several supplementary theoretical results. Appendix A.1 discusses how our identification result applies in environments with behavioral frictions. Appendix A.2 elaborates on the role played by invertibility. Finally Appendix A.3 gives further details for our illustrative example based on Smets and Wouters (2007).

A.1 Behavioral models

In this subsection we discuss to what extent our theoretical identification results are consistent with deviations from the usual full-information, rational-expectations (FIRE) benchmark. We first clarify what kinds of behavioral frictions are admissable (and which are not), and then explain why, to implement our methodology, researchers still always need to try to construct *full-information* forecasts, consistent with our main-text discussion.

Nested behavioral models. Every structural environment that can be written in the general form (2) - (3) is consistent with our identification results. Importantly, this contains models with behavioral frictions in which the deviation from FIRE is *independent of the policy rule*, in the sense that the behavioral friction is encoded in $\mathcal{H}_w, \mathcal{H}_x, \mathcal{H}_z, \mathcal{H}_e$, and does not change as the policy rule changes.

More formally, begin by considering a model with FIRE, and consider the ith equation in its non-policy block, written as

$$\mathcal{H}_{i,w}^R \boldsymbol{w} + \mathcal{H}_{i,x}^R \boldsymbol{x} + \mathcal{H}_{i,z}^R \boldsymbol{z} + \mathcal{H}_{i,e}^R e_0 = \mathbf{0}. \tag{A.1}$$

A typical example of such a block would be the aggregate consumption function, mapping sequences of household income and asset returns into a path of consumption. Our theory is consistent with behavioral frictions in which the model equation (A.1) is replaced by an alternative of the form

$$\mathcal{H}_{i,w}^B \boldsymbol{w} + \mathcal{H}_{i,x}^B \boldsymbol{x} + \mathcal{H}_{i,z}^B \boldsymbol{z} + \mathcal{H}_{i,e}^B e_0 = \mathbf{0}. \tag{A.2}$$

where the matrices $\mathcal{H}_{i,\bullet}^B$ are a *policy rule-invariant transformation* of $\mathcal{H}_{i,\bullet}^R$:

$$\mathcal{H}_{i,\bullet}^B(\theta) = f(\mathcal{H}_{i,\bullet}^R, \theta),$$

with the parameter vector θ governing the behavioral friction. Examples of behavioral frictions that can be written in this general way include sticky information, sticky expectations, cognitive discounting, level-k thinking, and diagnostic expectations; see Auclert et al. (2021) for further details. Crucially, in all of these cases, agent behavior continues to be shaped by policy only through the current and expected future (full-information) paths of $\{\boldsymbol{w}, \boldsymbol{x}, \boldsymbol{z}\}$, and so our identification results continue to apply.[24]

Implications for forecasting. The previous discussion reveals that, even if the underlying data-generating process features behavioral frictions (of the sort consistent with our identification result, of course), the forecasts that appear in (2) - (3) and thus our identification results are *full-information* forecasts. It follows that, when leveraging our "VAR-Plus" approach, researchers should aim to construct such full-information forecasts, and then note that their reported counterfactuals will be valid across a wide range of models with and without underlying behavioral frictions.

A.2 More on the role of invertibility

We begin by establishing, consistent with the intuition given throughout Section 2, that the sole purpose of invertibility is to generate full-information forecasts. Afterwards we provide some further discussion of what happens in the absence of invertibility.

Invertibility and forecasts. We provide a constructive argument showing that access to full-information forecasts is sufficient to recover our counterfactuals. For this it will prove convenient to reverse the order relative to the arguments in Proposition 1, beginning instead with counterfactuals for conditional episodes.

2. **Conditional episodes.**

   (i) *Conditional forecasts.* Recall that we need to construct $\tilde{y}^*_{t^*+h}$ and $\tilde{\Theta}_h \varepsilon_{t^*}$. Given the full-information forecasts $\mathbb{E}_{t^*-1}[y_{t+h}]$, $\tilde{y}^*_{t^*+h}$ can be constructed from Θ_ν exactly as in the proof of Proposition 1. Next note that $\Theta_h \varepsilon_{t^*}$ can be recovered as the *revision* in full-information forecasts

$$\Theta_h \varepsilon_{t^*} = (\mathbb{E}_{t^*} - \mathbb{E}_{t^*-1})[y_{t^*+h}].$$

[24] A simple example that violates this restriction is the misspecified learning model of Molavi (2019). Here, a change in policy rule affects agent learning and thus alters the response to any given set of (full-information) expected future time paths, breaking the policy rule independence that we require. Mathematically, this is isomorphic to how incomplete information as in Lucas (1972) breaks our identification results.

We can then just as before use Θ_ν to turn those expectation revisions into $\tilde{\Theta}_h \varepsilon_{t^*}$, completing the argument.

(ii) *Historical evolution.* We now need to construct $\tilde{y}_t^1$ as well $\sum_{\ell=0}^{t-t_1} \tilde{\Theta}_\ell \varepsilon_{t-\ell}$. As in the previous item, $\tilde{y}_t^1$ can still be computed directly from date-$t_1 - 1$ forecasts, as in the proof of Proposition 1. Next, for date t_1, we obtain $\Theta_h \varepsilon_{t_1}$ from forecast revisions as $(\mathbb{E}_{t_1} - \mathbb{E}_{t_1-1})[y_{t_1+h}]$, and then use Θ_ν to get the counterfactual $\tilde{\Theta}_h \varepsilon_{t_1}$, thus in particular giving $\tilde{y}_{t_1} = \tilde{\Theta}_0 \varepsilon_{t_1} + \tilde{y}_t^1$. Proceeding recursively, we for time $\check{t}$ obtain forecast revisions to get $\Theta_h \varepsilon_{\check{t}}$, and so as usual via Θ_ν recover $\tilde{\Theta}_h \varepsilon_{\check{t}}$. From here we then get the date-$\check{t}$ realized counterfactual outcome as

$$\tilde{y}_{\check{t}} = \tilde{\Theta}_0 \varepsilon_{\check{t}} + \underbrace{\sum_{\ell=1}^{\check{t}-t_1} \tilde{\Theta}_\ell \varepsilon_{\check{t}-\ell} + \tilde{y}_t^1}_{\text{from previous steps}},$$

completing the argument.

1. **Unconditional business cycles.** Proposition 1 presupposes knowledge of the autocovariance function of the observables y_t or, equivalently, access to an arbitrarily large sample of observations $\{y_t\}_{t=0}^\infty$. By the discussion in the previous item, knowledge of full-information forecasts suffices to instead construct an arbitrarily large counterfactual sample $\{\tilde{y}_t\}_{t=0}^\infty$, thus delivering the counterfactual autocovariance function $\tilde{\Gamma}(\ell)$.

From this discussion it follows that, conditional on full-information forecasts being observable, invertibility ceases to be necessary. Our empirical implementation of the VAR step is designed with this observation in mind.

Proposition 1 without invertibility. Without invertibility, the orthogonalized reduced-form residuals u_t satisfy (e.g., see Wolf, 2020)

$$u_t = P(L)\varepsilon_t.$$

The Wold lag polynomial $\Psi(L)$ then satisfies

$$\Psi(L)P(L) = \Theta(L),$$

Using that $P(L)P^*(L^{-1}) = I$, it then follows from the arguments in McKay and Wolf (2023) that the artificial Wold lag polynomial $\tilde{\Psi}(L)$ constructed in our proof of Proposition 1 satisfies

$$\tilde{\Psi}(L) = \tilde{\Theta}(L)P^*(L^{-1}).$$

Proceeding from here, however, the proof strategy of Proposition 1 now fails, as it is generally the case that $P^*(L^{-1})P(L) \neq I$.

The results in Section 2.4 furthermore reveal that it is not just our particular proof *strategy* that fails here—without invertibility, Wold-implied forecasts are generally not equal to full-information forecasts, and so the derived counterfactuals do not equal the truth (though they may be close, of course). Mathematically, the problem is that, while the true lag polynomial $\Theta(L)$ and the Wold lag polynomial $\Psi(L)$ generate the same autocovariance function, nothing guarantees that the counterfactual lag polynomials $\tilde{\Theta}(L)$ and $\tilde{\Psi}(L)$ will also generate the same second moments. It is only the assumption of invertibility—which ties the impulse responses in the lag polynomials $\Theta(L)$ and $\Psi(L)$ together in a particular way—that allows this argument to go through.

A.3 Counterfactual analysis in Smets and Wouters (2007)

Our laboratory data-generating process for the illustrations in Section 2.4 is the well-known model of Smets and Wouters, but with one minor change—we assume that the monetary authority follows rules of the form

$$i_t = \rho_i i_{t-1} + (1 - \rho_i)(\phi_\pi \pi_t + \phi_y y_t), \tag{A.3}$$

which is slightly simpler than the headline specification considered by Smets and Wouters.

Specifically, we assume that the researcher observes data generated from the posterior mode parameterization of the Smets and Wouters model, but with the monetary policy rule taking the particular form (A.3) with $\phi_\pi = 1.5$ and $\rho_i = \phi_y = 0$. She then wishes to predict the counterfactual second-moment properties of interest if instead the monetary authority followed the rule (A.3) but with $\rho_i = 0.8$, $\phi_\pi = 1.5$ and $\phi_y = 0.5$. We have chosen these two particular policy rules because they imply quite starkly different second-moment properties, with the first one aggressively stabilizing inflation, while the second one smoothes interest rates and also stabilizes output. This allows us to most transparently illustrate our results about counterfactual accuracy under non-invertibility, as displayed in Figure 1.

B Further details for our method

This appendix provides supplementary details for our proposed methodology. Appendix B.1 discusses estimation of reduced-form projections, Appendix B.2 elaborates on how to obtain the (asymptotic) distribution of the policy shock targets $\hat{\theta}_\nu$, and Appendix B.3 gives further details on the model estimation (i.e., "Plus") step. Finally, in Appendix B.4, we provide an extended discussion of well-documented vulnerabilities of the conventional DSGE approach.

B.1 Reduced-form projections

Our approach begins with estimation of the reduced-form VAR in (16). A textbook discussion of how to estimate reduced-form VARs and translate the autoregressive lag polynomial $A(L)$ into the implied Wold lag polynomial $\Psi(L)$ is provided, for example, in Kilian and Lütkepohl (2017). Appendix C.1 discusses the concrete implementation in our application, including details on data, lag length selection, and variable transformations.

B.2 Impulse response target estimation

Consistent with the recommendations of Plagborg-Møller and Wolf (2021) and Li et al. (2023), the researcher first uses a structural VAR to estimate the causal effects of policy shocks, identified using one (or several) of the usual semi-structural time series identification approaches. We propose to estimate this VAR using standard Bayesian techniques, delivering draws $i = 1, 2, \ldots, N$ of the policy shock causal effect vector $\hat{\theta}_{i,\nu}$.

We obtain $\hat{\theta}_\nu$ as the posterior mode of the estimated policy shock causal effects. For V_{θ_ν} we proceed as follows. We construct

$$\bar{V}_{\theta_\nu} \equiv \sum_{i=1}^{N} \left(\hat{\theta}_{i,\nu} - \hat{\theta}_\nu\right) \left(\hat{\theta}_{i,\nu} - \hat{\theta}_\nu\right)'.$$

Since the small-sample properties of estimating V_{θ_ν} in this way are poor, we instead work with a sample size-dependent transformation of $\bar{V}_{\theta_\nu}$, following Christiano et al. (2010):

$$V_{\theta_\nu} = f(\bar{V}_{\theta_\nu}, T)$$

where T is the sample size. The transformation $f(\bullet)$ has the following properties. First, V_{θ_ν} and $\bar{V}_{\theta_\nu}$ have the same diagonal entries. Second, for off-diagonal entries that correspond to the ℓth and jth lagged response of a common variable to a common shock, it scales down the entry of $\bar{V}_{\theta_\nu}$ by

$$\left(1 - \frac{|\ell - j|}{\bar{H}_T}\right)^{\eta_{1,T}}, \quad \ell, j = 1, 2, \ldots, \bar{H}_T \tag{B.1}$$

where $\bar{H}_T \leq H$ and $\bar{H}_T \to H$, $\eta_{1,T} \to 0$ as $T \to \infty$. Third, for all other off-diagonal entries corresponding ℓth and jth lagged responses, it scales down the entry of $\bar{V}_{\theta_\nu}$ by

$$\zeta_T \left(1 - \frac{|\ell - j|}{\bar{H}_T}\right)^{\eta_{2,T}}, \quad \ell, j = 1, 2, \ldots, \bar{H}_T \tag{B.2}$$

where $\zeta_T \to 1$ and $\eta_{2,T} \to 0$ as $T \to \infty$. Intuitively, this transformation dampens some (off-diagonal) elements in $\bar{V}_{\theta_\nu}$, with the dampening factor removed as the sample size increases. Finally, all covariances that are further apart than $\bar{H}_T$ periods are set to zero. One popular approach—followed, for example, in Christiano et al. (2005)—is to set $\eta_{1,T} = \infty$ and $\zeta_T = 0$ (thus $\eta_{2,T}$ and $\bar{H}_T$ are immaterial), so that V_{θ_ν} is simply a diagonal matrix composed of the diagonal components of $\bar{V}_{\theta_\nu}$. The opposite extreme is to not dampen at all, setting $V_{\theta_\nu} = \bar{V}_{\theta_\nu}$.

In our applications we will follow an intermediate strategy. We set $\zeta_T = 1$ in order to treat autocorrelations and correlations across different variables equally; we furthermore use a triangular kernel, so $\eta_{1,T} = \eta_{2,T} = 1$, and a bandwidth of $\bar{H}_T = 8$.[25] We depart from the standard diagonal weighting matrix because of the model selection step: using a diagonal matrix would lead to artificially sharp model selection, since small differences in fit of different models will lead to starkly different posterior odds. Accounting for the correlation patterns present in the IRF estimates reflects the informativeness of the data more accurately.

B.3 Model estimation

We here provide further implementation details for the model estimation step. We proceed in two steps. First, for a given model $\mathcal{M}_j$ and parameter vector ψ_j, we explain how to obtain $\theta_\nu(\psi_j, \mathcal{M}_j)$. This step is non-standard; in particular, we explain why we need not specify a policy rule to do so. Second, we discuss how we draw from the posterior and estimate the marginal likelihood. That step is instead entirely standard, so we will be brief.

Obtaining $\theta_\nu(\psi_j, \mathcal{M}_j)$. In order to evaluate the likelihood, we first need to obtain $\theta_\nu(\psi_j, \mathcal{M}_j)$. We obtain it in the following way:

[25] We note that our results are robust to different choices of bandwidth or to the use of other kernels.

1. Given $(\psi_j, \mathcal{M}_j)$, solve for impulse responses of the targeted outcome variables to policy news shocks for all horizons, $\boldsymbol{\nu}$. To do so we close the model with some determinacy-inducing policy rule; as discussed in McKay and Wolf (2023), the choice of that baseline rule is immaterial. Denote the (truncated) impulse response function matrices of interest as $\Theta_{\boldsymbol{x}_m,\boldsymbol{\nu}}(\psi_j, \mathcal{M}_j)$ for variable $\boldsymbol{x}_m$. Stack all of those impulse response matrices vertically in the same order as for $\hat{\theta}_\nu$, and denote the stacked matrix as $\Theta_{\boldsymbol{\nu}}(\psi_j, \mathcal{M}_j)$. This is a $(n_m T) \times T$ matrix, where T is the truncation horizon.[26]

2. We then, for each of the n_ν empirically identified policy shocks, find the unique *vector* of policy shocks in the model that matches the empirical impulse response targets as well as possible. Formally, for each empirical target shock $n = 1, \dots, n_\nu$, define a $T \times 1$ vector of news shocks $\tilde{\boldsymbol{\nu}}_n$. Vertically stack all these vectors of policy news shocks in the $(n_\nu T) \times 1$ vector $\tilde{\nu} = [\tilde{\boldsymbol{\nu}}'_1, \dots, \tilde{\boldsymbol{\nu}}'_{n_\nu}]'$. Define also for convenience the following $(n_m T) \times (n_\nu T)$ matrix: $\boldsymbol{\Phi}(\psi_j, \mathcal{M}_j) = I_{n_\nu} \otimes \Theta_{\boldsymbol{\nu}}(\psi_j, \mathcal{M}_j)$ where I_{n_ν} is an n_ν-dimensional identity matrix. We then obtain the best-fit vector of news shocks $\tilde{\nu}^*$ as

$$\begin{aligned} \tilde{\nu}^*(\psi_j, \mathcal{M}_j) \quad &= \quad \underset{\tilde{\nu}}{\operatorname{argmax}} \;\; \tilde{p}(\hat{\theta}_\nu, \tilde{\theta}_\nu, V_{\theta_\nu}) \\ \text{s.t} \quad & \quad\;\; \tilde{\nu}_{H+1:T,n} = 0 \quad \text{for all } 1, \dots, n_\nu \\ & \quad\;\; \tilde{\theta}_\nu = \boldsymbol{\Phi}(\psi_j, \mathcal{M}_j)\tilde{\nu} \end{aligned}$$

where $\tilde{p}(\hat{\theta}_\nu, V_{\theta_\nu}, \tilde{\theta}_\nu)$ is the assumed density for "data" $\hat{\theta}_\nu$ with mean $\tilde{\theta}_\nu$ and covariance matrix V_{θ_ν}, and $\tilde{\nu}_{H+1:T,n}$ denotes elements $H+1, H+2, \dots, T$ of vector $\tilde{\boldsymbol{\nu}}_n$.[27] Given that f is assumed to be the density of a multivariate normal and V_{θ_ν} is taken as given, the maximizer $\tilde{\nu}^*(\psi_j, \mathcal{M}_j)$ can be found in closed form (since the maximization problem is a simple restricted linear quadratic problem).

In our empirical applications, we use only one identified monetary policy shock. To gain further intuition it is instructive to analyze this one-shock case in more detail. Take the top left $H \times H$ elements of each of the stacked impulse response matrices $\Theta_{\boldsymbol{x}_m,\boldsymbol{\nu}}(\psi_j, \mathcal{M}_j)$, stack them vertically, and denote the resulting matrix by $\Theta^H_{\boldsymbol{\nu}}(\psi_j, \mathcal{M}_j)$. Replacing the multivariate normal density, transforming appropriately, and focusing

[26] We set a truncation horizon of $T = 300$. Our results are insensitive to that choice.

[27] We impose this constraint to avoid overfitting: in order to match the IRF up to horizon H, we can only use the news shocks up to horizon H, and all other news shocks are set to zero.

only on the first H news shocks (since all others are set to zero) the problem to be solved can be written as:

$$\tilde{\nu}^*_{1:H}(\psi_j, \mathcal{M}_j) = \underset{\tilde{\nu}_{1:H}}{\text{argmax}} -\frac{1}{2}\left(\hat{\theta}_\nu - \Theta^H_{\boldsymbol{\nu}}(\psi_j, \mathcal{M}_j)\tilde{\nu}_{1:H}\right)' V^{-1}_{\theta_\nu}\left(\hat{\theta}_\nu - \Theta^H_{\boldsymbol{\nu}}(\psi_j, \mathcal{M}_j)\tilde{\nu}_{1:H}\right)$$

It is straightforward to show that the solution in this case is given by:

$$\tilde{\nu}^*_{1:H}(\psi_j, \mathcal{M}_j) = \left(\Theta^H_{\boldsymbol{\nu}}(\psi_j, \mathcal{M}_j)' V^{-1}_{\theta_\nu}\Theta^H_{\boldsymbol{\nu}}(\psi_j, \mathcal{M}_j)\right)^{-1}\left(\Theta^H_{\boldsymbol{\nu}}(\psi_j, \mathcal{M}_j)' V^{-1}_{\theta_\nu}\hat{\theta}_\nu\right)$$

In words, we can find the best-fitting shock vector $\tilde{\nu}^*_{1:H}$ through a "regression" of the empirical target impulse responses on the space of impulse response sequences implementable through policy.

3. With $\tilde{\nu}^*$ at hand, compute the model-implied impulse response functions as $\theta_\nu(\psi_j, \mathcal{M}_j) = \boldsymbol{\Phi}(\psi_j, \mathcal{M}_j)\tilde{\nu}^*$.

We note that this way of constructing the model-implied impulse responses $\theta_\nu(\psi_j, \mathcal{M}_j)$ differs from the standard approach of first (i) specifying a policy rule and then (ii) assuming that the identified policy shock corresponds to a time-0 shock under that rule (e.g., as in Christiano et al., 2005). For this approach to be valid, the assumed rule has to be correctly specified. In contrast, our approach does not require assumptions about the policy rule—we simply construct a sequence of contemporaneous and news policy shocks $\tilde{\nu}^*$ that perturbs the expected path of the policy instrument analogously to the empirically estimated policy instrument impulse response.[28]

Posterior distribution & marginal likelihood. We use a standard Random Walk Metropolis Hastings algorithm, with a multivariate normal for the proposal distribution. The variance-covariance matrix is initially assumed to be equal to the prior variance-covariance matrix, scaled by a constant c_1^2.[29] We use the first N_a draws to estimate the variance-covariance matrix of the proposal distribution, updating the proposal variance-covariance matrix to the observed variance-covariance matrix of parameters in the first N_a draws (scaled

[28] This claim works exactly in population as $T, H \to \infty$. However, due to the finite horizon of the impulse-response matching, the baseline assumed rule may matter due to truncation. In the models we consider, the matched impulse-response and inferred structural parameters are almost exactly the same under a variety of parameters for the assumed determinacy-inducing rule, consistent with the exact population result.

[29] For our HANK models, we in this step use a standard deviation of 0.1 for the informational stickiness parameter (instead of 0.2, see Table C.4), to avoid getting too many draws outside of the parameter support.

by c_2^2). Once updated, we sample another $N_b + N_c$ draws, burn the first N_b and keep the last N_c draws, which we use as our posterior distribution. We set $N_a = N_c = 100000$, $N_b = 50000$, $c_1 = 0.8$ and $c_2 = 0.7$ for all models. Our acceptance rates for all models considered range between 20 and 30 percent.

In order to then implement our applications, we need to store the impulse response matrices of the outcomes of interest with respect to the full sequence of news shocks. Given that storing hundreds of thousands of draws of $T \times T$ matrices is very expensive in terms of memory, we store only the $T_u \times T_u$ top left elements, for only a number of N_d draws. We set $T_u = 200$ and $N_d = 1000$. Specifically, we store one draw out of each $N_c/N_d = 100$, to get draws that are closer to uncorrelated. Finally, given those posterior draws, we estimate the marginal likelihood using the harmonic mean estimator of Geweke (1999).[30]

B.4 Vulnerabilities of the quantitative DSGE approach

We here provide some further details supplementing the discussion in Section 3.4, elaborating on some well-known vulnerabilities of the full-information quantitative DSGE approach.

Model mis-specification and inference. Under standard full-information approaches to model estimation (like, e.g., Smets and Wouters, 2007), mis-specification in one part of the model will affect inference for the other parts. The argument is straightforward, so our discussion here will be brief; we will furthermore focus our discussion on mis-specification in shock processes, as such mis-specification is particularly likely in practice (Chari et al., 2009). Analogous arguments apply to mis-specification in policy rules.

Suppose the true data-generating process is

$$y_t = \Theta^*(L)\xi_t \tag{B.3}$$

$$\xi_t = B^*(L)\varepsilon_t \tag{B.4}$$

where $\varepsilon_t \sim N(0, I)$. Relative to (1), the system (B.3) - (B.4) is written to explicitly separate the exogenous process (i.e., equation (B.4)) from the endogenous model propagation (i.e., equation (B.3)). For example, ε_t could be an innovation to total factor productivity, while ξ_t is the exogenous TFP level itself. For future reference we define $\Psi^*(L) = \Theta^*(L)B^*(L)$.

[30] We set the truncation parameter such that we use only half of the sample. We use the full sample consisting of N_c draws to estimate the marginal likelihoods.

The researcher instead entertains models indexed by parameters $\psi = (\psi_1', \psi_2')'$:

$$y_t = \Theta_{\psi_1}(L)\xi_t \tag{B.5}$$

$$\xi_t = B_{\psi_2}(L)\varepsilon_t \tag{B.6}$$

where again $\varepsilon_t \sim N(0, I)$. We assume that there is no mis-specification in the endogenous propagation part of the model: there is a (in fact unique) ψ_1^* such that $\Theta_{\psi_1^*}(L) = \Theta^*(L)$. Shock propagation, however, is mis-specified; for example, the researcher may assume that all shocks follow AR(1) processes, while in fact they follow richer ARMA(p,q) processes. For future reference we again write $\Psi_\psi(L) = \Theta_{\psi_1}(L)B_{\psi_2}(L)$.

Finally, to make our arguments as stark as possible, we suppose that there exists a unique $\psi^\dagger$ such that

$$\Psi^*(e^{-i\omega})\Psi^*(e^{-i\omega})' = \Psi_{\psi^\dagger}(e^{-i\omega})\Psi_{\psi^\dagger}(e^{-i\omega})' \qquad \forall \omega \in [0, 2\pi].$$

Thus, when evaluated at $\psi^\dagger$ (and only then), the two processes (B.3) - (B.4) and (B.5) - (B.6) imply the exact same second moments, so conventional likelihood-based estimation will asymptotically yield $\psi = \psi^\dagger$. But since $B^*(L) \neq B_{\psi_2^\dagger}(L)$, we will generically have $\Theta^*(L) \neq \Theta_{\psi_1^\dagger}(L)$—i.e., mis-specification in the endogenous shock propagation part, including in particular the policy space Θ_ν. Since our approach does not require the researcher to take any stance on the shock process part $B(L)$, it is by design robust to such concerns.

A concrete illustration of this abstract discussion is provided by the model of Smets and Wouters. In that model, the exogenous shocks driving inflation already induce hump shapes (they follow ARMA(1,1)'s), and so other shocks—like monetary shocks—induce much weaker hump shapes than observed in the data; we thank Simon Gilchrist for making this point.

Weak identification. Standard full-information approaches to estimation of DSGE models are also often subject to concerns of weak identification (e.g., see Fernández-Villaverde et al., 2016). Our approach is arguably less subject to this concern, simply because it only requires the researcher to partially specify the model, thus reducing the number of parameters that need to be identified. We here provide a simple example illustration of this insight.

Consider the following two-variable, two-equation static model:

$$\begin{aligned} y_t &= -\frac{1}{\gamma} i_t + \sigma_d \varepsilon_t^d, \\ i_t &= \phi_y y_t + \sigma_m \varepsilon_t^m, \end{aligned}$$

where y_t and i_t denote outcome variables (output and interest rates), and $(\varepsilon_t^d, \varepsilon_t^m)$ are shocks. Note that the solution is given as

$$\begin{pmatrix} y_t \\ i_t \end{pmatrix} = \underbrace{\frac{1}{1+\frac{\phi_y}{\gamma}} \begin{pmatrix} -\frac{1}{\gamma}\sigma_m & \sigma_d \\ \sigma_m & \phi_y \sigma_d \end{pmatrix}}_{\equiv \Theta} \begin{pmatrix} \varepsilon_t^m \\ \varepsilon_t^d \end{pmatrix}$$

Consider first a researcher following our approach. The ratio of the impulse responses of interest rates and output to a monetary policy shock ε_t^m point-identifies γ, and so the space of output and interest rate allocations implementable through policy, as required by our identification result. Now consider instead identification based on second moments; i.e., we seek to find a tuple $\{\gamma, \phi_y, \sigma_d, \sigma_m\}$ such that

$$\Sigma = \Theta(\gamma, \phi_y, \sigma_d, \sigma_m)\Theta(\gamma, \phi_y, \sigma_d, \sigma_m)'$$

where $\Sigma \equiv \Theta\Theta'$ is the true variance-covariance matrix. It is straightforward to verify that these moment conditions are insufficient to point-identify the model, and in particular do not point-identify γ.[31]

[31] To see this, start with some arbitrary $\gamma > 0$. Note that

$$\frac{\text{Var}(i_t) + \gamma\,\text{Cov}(y_t, i_t)}{\text{Var}(y_t) + \frac{1}{\gamma}\,\text{Cov}(y_t, i_t)} = \frac{\phi_y^2 + \phi_y \gamma}{1 + \frac{\phi_y}{\gamma}}$$

Solve this equation for ϕ_y, recover σ_d from $\text{Var}(i_t) + \gamma\,\text{Cov}(y_t, i_t)$, and finally get σ_m from $\text{Var}(i_t)$. The resulting parameter vector leads the model to correctly match the desired Σ.

C Monetary policy "VAR-Plus" inputs

This appendix complements the discussion in Section 4 on the "VAR-Plus" inputs for our applications. First, in Appendix C.1, we begin with the reduced-form projections. Second, in Appendix C.2, we then present additional details on our empirical monetary shock estimation, completing the VAR step. Third, in Appendices C.3 and C.4, we provide the detailed equations for our list of models $\mathcal{M}$, and discuss their estimation.

C.1 Reduced-form projections

We provide supplementary details on how we construct our reduced-form projections. We elaborate on data construction and econometric implementation, and also compare the implied forecasts with other approaches.

DATA. We consider the same ten observables y_t as in Angeletos et al. (2020). The series are constructed as follows. Unless indicated otherwise, each series is transformed to stationarity following Hamilton (2018), and series names refer to FRED mnemonics.

- *Unemployment rate.* We take the series `UNRATE` from FRED. We do not transform this series further.
- *Output gap.* We take log output per capita from FRED (`A939RX0Q048SBEA`). We interpret the stationarity-transformed series as a measure of the output gap.
- *Investment.* We compute log investment per capita, where investment is defined as the sum of durables and gross private domestic investment. We construct this series as (`PCDG`+`GPDI`)*`A939RX0Q048SBEA/GDP`.
- *Consumption.* We compute log consumption per capita, where consumption is defined as the sum of nondurables and services. We construct this series as (`PCND`+`PCESV`) * `A939RX0Q048SBEA/GDP`.
- *Hours.* We compute log hours worked, where total hours worked are constructed as `PRS85006023` * `CE16OV/CNP16OV`.
- *Utilization-adjusted TFP.* We compute the cumulative sum of the series `DTFPu`, from John Fernald's webpage (`https://www.johnfernald.net/TFP2023.03.07revision`).

- *Labor productivity.* We compute log labor productivity, where labor productivity is obtained as `OPHNFB`.

- *Labor share.* We compute the log labor share, with `PRS85006173` as the labor share.

- *Inflation.* We compute the log-differenced GDP deflator (`GDPDEF`), and then annualize, without further transformations.

- *Federal funds rate.* We obtain the series `FEDFUNDS`, without further transformations.

All series are quarterly. For the applications in Sections 5.2 and 5.3, we consider samples from 1960:Q1—2019:Q4. For the covid inflation counterfactual in Section 5.4, we extend the sample to 2021:Q2, the contemplated forecasting date.

Econometric implementation. We restrict attention to OLS point estimates. We always include a constant and a linear time trend. For the second-moment counterfactual in Section 5.2 we include four lags, to allow for an accurate fit of second moments. For the forecast-based counterfactuals in Sections 5.3 and 5.4, we include two lags.

Comparison with alternative forecasts. We now perform two additional checks to demonstrate the good forecasting performance of our reduced-form VAR: we (i) check that the forecast accuracy is similar to that of the Survey of Professional Forecasters (SPF); and (ii) show that our 10-variable system contains nearly all of the information in the eight business cycle factors that Stock and Watson (2016) computed from a large set of macroeconomic and financial variables.

1. *Comparison with the SPF.* We assess forecast accuracy starting in 1981:Q3 (when the SPF forecast for the T-Bill rate becomes available) and ending in 2007:Q3 (before the onset of the Great Recession and the ZLB period).[32] Table C.1 shows the mean squared errors of the one- and four-quarter-ahead forecasts from our VAR and from the SPF, for our three main series of interest. Our VAR evidently performs well by this metric.

2. *Information content of the Stock and Watson factors.* Stock and Watson (2016) estimate 8 factors that drive the bulk of the variation in a database of 207 quarterly time series on

[32] In order to allow an apples-to-apples comparison with the SPF, we need to slightly modify our VAR. Specifically, we use raw GDP data (not per capita) and the T-Bill rate in place of the federal funds rate, as those are the variables that appear in the SPF. As in the baseline VAR analysis, we detrend all non-stationary series, but to compare to the SPF we add the trends back to the VAR-implied forecasts.

Variable	1-quarter ahead		4-quarter ahead	
	VAR	**SPF**	**VAR**	**SPF**
GDP	0.569	0.327	2.76	2.97
Inflation	0.357	0.934	0.646	1.85
T-Bill rate	0.461	0.740	1.66	3.17

Table C.1: Mean squared error of 1- and 4-quarter ahead forecasts.

	1-quarter ahead		4-quarter ahead	
	w/o f	w/ f	w/o f	w/ f
Output gap	0.918	0.935	0.648	0.774
Inflation	0.826	0.838	0.717	0.732
Interest rate	0.945	0.953	0.759	0.781

Table C.2: Assessing the incremental information content of the Stock-Watson factors: forecasting R^2 with and without inclusion of factors in the VAR.

the U.S. macro-economy and financial markets. We now ask whether adding these factors to the information set of our VAR would lead to a substantial improvement in forecasting performance. Specifically, let the variables in our VAR be represented by the vector y_t and the 8 factors be represented by the vector f_t. For horizon $h \in \{1, 4\}$, we consider a regression of the form

$$y_{t+h} = B_0 y_t + B_1 y_{t-1} + B_f f_t,$$

and then assess the implications of setting $B_f = 0$. Table C.2 shows the results for our core observables. We see that, with the exception of the 4-quarter-ahead forecast of the output gap, the increase in R^2 from including the factors is quite small. We thus judge that including the Stock-Watson factors would not lead to materially different forecasts.

C.2 Empirical evidence on monetary shock propagation

We provide further details on how we construct our monetary shock estimation targets in Section 4.2. We elaborate on data construction and econometric implementation.

Data. We are interested in impulse responses of three outcome variables: the output gap, inflation, and the policy rate. All series are constructed as for the first part of the VAR step, see Appendix C.1. Our measure of a monetary shock series is obtained from the replication files of Aruoba and Drechsel (2022). We aggregate by averaging the monthly series, and we set all missing values of this monetary shock IV to zero, as in Känzig (2021).

Econometric implementation. We estimate a VAR in the shock series together with our three outcome variables of interest, consistent with the recommendations of Li et al. (2023). As in Plagborg-Møller and Wolf (2021), we order the shock series first in a recursive identification of our VAR, delivering invertibility-robust estimates of the causal effects of the monetary shock. We include two lags, a linear time trend, and use a uniform-normal-inverse-Wishart distribution over the orthogonal reduced-form parameterization (Arias et al., 2018). Our estimation results are robust to these particular choices. This procedure yields draws of the policy shock causal effect vector $\hat{\theta}_\nu$, which are then used to construct V_{θ_ν} following the steps outlined in Appendix B.2.

C.3 Models of monetary policy transmission

This section provides some supplementary details for our structural models of monetary policy transmission sketched in Section 4.3. We list all model equations; however, since the models are relatively standard, the derivations will be rather brief. Throughout this section, we use tildes to denote log-deviations from steady state.

C.3.1 Baseline RANK

Households & unions. Households choose sequences of consumption c_t and assets a_t^H to maximize lifetime utility, given by

$$\mathbb{E}_0 \left[\sum_{t=0}^{\infty} \beta^t \left[u\left(c_t - hc_{t-1}\right) - v\left(\ell_t\right) \right] \right], \tag{C.1}$$

subject to a standard no-Ponzi condition as well as the budget constraint

$$c_t + a_t^H = w_t(1-\tau_t^\ell)\ell_t + d_t^H - \tau_t + \frac{1+r_{t-1}^n}{1+\pi_t} a_{t-1}^H, \tag{C.2}$$

where w_t is the real wage, τ_t^ℓ is the labor tax rate, d_t^H is real dividend income, τ_t is a transfer, r_t^n is the nominal interest rate, and π_t is the price inflation rate. We assume that $u(x) = \frac{x^{1-\gamma}}{1-\gamma}$

and $v(x) = \nu \frac{x^{1+\varphi}}{1+\varphi}$. The Euler Equation in log-deviations from steady state is:

$$\tilde{\lambda}_t = \mathbb{E}_t[\tilde{r}_{t+1} + \tilde{\lambda}_{t+1}]$$

with $\tilde{r}_{t+1} = \tilde{r}^n_t - \pi_{t+1}$, $\frac{P_{t+1}}{P_t} = \exp(\pi_{t+1})$, and

$$\tilde{\lambda}_t = -\frac{1}{(1-\beta h)(1-h)}\gamma(\tilde{c}_t - h\tilde{c}_{t-1}) + \frac{1}{(1-\beta h)(1-h)}\beta h\gamma(\mathbb{E}_t\left[\tilde{c}_{t+1}\right] - h\tilde{c}_t).$$

A detailed derivation of the wage Phillips curve—which summarizes the labor supply block—is deferred until Appendix C.3.3, given that the full information case is nested in the derivation that includes cognitive discounting.

Production and pricing. The production function for an intermediate good producer i is:

$$Y_t(i) = \bar{A}(u_t(i)k_{t-1}(i))^{\alpha}(\ell_t(i))^{1-\alpha}$$

where $\bar{A}$ denotes aggregate productivity, $k_{t-1}(i)$ is capital stock of firm i, $u_t(i)$ is capacity utilization, and $\ell_t(i)$ denotes labor hired. All intermediate good producers are symmetric and so we drop the i subscript. Capital is purchased one period in advance. The intermediate good producer solves:[33]

$$\max_{\ell_t, k_t, u_t} \mathbb{E}_0\left[\sum_{t=0}^{\infty}(\Pi_{j=0}^{t}(1+r_j))^{-1}[p^I_t Y_t - w_t\ell_t - a(u_t) - q_t(k_t - (1-\delta)k_{t-1})\right]$$

where $a(u_t)$ is an utility cost of adjusting capacity, and $q_t k_t$ is the total cost of capital purchases for next period.[34] The first-order conditions are:

$$\begin{aligned}
w_t &= p^I_t(1-\alpha)\bar{A}\left(\frac{\ell_t}{u_t k_{t-1}}\right)^{-\alpha} \\
a'(u_t) &= p^I_t\alpha\bar{A}\left(\frac{\ell_t}{u_t k_{t-1}}\right)^{1-\alpha} \\
q_t &= \mathbb{E}_t\left(\frac{1}{1+r_{t+1}}\left[p^I_{t+1}\alpha\bar{A}\left(\frac{\ell_{t+1}}{u_{t+1}k_t}\right)^{1-\alpha} + (1-\delta)q_{t+1}\right]\right)
\end{aligned}$$

[33]We discount future pay-offs using the real rate of interest. Up to first order, this is equivalent to using the representative household's implied stochastic discount factor.

[34]The cost is written in terms of utility, so it does not enter the market-clearing condition.

Log-linearizing around the steady state:

$$\begin{aligned}
\tilde{y}_t &= \alpha(\tilde{u}_t + \tilde{k}_{t-1}) + (1-\alpha)\tilde{\ell}_t \\
\tilde{w}_t &= \tilde{p}_t^I + \alpha(\tilde{u}_t + \tilde{k}_{t-1}) - \alpha\tilde{\ell}_t \\
\zeta\tilde{u}_t &= \tilde{p}_t^I + (\alpha-1)(\tilde{u}_t + \tilde{k}_{t-1}) + (1-\alpha)\tilde{\ell}_t \\
\tilde{q}_t &= \mathbb{E}_t\left[-\tilde{r}_{t+1} + \left(1 - \frac{1-\delta}{1+\bar{r}}\right)(\tilde{p}_{t+1}^I + (\alpha-1)(\tilde{k}_t + \tilde{u}_{t+1}) + (1-\alpha)\tilde{\ell}_{t+1}) + \frac{1-\delta}{1+\bar{r}}\tilde{q}_{t+1}\right]
\end{aligned}$$

where $\zeta = a''(1)/a'(1)$ is the curvature parameter of the capacity utilization cost function. Following Smets and Wouters (2007), we parametrize $\zeta = \frac{\psi}{1-\psi}$ and then use the same prior on ψ as in that paper.

Retail firms solve their dynamic pricing problem subject to Calvo frictions. Detailed derivations are deferred until Appendix C.3.3.

Capital good producers solve

$$\max_{i_t} \mathbb{E}_0\left[\sum_{t=0}^{\infty}(\Pi_{j=0}^t(1+r_j))^{-1}\left(q_t i_t - S\left(\frac{i_t}{i_{t-1}}\right)\right)\right],$$

where i_t is the production of new capital goods (sold to the intermediate goods producers), and $S(x)$ is the adjustment cost function. The first-order condition is given by:

$$q_t = \frac{1}{i_{t-1}}S'\left(\frac{i_t}{i_{t-1}}\right) - \mathbb{E}_t\left[\frac{1}{1+r_{t+1}}S'\left(\frac{i_{t+1}}{i_t}\right)\frac{i_{t+1}}{i_t^2}\right]$$

We assume that $S(1) = S'(1) = 0$ and $\kappa = S''(1) > 0$. Log-linearizing around the steady state yields:

$$q_t = \kappa(\tilde{i}_t - \tilde{i}_{t-1}) - \frac{\kappa}{1+\bar{r}}(\tilde{i}_{t+1} - \tilde{i}_t)$$

Finally, capital evolves according to $k_t = (1-\delta)k_{t-1} + i_t$ or in log-linearized terms:

$$\tilde{k}_t = (1-\delta)\tilde{k}_{t-1} + \delta\tilde{i}_t$$

We note that therefore goods market-clearing implies that, to first order:

$$\tilde{y}_t = \bar{c}\tilde{c}_t + \bar{i}\tilde{i}_t$$

Policy. The government budget constraint is

$$w_t \ell_t \tau_t^\ell + b_t = (1 + r_t) b_{t-1} + \tau_t + g_t$$

where r_t is the real return on government debt b_t, τ_t denotes lump-sum transfers, τ_t^ℓ denotes distortionary labor taxes, and g_t denotes government expenditure. Log-linearizing:

$$\bar{w}\bar{\ell}\bar{\tau}^\ell(\tilde{w}_t + \tilde{\ell}_t + \tilde{\tau}_t^\ell) + \bar{b}\tilde{b}_t = (1 + \bar{r})\bar{b}(\tilde{b}_{t-1} + \tilde{r}_t) + \bar{\tau}\tilde{\tau}_t + \bar{g}\tilde{g}_t$$

Second, the realized real return on government debt satisfies

$$1 + r_t = \frac{\bar{r} + \eta}{\exp(\pi_t)} \frac{1}{p_{t-1}} + \frac{1 - \eta}{\exp(\pi_t)} \frac{p_t}{p_{t-1}}$$

where p_t is the real relative price of government debt and η is the decay rate of the coupon, with $\eta = 0$ corresponding to perpetuities and $\eta = 1$ corresponding to one-period debt. Log-linearizing:

$$\tilde{r}_t = -\pi_t - \tilde{p}_{t-1} + \frac{1 - \delta}{1 + \bar{r}} \tilde{p}_t$$

The central bank sets the nominal rate on one-period government debt, which is in zero net supply. By perfect foresight arbitrage we have

$$1 + r_t = \frac{1 + r_{t-1}^n}{\exp(\pi_t)}, \quad t = 1, 2, \ldots$$

and so, in log-deviations

$$\tilde{r}_t = \tilde{r}_{t-1}^n - \pi_t, \quad t = 1, 2, \ldots$$

or

$$\tilde{r}_t^n = -\tilde{p}_t + \frac{1 - \eta}{1 + \bar{r}} \mathbb{E}_t \left[\tilde{p}_{t+1}\right]$$

It remains to determine how taxes are set. We assume:

$$\begin{aligned} \tilde{\tau}_t &= \tilde{g}_t = 0 \\ \bar{w}\bar{\ell}\tilde{\tau}_t^\ell &= \bar{b}\tau_b^\ell \tilde{b}_{t-1} \end{aligned}$$

That is, all the adjustment is done via distortionary taxes. The resulting law of motion for government debt is

$$\tilde{b}_t = (1 + \bar{r} - \bar{\tau}^\ell \tau_b^\ell)\tilde{b}_{t-1} + (1 + \bar{r})\tilde{r}_t.$$

Policy rule for computation. For our numerical analysis, we close the model with a determinacy-inducing Taylor rule, as discussed in Appendix B.3:

$$\tilde{r}_t^n = (1-\rho)\left(\rho\tilde{r}_{t-1}^n + \phi_\pi\pi_t + \phi_y\tilde{y}_t + \phi_{\Delta y}(\tilde{y}_t - \tilde{y}_{t-1})\right)$$

As emphasized throughout, our model estimation step and policy counterfactual applications do not depend on this choice of basis rule, simply because we allow for arbitrarily general policy shocks, allowing us to implement arbitrary paths of interest rates. For Figures 3 and 4, we subject this rule to ten-quarter-ahead forward guidance shocks.

Steady state. We normalize the level of disutility of labor such that $\bar{\ell} = 1$. Given that assumption, the Euler equation pins down the real rate as $1 + \bar{r} = \beta^{-1}$. We can then find $\bar{k}$, which immediately yields $\bar{i}$, $\bar{y}$ and $\bar{w}$. We calibrate the level of outstanding government debt, labor taxes and transfers (see Appendix C.4), and pick the steady state level of government consumption such that the intertemporal government budget constraint holds.

C.3.2 Baseline HANK

The only two differences relative to the baseline RANK model are that: (i) we replace the representative agent with a heterogeneous agents block, as already described in the main text; (ii) we now need to specify how dividends are paid to the households.

Household and unions. Households are subject to idiosyncratic income risk (with the risk process taken from Kaplan et al., 2018), and hours worked are intermediated by labor unions, as in the baseline representative-agent model.[35] Households save in government bonds, while firm capital and equity is held by financial intermediaries; those intermediaries gradually pay out dividends to households in proportion to their productivity. Letting $1 - \theta$ denote the probability that a household updates its information about aggregate conditions, and letting s denote the number of periods since the last update, the consumption-savings problem can be stated recursively as

$$V_t(a, e, s) = \max_{c,a'} \left\{u(c) - v(\ell_t) + \beta\mathbb{E}_{t-s}\left[\theta V_{t+1}(a', e', s+1) + (1-\theta)V_{t+1}(a', e', 0)\right]\right\}$$

[35] We assume that unions evaluate the marginal utility of income using $c^{-\gamma}$ where c is aggregate consumption. As the Phillips curves are then unchanged, this assumption limits the effects of inequality to the demand side of the model (as in McKay and Wolf, 2022).

subject to the budget constraint

$$c + a' = \left((1 - \tau_{\ell,t}) w_t \ell_t + d_t^H\right) e + \frac{1 + r_{t-1}^n}{1 + \pi_t} a + \tau_t$$

and the borrowing constraint $a' \geq \underline{a}$, and where e denotes idiosyncratic household productivity. The borrowing constraint $\underline{a}$ is set as in Kaplan et al. (2018). In order to compute the solution with informational rigidities, we follow Auclert et al. (2020): we first solve for the Jacobians of the household block under full information, and then transform them to obtain the solution under sticky information.

Dividend distribution. Households receive dividends through a financial intermediary. Let a_t^I denote total assets held by the financial intermediary. Those assets evolve as

$$a_t^I = (1 + r_t) a_{t-1}^I + (d_t - d_t^H)$$

where d_t denotes dividends paid by firms to the intermediary and d_t^H denotes payments from the intermediary to the households. We assume the following distribution rule:

$$(d_t^H - \bar{d}) = \delta_1 (d_t - \bar{d}) + \delta_2 (1 + r_t) a_{t-1}^I$$

Note that $\delta_1 = 1$ corresponds to the usual case of dividends paid out straight to households, with $a_t^I = 0$ always. The linearized relations are

$$\hat{a}_t^I = (1 - \delta_2)(1 + \bar{r}) \hat{a}_{t-1}^I + (1 - \delta_1) \bar{d} \tilde{d}_t$$

and

$$\bar{d} \tilde{d}_t^H = \delta_1 \bar{d} \tilde{d}_t + \delta_2 (1 + \bar{r}) \tilde{a}_{t-1}^I$$

where $\hat{x} = x - \bar{x}$. We linearize (instead of log-linearizing) with respect to a_t^I since $\bar{a}^I = 0$.

Steady state. We proceed exactly as in the RANK case. Given a calibrated real interest rate, we pick β such that in equilibrium households want to hold the calibrated level of liquid assets, which are given by the outstanding stock of government debt. Apart from the value of β, the steady state is exactly the same as in the RANK case.

C.3.3 Adding cognitive discounting

This subsection derives the price- and wage-NKPCs under cognitive discounting and price indexation. We derive the NPKCs under partial indexation and cognitive discounting, where ζ and ζ_w are the degrees of price indexation; $\zeta = \zeta_w = 1$ corresponds to the case considered in our main analysis.

Pricing. The problem of a retailer is to choose P_t^* to maximize

$$\mathbb{E}_t \sum_{\tau \geq t}^{\infty} (\bar{\beta}\theta_p)^{\tau-t} M_{\tau|t} \left(\frac{P_{\tau|t}}{P_\tau} - \mu_\tau \right) \left(\frac{P_{\tau|t}}{P_\tau} \right)^{-\epsilon_p} Y_\tau,$$

where $\bar{\beta} = \frac{1}{1+\bar{r}}$, $P_{\tau|t}$ is the price at date τ of a firm that last updated its price at t, μ_τ is the real marginal cost of producing at τ, P_τ is the aggregate price index, Y_τ is aggregate demand, $M_{\tau|t} = u_c(c_\tau)/u_c(c_t)$, and $1 - \theta_p$ is the probability of resetting the price. Due to price indexation, we have

$$P_{\tau|t} = P_t^* \underbrace{\exp\left(\zeta(\pi_t + \pi_{t+1} + \cdots \pi_{\tau-1})\right)}_{\equiv I_{\tau|t}}.$$

The first-order condition of the price-setting problem is

$$(\epsilon_p - 1)\mathbb{E}_t \sum_{\tau \geq t}^{\infty} (\bar{\beta}\theta_p)^{\tau-t} M_{\tau|t} \left(\frac{P_{\tau|t}}{P_\tau} \right)^{-\epsilon_p} Y_\tau \frac{I_{\tau|t}}{P_\tau} = \epsilon_p \mathbb{E}_t \sum_{\tau \geq t}^{\infty} (\bar{\beta}\theta_p)^{\tau-t} M_{\tau|t}\mu_\tau \left(\frac{P_{\tau|t}}{P_\tau} \right)^{-\epsilon_p - 1} Y_\tau \frac{I_{\tau|t}}{P_\tau}.$$

Log-linearizing both sides of this equation around a zero-inflation steady state we have

$$\mathbb{E}_t \sum_{\tau \geq t}^{\infty} (\bar{\beta}\theta_p)^{\tau-t} \left[\tilde{\mu}_\tau - \tilde{P}_{\tau|t} + \tilde{P}_\tau \right] = 0$$

or

$$\mathbb{E}_t \sum_{\tau \geq t}^{\infty} (\bar{\beta}\theta_p)^{\tau-t} \left[\tilde{\mu}_\tau - \tilde{P}_t^* - \sum_{s=t+1}^{\tau} \zeta\pi_{s-1} + \tilde{P}_\tau \right] = 0$$

and so

$$\tilde{P}_t^* - \tilde{P}_t = (1 - \bar{\beta}\theta_p)\mathbb{E}_t \sum_{\tau \geq t}^{\infty} (\bar{\beta}\theta_p)^{\tau-t} \left[\tilde{\mu}_\tau - \sum_{s=t+1}^{\tau} \zeta\pi_{s-1} + \tilde{P}_\tau - \tilde{P}_t \right]$$

or

$$\tilde{P}_t^* - \tilde{P}_t = (1-\bar{\beta}\theta_p)\mathbb{E}_t \sum_{\tau \geq t}^{\infty} (\bar{\beta}\theta_p)^{\tau - t} \left[\tilde{\mu}_\tau + \sum_{s=t+1}^{\tau} (1-\zeta L)\pi_s \right]$$

where L is the lag operator. We now apply cognitive discounting (as in Gabaix, 2020):

$$\tilde{P}_t^* - \tilde{P}_t = (1-\bar{\beta}\theta_p) \sum_{\tau \geq t}^{\infty} (\bar{\beta}\theta_p m)^{\tau - t} \left[\tilde{\mu}_\tau + \mathbb{E}_t \sum_{s=t+1}^{\tau} (1-\zeta L)\pi_s \right] \tag{C.3}$$

where m is the cognitive discount factor.

The aggregate price index evolves as

$$P_t = \left[\theta_p (P_{t-1}(\exp(\zeta \pi_{t-1})))^{1-\varepsilon} + (1-\theta_p)(P_t^*)^{1-\varepsilon} \right]^{1/(1-\varepsilon)}$$

Solving this for P_t^*:

$$P_t^* = \left[\frac{P_t^{1-\varepsilon} - \theta_p (P_{t-1}(\exp(\zeta \pi_{t-1}))^{1-\varepsilon}}{1-\theta_p} \right]^{1/(1-\varepsilon)}$$

Dividing by P_t:

$$\frac{P_t^*}{P_t} = \left[\frac{1 - \theta_p (\exp(\pi_t))^{\varepsilon - 1} (\exp(\zeta \pi_{t-1}))^{1-\varepsilon}}{1-\theta_p} \right]^{1/(1-\varepsilon)}$$

Re-arranging:

$$(1-\theta_p) \left(\frac{P_t^*}{P_t} \right)^{1-\varepsilon} = 1 - \theta_p (\exp(\pi_t))^{\varepsilon - 1} (\exp(\zeta \pi_{t-1}))^{1-\varepsilon}$$

Log-linearizing:

$$\begin{aligned} \pi_t &= \frac{1-\theta_p}{\theta_p} \left(\tilde{P}_t^* - \tilde{P}_t \right) + \zeta \pi_{t-1} \\ (1-\zeta L)\pi_t &= \frac{1-\theta_p}{\theta_p} \left(\tilde{P}_t^* - \tilde{P}_t \right) \end{aligned} \tag{C.4}$$

Combining (C.3) and (C.4) we arrive at

$$(1-\zeta L)\pi_t = \frac{(1-\theta_p)(1-\bar{\beta}\theta_p)}{\theta_p} \mathbb{E}_t \sum_{\tau \geq t}^{\infty} (\bar{\beta}\theta_p m)^{\tau - t} \left[\tilde{\mu}_\tau + \sum_{s=t+1}^{\tau} (1-\zeta L)\pi_s \right]. \tag{C.5}$$

Define $\breve{\pi}_t = (1-\zeta L)\pi_t$ as the quasi-differenced rate of inflation. We can then rewrite the preceding equation as

$$\check{\pi}_t = \frac{(1-\theta_p)(1-\bar{\beta}\theta_p)}{\theta_p}\mathbb{E}_t\left[\sum_{\tau=t}^{\infty}(\bar{\beta}\theta_p m)^{\tau-t}\tilde{\mu}_\tau + \frac{1}{1-\bar{\beta}\theta_p m}\sum_{\tau=t+1}^{\infty}(\bar{\beta}\theta_p m)^{\tau-t}\check{\pi}_\tau\right]$$

or

$$\check{\pi}_t = \underbrace{\frac{(1-\theta_p)(1-\bar{\beta}\theta_p)}{\theta_p}}_{\kappa_p}\mathbb{E}_t\left[\sum_{\tau=t}^{\infty}(\bar{\beta}\theta_p m)^{\tau-t}\left(\tilde{\mu}_\tau + \frac{\check{\pi}_\tau}{1-\bar{\beta}\theta_p m}\right) - \frac{\check{\pi}_t}{1-\bar{\beta}\theta_p m}\right]$$

and so

$$\check{\pi}_t\left[1+\frac{\kappa_p}{1-\bar{\beta}\theta_p m}\right] = \kappa_p\left[\sum_{\tau=t}^{\infty}(\bar{\beta}\theta_p m)^{\tau-t}\left(\tilde{\mu}_\tau + \frac{\check{\pi}_\tau}{1-\bar{\beta}\theta_p m}\right)\right]$$

Differencing forward and re-arranging:

$$\left[1+\frac{\kappa_p}{1-\bar{\beta}\theta_p m}\right]\left(\check{\pi}_t - \bar{\beta}\theta_p m\mathbb{E}_t\check{\pi}_{t+1}\right) = \kappa_p\left(\tilde{\mu}_t + \frac{\check{\pi}_t}{1-\bar{\beta}\theta_p m}\right)$$

and so

$$\check{\pi}_t = \kappa_p\tilde{\mu}_t + \bar{\beta}\theta_p m\left[1+\frac{\kappa_p}{1-\bar{\beta}\theta_p m}\right]\mathbb{E}_t\check{\pi}_{t+1}$$

Replacing the definition of $\tilde{\pi}_t$ and noting that $\tilde{\mu}_t = \tilde{p}_t^I$ yields the price-NKPC:

$$\pi_t - \pi_{t-1} = \kappa_p p_t^I + \beta^p\mathbb{E}_t\left[\pi_{t+1} - \pi_t\right] \tag{C.6}$$

Wage-setting. For tractability we assume that unions evaluate household utility at average consumption and hours worked (rather than averaging across individual household utilities), as in McKay and Wolf (2022). When a union does not update its wage, it adjusts it to $W_{j,t} = W_{j,t-1}(\exp(\zeta_w\pi_{t-1}))$, where π_t is price inflation. We will use the notation

$$W_{\tau|t} \equiv W_t^* \exp(\zeta_w(\pi_t + \cdots + \pi_{\tau-1})$$

for the nominal wage at date τ for a union that set its wage at date t. As before we derive everything allowing for partial indexation, with our analysis in the main text corresponding to the special case of full indexation ($\zeta_w = 1$). Real earnings for union j are

$$\frac{W_{\tau|t}}{P_\tau}\ell_{j\tau} = \left(\frac{W_{\tau|t}}{P_\tau}\right)\left(\frac{W_{\tau|t}}{W_\tau}\right)^{-\epsilon_w} L_\tau = \left(\frac{W_\tau}{P_\tau}\right)\left(\frac{W_{\tau|t}}{W_\tau}\right)^{1-\epsilon_w} L_\tau.$$

Note that $\ell_{j,t}$ denotes hours worked for union j, ℓ_τ is total hours worked by the households, and L_τ is the effective aggregate labor supply. Wage dispersion implies $L_\tau \leq \ell_\tau$; however, since we consider first-order approximations, we can proceed as if $L_\tau = \ell_\tau$.

The union's problem is to choose the nominal reset wage W_t^* to maximize

$$\mathbb{E}_t \sum_{\tau \geq t} \left(\bar{\beta}\theta_w\right)^{\tau-t} \left[\lambda_t \left(\frac{W_\tau}{P_\tau}\right) \left(\frac{W_{\tau|t}}{W_\tau}\right)^{1-\epsilon_w} - \nu_\ell\left(\ell_\tau\right) \left(\frac{W_{\tau|t}}{W_\tau}\right)^{-\epsilon_w}\right] L_\tau$$

where λ_t is the relevant aggregate marginal utility, and $\bar{\beta}$ is the time discount factor used by the union, assumed to equal the one used by the firm.[36]

The first-order condition is

$$\mathbb{E}_t \sum_{\tau \geq t} \left(\bar{\beta}\theta_w\right)^{\tau-t} \nu_\ell\left(\ell_\tau\right) \ell_\tau \epsilon_w W_\tau^{\epsilon_w} \prod_{s=t+1}^{\tau} \exp(\zeta_w \pi_{s-1})$$
$$= \mathbb{E}_t \sum_{\tau \geq t} \left(\bar{\beta}\theta_w\right)^{\tau-t} u_c(c_\tau)(\epsilon_w - 1) \frac{W_{\tau|t}}{P_\tau} W_\tau^{\epsilon_w} \ell_\tau \prod_{s=t+1}^{\tau} \exp(\zeta_w \pi_{s-1}).$$

Log-linearizing the first-order condition around a zero-inflation steady state:

$$\mathbb{E}_t \sum_{\tau=t}^{\infty} \left(\beta\theta_w\right)^{\tau-t} \left(\phi\tilde{\ell}_\tau - \tilde{W}_{\tau|t} + \tilde{p}_\tau - \tilde{\lambda}_\tau\right) = 0$$

or

$$\mathbb{E}_t \sum_{\tau=t}^{\infty} \left(\beta\theta_w\right)^{\tau-t} \left(\phi\tilde{\ell}_\tau - \tilde{W}_t^* - \sum_{s=t+1}^{\tau} \zeta_w \pi_{s-1} + \tilde{p}_\tau - \tilde{\lambda}_\tau\right) = 0,$$

where $\phi \equiv \frac{\nu_{\ell\ell}(\bar{\ell})\bar{\ell}}{\nu_\ell(\bar{\ell})}$. Re-arranging

$$\tilde{W}_t^* - \tilde{W}_t = (1 - \bar{\beta}\theta_w)\mathbb{E}_t \sum_{\tau \geq t} \left(\bar{\beta}\theta_w\right)^{\tau-t} \left(\phi\tilde{\ell}_\tau - \tilde{\lambda}_\tau - \sum_{s=t+1}^{\tau} \zeta_w \pi_{s-1} - \tilde{W}_t + \tilde{p}_\tau\right)$$

[36] In the case of RANK, λ_t is as discussed in Appendix C.3.1, and the firm and union discount factors are always identical. In the case of HANK, we use the marginal utility evaluated at aggregate consumption (i.e., $\lambda_t = c_t^{-\gamma}$), as in McKay and Wolf (2022), and we just set the discount factor for unions equal to the one for firms to keep the models as comparable as possible.

and so

$$\tilde{W}_t^* - \tilde{W}_t = (1-\bar{\beta}\theta_w)\mathbb{E}_t \sum_{\tau \geq t} \left(\bar{\beta}\theta_w\right)^{\tau-t} \left(\phi\tilde{\ell}_\tau - \tilde{\lambda}_\tau + \sum_{s=t+1}^{\tau} (\pi_s^w - \zeta_w \pi_{s-1}) - \tilde{w}_\tau\right),$$

where $\tilde{w}_\tau \equiv \tilde{W}_\tau - \tilde{p}_\tau$. We will define $\chi_\tau = \varphi\tilde{\ell}_\tau - \tilde{\lambda}_\tau - \tilde{w}_\tau$ to be the labor wedge. Recall that, under our assumptions, we in the HANK model have that $\tilde{\lambda}_t = -\gamma\tilde{c}_t$ where $\tilde{c}_t$ is log-deviations of aggregate consumption.

From the definition of the wage index we have

$$\pi_t^w = \frac{1-\theta_w}{\theta_w}(\tilde{W}_t^* - \tilde{W}_t) + \zeta_w \pi_{t-1}.$$

Combining these relations we get

$$\pi_t^w - \zeta_w \pi_{t-1} = \frac{(1-\theta_w)(1-\beta\theta_w)}{\theta_w}\mathbb{E}_t \sum_{\tau \geq t} \left(\bar{\beta}\theta_w\right)^{\tau-t} \left(\chi_\tau + \sum_{s=t+1}^{\tau} (\pi_s^w - \zeta_w \pi_{s-1})\right)$$

Applying cognitive discounting:

$$\pi_t^w - \zeta_w \pi_{t-1} = \frac{(1-\theta_w)(1-\bar{\beta}\theta_w)}{\theta_w}\mathbb{E}_t \sum_{\tau \geq t} \left(\bar{\beta}\theta_w m\right)^{\tau-t} \left(\chi_\tau + \sum_{s=t+1}^{\tau} (\pi_s^w - \zeta_w \pi_{s-1})\right)$$

This expression has the same structure as (C.5). Operating exactly in the same way as before we obtain

$$\pi_t^w - \zeta_w \pi_{t-1} = \kappa_w \chi_t + \beta\theta_w m \left[1 + \frac{\kappa_w}{1-\beta\theta_w m}\right] \mathbb{E}_t \left[\pi_{t+1}^w - \zeta_w \pi_t\right]$$

With full wage indexation this gives the wage-NKPC used in our main analysis:

$$\pi_t^w - \pi_{t-1} = \kappa_w \chi_t + \beta^w \mathbb{E}_t \left[\pi_{t+1}^w - \pi_t\right] \tag{C.7}$$

C.4 Model calibration and estimation

We now discuss the parameterization of our models. We proceed in two steps—first the calibration part, and then the estimation.

Parameter	Description	Value	Target
$1/\gamma$	EIS	0.5	Standard
$1/\varphi$	Frisch elasticity	0.5	Standard
$\bar{r}$	Real interest rate (annual)	0.04	Real interest rate
α	Capital share	0.36	Christiano et al. (2005)
δ	Depreciation rate (anual)	0.1	Christiano et al. (2005)
δ_1, δ_2	Dividend pay-out process	$0.2, 0.05$	Capital Gains MPC
$\bar{\tau}_\ell$	Labor tax rate	0.3	Average Labor Tax
$\bar{\tau}/\bar{y}$	Transfers	0.05	Wolf (2023)
$\bar{b}/\bar{y}$	Steady state liquid assets	1.04	Kaplan et al. (2018)
$1/\eta$	Liquid assets duration (quarters)	5	Kaplan et al. (2018)
τ_b^ℓ	Speed of fiscal adjustment	0.15	Gradual fiscal adjustment

Table C.3: Calibrated model parameters.

Calibration. For all three models, we calibrate the elasticity of intertemporal substitution and the Frisch elasticity to be $\frac{1}{2}$, which are standard values in the literature. For RANK, we set $\beta = 0.99$ (quarterly) in order to get a real interest rate of 4 percent annualized. For HANK, we pick β in order to match the same steady-state level of assets for all models. We calibrate the idiosyncratic income process for HANK from Kaplan et al. (2018).

We set the capital share to $\alpha = 0.36$ and depreciation rate to $\delta = 0.025$ quarterly, which is consistent with the values used in Christiano et al. (2005). The dividend distribution process is parameterized by assuming $\delta_1 = 0.2$ and $\delta_2 = 0.05$, which ensures a gradual payment of dividends and therefore low consumption response from capital gains.[37]

We follow Wolf (2023) for the steady state calibration of the fiscal side. We assume a labor tax rate $\bar{\tau}_\ell$ of 0.3, and set transfers to be 5 percent of GDP. The steady state level of nominal assets is set to 27 percent of annual GDP, as in Kaplan et al. (2018). Government debt maturity is calibrated to $\eta = 0.2$, which implies an average debt duration of 5 quarters. The steady-state level of government expenditure is set such that the budget constraint holds in steady state, which yields $\frac{\bar{g}}{\bar{y}} = 0.1395$. We assume that all dynamic fiscal adjustment is done via labor taxes, with $\tau_b^\ell = 0.15$. This implies gradual fiscal adjustment, in line with the range considered in Auclert et al. (2020).

A summary of the calibrated parameter values is provided in Table C.3.[38]

[37] As long as the pay-out is gradual, our results are not sensitive to the specific values used.

[38] The baseline determinacy-induced monetary policy rule that we consider sets $\rho = 0.85$, $\phi_\pi = 2$, $\phi_y = 0.25$, and $\phi_{\Delta y} = 0.3$. Recall that this choice of rule only matters for our illustrative results in Figures 3 and 4.

Estimation. We estimate all models to ensure consistency with the empirical monetary policy shock impulse response targets $\hat{\theta}_\nu$. For the baseline RANK model we estimate five parameters: the strength of habits (h), the degrees of price as well as wage rigidity (θ_p and θ_w), the curvature of investment adjustment costs (κ), and the curvature of capacity utilization costs (ζ). For the baseline HANK model, the household information stickiness parameter (θ) replaces the degree of habit formation (h). Finally, for the behavioral models, consider the case of m fixed and set to $m = 0.65$, at the lower end of the range considered by Gabaix (2020). We make this choice because our data are only weakly informative about m, as is implicit in the results displayed in Table 4.1.

Table C.4 summarizes the posterior distributions of all estimated parameters. We see that, for h, κ and ψ, posterior distributions are relatively close to the prior. On the other hand, the distributions of θ_p, θ_w and θ are meaningfully affected. In the cases of θ_p and θ_w, the level of price and wage stickiness required to fit the impulse responses is relatively large, especially for prices; this reflects the known mismatch between micro level and macro level estimates of price rigidity, with macro estimates pointing towards much stickier prices than micro evidence. For the case of θ, a higher degree of informational stickiness is required to fit the empirical impulse responses than the one encoded in the prior. The degree of information rigidity is close to the one inferred in Auclert et al. (2020).

Model	Parameter	Prior			Posterior				
		Dist.	Mean	St. Dev	Mode	Mean	Median	5 percent	95 percent
RANK - RE	h	Beta	0.70	0.10	0.7240	0.7082	0.7157	0.5335	0.8571
	θ_p	Beta	0.67	0.20	0.9485	0.8622	0.9136	0.5598	0.9804
	θ_w	Beta	0.67	0.20	0.8860	0.7544	0.8091	0.3706	0.9657
	κ	Normal	5.00	1.50	5.0527	5.2668	5.2512	2.9409	7.6503
	ψ	Beta	0.50	0.15	0.4621	0.4665	0.4645	0.2247	0.7175
HANK - RE	θ	Beta	0.70	0.20	0.9526	0.7985	0.8414	0.4655	0.9813
	θ_p	Beta	0.67	0.20	0.9521	0.8527	0.9070	0.5381	0.9822
	θ_w	Beta	0.67	0.20	0.9031	0.7664	0.8215	0.3860	0.9683
	κ	Normal	5.00	1.50	5.3003	5.2477	5.2396	2.8847	7.6453
	ψ	Beta	0.50	0.15	0.4654	0.4678	0.4648	0.2247	0.7212
RANK - CD	h	Beta	0.70	0.10	0.7102	0.7112	0.7187	0.5373	0.8609
	θ_p	Beta	0.67	0.20	0.8641	0.8664	0.9061	0.6166	0.9769
	θ_w	Beta	0.67	0.20	0.9462	0.7564	0.8118	0.3691	0.9646
	κ	Normal	5.00	1.50	5.2998	5.3312	5.3211	3.0123	7.6935
	ψ	Beta	0.50	0.15	0.4701	0.4714	0.4686	0.2292	0.7244
HANK - CD	θ	Beta	0.70	0.20	0.9600	0.8112	0.8553	0.4866	0.9847
	θ_p	Beta	0.67	0.20	0.8544	0.8467	0.8917	0.5507	0.9794
	θ_w	Beta	0.67	0.20	0.9511	0.7860	0.8483	0.4005	0.9711
	κ	Normal	5.00	1.50	5.3222	5.3264	5.3186	2.9636	7.7030
	ψ	Beta	0.50	0.15	0.4692	0.4607	0.4571	0.2244	0.7119

Table C.4: Prior and posterior distributions of structural parameters. RE denotes that the model assumes rational expectations ($m = 1$), whereas CD indicates that the model features cognitive discounting in price and wage setters (with $m = 0.65$).

D Supplementary details for empirical applications

This appendix contains supplementary results for our three monetary policy counterfactual applications in Section 5.

D.1 Average business cycle

We here substantiate our claims that the headline finding of Section 5.2—i.e., that meaningful volatility reductions in output and, to a lesser extent, inflation would have been feasible—are driven by neither the Great Recession nor by the policy causal effect extrapolation embedded in our structural models (i.e, by the Plus step).

We begin in Figure D.1 by instead constructing counterfactual volatilities with reduced-form forecasts obtained on a sample that only stretches to 2007:Q1. We see that the picture is essentially unchanged relative to Figure 5: inflation and in particular output gap volatility reductions are feasible, at the cost of somewhat more volatile interest rates. This robustness is not surprising: the main business-cycle shock of Angeletos et al. (2020) meaningfully moves aggregate output even on pre-ZLB samples (while having rather little effect on inflation), so the exact same logic from our discussion in Section 5.2 continues to apply.

Next, in Figure D.2, we repeat our baseline analysis, but now minimizing (19) using only the matched policy shock impulse responses $\hat{\theta}_\nu$, rather than the entirety of the model-implied policy causal effect matrices Θ_ν. Qualitatively, the exact same picture emerges as before: inflation and in particular output are less volatile, while interest rates are somewhat more volatile. The intuition is yet again immediate from our analysis of the main business-cycle shock in Figure 6: through nominal interest rate cuts—including those directly matched in $\hat{\theta}_\nu$—the policymaker can meaningfully reduce the volatility of output fluctuations.[39] It is thus the empirical VAR step—and not the additional structural assumptions embedded in the Plus step—that drive the headline takeaways of Section 5.2, as claimed.

D.2 Great Recession

Figure D.3 constructs the Great Recession counterfactual using only the matched policy shock impulse responses $\hat{\theta}_\nu$ (rather than all of Θ_ν). As before, we here use $\hat{\theta}_\nu$ to enforce the counterfactual policy rule of interest as well as possible.

[39] The only difference is quantitative: the entire matrix of monetary policy causal effects Θ_ν allows the policymaker to tailor her interest rate response even better.

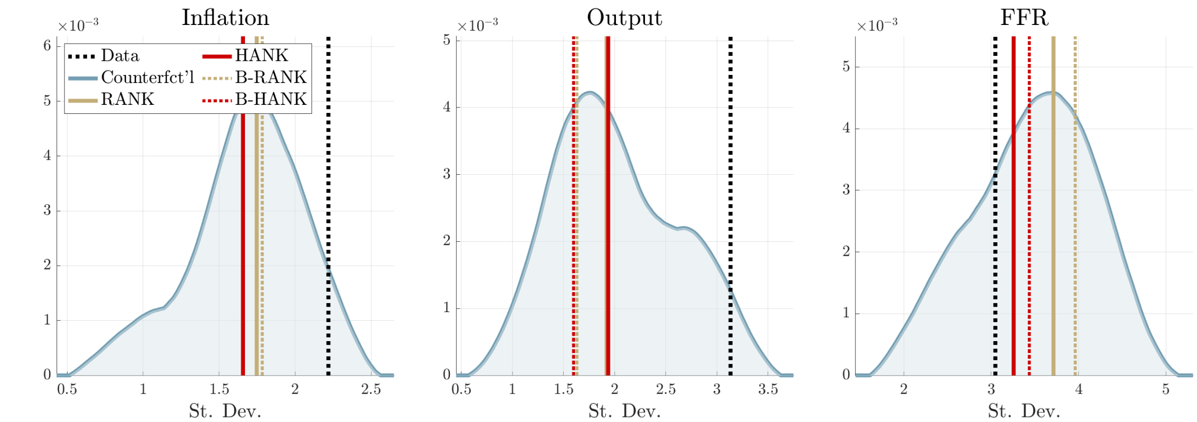

Figure D.1: Counterfactual early-sample (1960:Q1 – 2007:Q1) average volatilities of inflation, output, and the federal funds rate, under the policy rule that minimizes (19). Black dashed: data point estimate under observed policy. Blue: posterior Kernel density of counterfactual volatilities drawing from posterior across all models and parameters. Beige: posterior mode of counterfactual using RANK models (baseline and behavioral). Red: posterior mode of counterfactual using HANK models (baseline and behavioral).

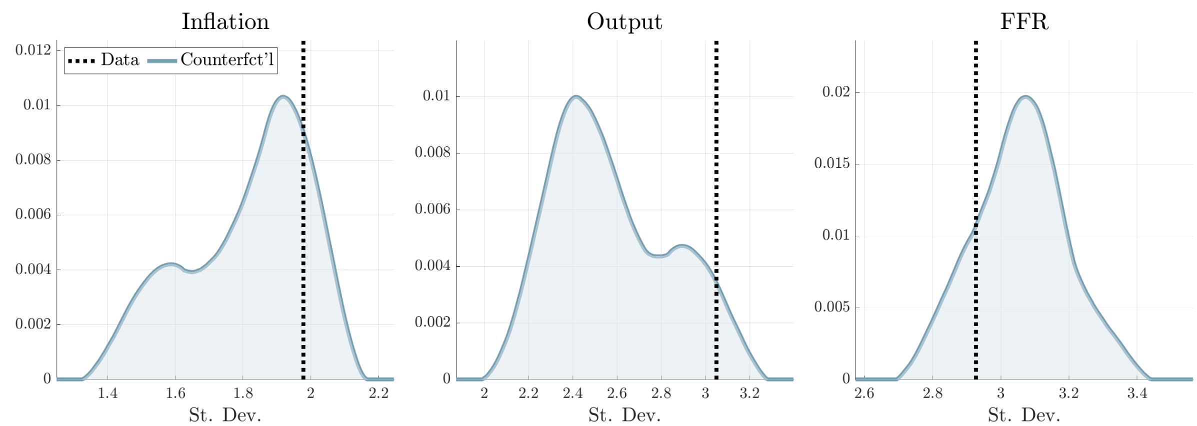

Figure D.2: Counterfactual average volatilities of inflation, output, and the federal funds rate, under the policy that minimizes (19). Black dashed: data point estimate. Blue: posterior Kernel density of estimates drawing from posterior across all models and parameters, using only the matched policy shock impulse responses $\hat{\theta}_\nu$.

The resulting counterfactuals are again broadly similar to our main results in Figure 7: the nominal interest rate is cut aggressively, leading to more stable output, at the cost of elevated inflation. Moving from the restricted policy causal effect space $\hat{\theta}_\nu$ to the entirety of Θ_ν smoothes out the rate cut and helps somewhat better stabilize output; that being said, the differences are relatively small, suggesting that the counterfactual policy does not rely much on model-implied extrapolation to the causal effects of interest rates forward guidance.

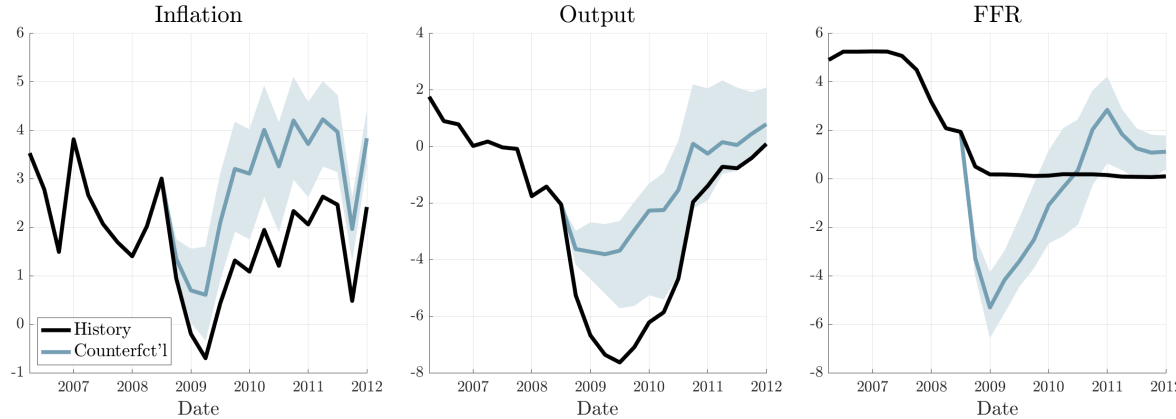

Figure D.3: Counterfactual evolution of inflation, output, and the federal funds rate in the Great Recession, under the policy that minimizes (19) without any effective lower bound on rates. Black: data. Blue: posterior median (solid) and 16th and 84th percentile bands (shaded), using only the matched policy shock impulse responses $\hat{\theta}_\nu$.

It is thus yet again the VAR step—and not the additional structure of the Plus step—that drives our findings in this second application.

D.3 Post-covid inflation

Figure D.4 complements the analysis in Section 5.4 by constructing the post-covid inflation forecasting counterfactual using only the matched monetary policy shock impulse responses $\hat{\theta}_\nu$ (and not all of Θ_ν). Recall that the disagreement across models discussed in Section 5.4 was precisely related to the model-based extrapolation from $\hat{\theta}_\nu$ to the rest of Θ_ν: in the rational-expectations models, the policymaker leverages small real rate hikes in the far future to stabilize inflation today; in the behavioral models, this is not possible, so interest rates are instead hiked already today.

As expected, Figure D.4 reveals that, when not relying on any extrapolation, the counterfactual looks much closer to the behavioral case: interest rates are actually hiked more aggressively, and that then leads to moderate declines in inflation and output. This third application is thus a clear example of where the additional extrapolation of the Plus step matters greatly for the final results.

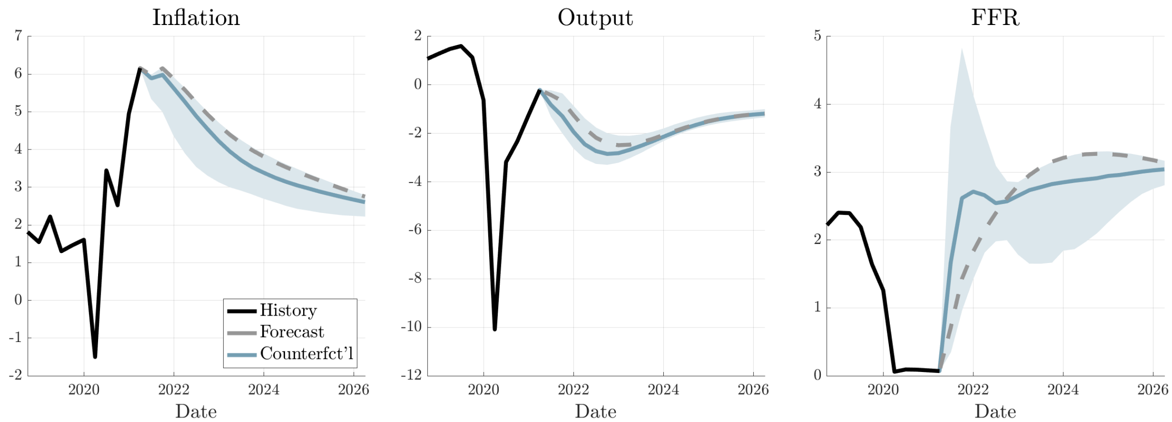

Figure D.4: Counterfactual projections of inflation, output, and the federal funds rate in the post-covid inflationary episode (from 2021:Q2), under the policy that minimizes (19). Policy causal effects from posterior across all models and parameters, but using only the matched policy shock impulse responses $\hat{\theta}_{\nu}$. Black: data. Grey: actual forecast. Blue: posterior median (solid) and 16th and 84th percentile bands (shaded).