Evaluating Monetary Policy Counterfactuals: When Do We Need Structural Models?†

Tomás E. Caravello   Alisdair McKay   Christian K. Wolf

MIT   FRB Minneapolis   MIT & NBER

September 4, 2026

Abstract: We give conditions under which knowledge of the effects of monetary policy shocks suffices to evaluate policy counterfactuals that change not just the monetary rule, but also the nature of equilibrium selection. For example, the effects of monetary policy shocks in a regime of monetary dominance can be used to evaluate counterfactual outcomes even under a regime switch to fiscal dominance. Since the empirical literature delivers the causal effects of short-lived monetary shocks, the sole remaining role of model structure in evaluating such counterfactuals is thus to extrapolate from the effects of transitory to those of more persistent policy rate changes. Among popular models of monetary policy transmission, household heterogeneity (as in the burgeoning "HANK" literature) does not change this extrapolation very much, while behavioral frictions do.

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

†Email: tomasec@mit.edu, alisdair.mckay@mpls.frb.org, and ckwolf@mit.edu. We received helpful comments from the editor, Robert Barro, four anonymous referees, Marios Angeletos, Régis Barnichon, Francesco Bianchi, Luigi Bocola, Ricardo Caballero, Gabriel Chodorow-Reich, Marco Del Negro, Martin Eichenbaum, Xavier Gabaix, Simon Gilchrist, Cosmin Ilut, Geert Mesters, Mikkel Plagborg-Møller, Giorgio Primiceri, Valerie Ramey, Matt Rognlie, Karthik Sastry, 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

A central challenge in monetary economics is how to construct credible counterfactuals for alternative policy regimes. Following Lucas (1976), the standard way to evaluate such counterfactuals has been to build a micro-founded model of macroeconomic fluctuations, change monetary policy in that model, and re-solve. A different tradition, going back to Sims (1980), instead directly leverages empirical evidence associated with historical policy variation. Recent work (see McKay and Wolf, 2023; Barnichon and Mesters, 2023) has demonstrated that such empirical evidence, together with weak structural assumptions, can suffice to construct valid, so-called "semi-structural" policy counterfactuals.

While increasingly popular, this semi-structural approach has been limited by two fundamental challenges. The first challenge is conceptual: prior work can speak to policy changes *within* a policy regime, e.g., covering changes in interest rate feedback rules conditional on monetary dominance, but not *across* regimes, e.g., switching to fiscal dominance. The second challenge is practical: the researcher needs the dynamic causal effects of changes in policy rates at all current and future horizons, a clearly formidable requirement.

Our principal contribution is to revisit and at least partially circumvent these limitations. We first generalize the identification results of the semi-structural approach to also allow the researcher to evaluate the effects of changes across policy regimes, notably between monetary and fiscal dominance. Turning to the challenge of getting the full menu of policy rate causal effects, we show that the empirical evidence on monetary policy shocks chiefly pins down the causal effects of transitory rate changes, leaving a single remaining role for model structure: to extrapolate beyond the observed support of the data to deliver the causal effects of gradual rate changes. We compare how various frontier models of monetary policy transmission achieve this extrapolation, and analyze what frictions matter most for this key purpose. We finally turn to several applications that illustrate our results, including one to post-2021 inflation dynamics under either monetary or fiscal dominance.

Evaluating monetary policy counterfactuals. Our object of interest throughout this paper is the counterfactual evolution of the macro-economy under alternative assumptions on systematic monetary policy conduct, both on average, for how the "typical" business cycle would have unfolded, and for particular historical episodes. Recent contributions to the semi-structural approach, as referenced above, generally focus on counterfactual linear feedback rules that leave equilibrium selection unchanged, e.g., changing from one Taylor-type interest rate rule to another. For a relatively general family of linearized structural macroeconomic models, prior work has shown that such policy counterfactuals are pinned down by two sets of "sufficient statistics." The first set consists of *reduced-form projections*: impulse responses to Wold innovations for the average business cycle, and forecasts for conditional historical counterfactuals. The second consists of *policy causal effects*, i.e., the causal effects of current and expected future policy rate changes on the macro-economy.

Our first contribution is to establish that, for counterfactual policies that are consistent with a range of possible equilibria, the sufficient statistics of the semi-structural approach identify *all* of the associated equilibrium outcomes. We can then select among those equilibria using any desired criterion, ranging from monetary dominance, to continuity at the boundary of determinacy (as considered in Lubik and Schorfheide, 2004), to the fiscal theory of the price level (FTPL, see Cochrane, 2023). Remarkably, these results imply that the effects of shocks to an active monetary policy rule also suffice to recover the counterfactual outcomes in the very different fiscal dominance regime (as in Leeper, 1991)—perhaps the most widely studied example of a large-scale policy regime shift. In other words, we here add to the semi-structural literature the insight that policy causal effects are actually portable across seemingly very different policy regimes.

We next establish that the same sufficient statistics allow evaluation not just of permanent changes in equilibrium selection, but even of stochastic switches between policy regimes, as commonly assumed in the literature on the history of U.S. monetary policy conduct (e.g., see Davig and Leeper, 2007; Bianchi, 2013; Bianchi and Melosi, 2017, among many others). Intuitively, given linearity, the causal effects of current and expected future interest rates summarize all the ways in which interest rate policy impacts the macro-economy. Moving to stochastic, recurring regime switches complicates the process governing expectations for the policy instrument; it does not, however, change the fact that the sufficient statistics are still all you need to evaluate that process and its effects.

Reduced-form projections. The preceding identification results require reduced-form projections formed with respect to an information set that spans the entire history of shocks hitting the macro-economy, i.e., the assumption of "invertibility," familiar from and often criticized in the literature on shock identification (e.g., see Plagborg-Møller and Wolf, 2022). The purpose of invertibility here, however, is very different from its role in that literature: it simply ensures that projections are formed using a complete information set, i.e., "oracle" forecasts (as in Barnichon and Mesters, 2023). Our second contribution is to argue theoretically and through simulations that even projections with respect to a relatively small information set can actually deliver accurate policy counterfactuals; what matters is that the information set contains the main predictors for the macroeconomic outcomes of interest.

Monetary policy causal effects. It remains to recover the second sufficient statistic: the dynamic causal effects of changes in the policy rate. We begin by reviewing the available empirical evidence on monetary policy shock propagation, and argue that it pins down the effects of transitory interest rate changes. In principle, each empirically identified monetary policy shock may capture a different policy "treatment"—some may only revise the policy rate today, while others may change expectations of rates in the future. In practice, however, that is unfortunately not the case, with most of the available data variation speaking only to the effects of transitory rate changes. Representative examples include Romer and Romer (2004), who estimate the effects of a short-lived change in rates, and the recent refinement in Aruoba and Drechsel (2026). Many of the other popular monetary shock series, e.g., Gertler and Karadi (2015), Miranda-Agrippino and Ricco (2021), or even the "forward guidance" disturbance in Swanson (2024), behave similarly.

With historical data uninformative about the effects of more delayed or persistent changes in policy rates, a role for model structure emerges: to substitute for that missing data variation. We thus consider several frontier quantitative models that deliver this mapping, from representative-agent (RANK) to heterogeneous-agent (HANK) models, with and without behavioral frictions. Consistent with prior work, we document that RANK and HANK models can both match the available evidence on the aggregate effects of transitory monetary shocks. We then show that, once the two models are disciplined in this way, they also approximately agree on the unknown aggregate effects of persistent policy changes. Intuitively, while it is well-known that market incompleteness can in principle dampen or amplify monetary policy propagation (Auclert, 2019), our insistence on identical responses to the observed short-lived changes in policy means that any heterogeneity-related channels largely offset, delivering as-if aggregation that echoes Werning (2015).[1] In particular we find that, in both RANK and HANK, short-run economic conditions are highly sensitive to assumptions about far-ahead nominal interest rates; if anything, forward guidance is actually even a bit more powerful in HANK, consistent with Auclert et al. (2020). This contrasts sharply with alternative models featuring strong behavioral frictions in the expectations of price setters (as in Gabaix, 2020),

[1]This equivalence, also reported in Auclert et al. (2025), reflects the fact that our RANK and HANK models differ *only* on the demand side. Changes in supply, e.g., via household labor supply, may break it.

as familiar from the literature on the forward guidance puzzle (Del Negro et al., 2023).

Applications. With our identification results and sufficient statistics in hand, we turn to several applications, designed to illustrate the main lessons of our theoretical analysis.

The first application revisits post-2021 inflation dynamics. A common criticism, both at the time and also ex post, was that the Federal Reserve was slow to hike its policy rate, in contrast to the prescriptions of standard Taylor-type rules. We evaluate this counterfactual using our approach and find that, largely independently of what model is used for monetary policy causal effect extrapolation, following a conventional Taylor-type rule would have, under monetary dominance, indeed led to materially lower inflation, though at the cost of lower output (echoing the earlier findings of Barnichon, 2022). Next, using our new results, we also evaluate a counterfactual for the same aggressive policy rate path, but under fiscal dominance. Relying on the exact same monetary policy causal effects as before, but switching assumptions on equilibrium selection, we see that the rate hikes if anything increase inflation, reflecting the familiar "stepping on a rake" effect of the FTPL (Sims, 2011).

In our second application we argue that monetary policy, assuming a consistent "passive" fiscal backing, could have materially reduced the volatility of output over a long post-war sample period, while at the same time also reducing inflation volatility. This conclusion follows straight from empirical measurement: it obtains using reduced-form projections under either small or large information sets, illustrating the (non-)invertibility result, and relies chiefly on the empirically estimable effects of transitory policy rate changes, with causal effect extrapolation relatively unimportant. Key to this finding is that most post-war output fluctuations look demand-driven (Angeletos et al., 2020), and can thus be moderated effectively through interest rate policy. This second application in particular showcases that our approach to policy counterfactual evaluation collapses to the purely semi-structural method of prior work whenever empirical evidence alone already suffices.

For the third application we consider commitment, either perfect or stochastic, to forward guidance beyond the duration of the Great Recession. This final application not only illustrates the applicability of our method to non-linearities and stochasticity in policy conduct, but is also the first instance in which assumptions for causal-effect extrapolation take center stage: under rational expectations, whether commitment to forward guidance is perfect or stochastic matters greatly; with behavioral frictions, it is much less important, reflecting a reduced power of forward guidance. Assumptions about the strength of behavioral frictions thus in particular matter much more than the choice between RANK and HANK, revealing that differences in transmission *channels*—the direct-indirect dichotomy stressed in Kaplan et al. (2018)—may not matter all that much for macroeconomic dynamics.

Scope and limitations. Our analysis addresses some of the main limitations of recent incarnations of the semi-structural approach. A key constraint that we do not relax, however, is linearity of the non-policy block, allowing us to restrict attention to the effects of (expected) rate paths, without any regard for their size, sign, or the state of the economy. This focus is shared by a large literature that uses fully specified general equilibrium models for monetary policy counterfactual evaluation, but it does come with material restrictions, notably ruling out changes in the economy's steady state (e.g., changes in the inflation target), or time and state dependence in the rest of the economy (e.g., see Gagliardone et al., 2026, for pricing).[2]

Further literature. We already discussed how our analysis relates to the semi-structural literature on policy evaluation; here, we elaborate further on the relation to the Lucas program. Contributions to that research program broadly fall into two camps. The first one leverages general equilibrium models estimated via full-information, likelihood-based approaches (e.g., Smets and Wouters, 2007). Our results reveal that this approach is actually more robust than commonly believed: likely model misspecification, as criticized by some commentators (e.g., Chari et al., 2009), distorts conclusions if, *and only if*, it either leads to data-inconsistent reduced-form projections (i.e., second moments), or to incorrect policy instrument causal effects. Conditional on those two "sufficient statistics," and for a space of policy counterfactual questions that spans many of those studied in prior structural work, the details of the assumed model structure then actually play no further role. Of course, if the model is misspecified, then matching data second moments—which is what likelihood-based estimation achieves—may distort policy causal effects.

Our proposed approach to recovering rate causal effects outside the support of the data instead directly builds on the limited-information strategy of the second camp: impulse response matching (Christiano et al., 2010). As this strategy only leverages a partial description of the economy, it is on its own insufficient to speak to policy counterfactual questions; our analysis reveals, however, that the assumed structure is just what is needed to deliver the sole input that the pure semi-structural approach misses.

[2]Notable examples of prior structural work that evaluates policy counterfactuals similar to ours (e.g., stochastic regime-switching and fiscal dominance), and linearizes the non-policy block, include Davig and Leeper (2007), Cogley et al. (2010), Bianchi and Melosi (2017), and Bocola et al. (2024).

2 Evaluating monetary policy counterfactuals

This section presents our monetary policy counterfactual identification results. Section 2.1 briefly reviews the recent semi-structural approach. Sections 2.2 and 2.3 then continue with our main novel results on policy regime change, both permanent and stochastic.

2.1 Environment and baseline identification results

We state identification results for the counterfactual evolution of the macro-economy under alternative assumptions on systematic policy design, both on average, for how the "typical" business cycle would have unfolded, and for particular historical episodes. The statements are a synthesis and minor extension of those in earlier work (notably McKay and Wolf, 2023; Barnichon and Mesters, 2023, 2026), so the discussion will be brief, but self-contained.

Environment. We consider a stochastic economy that admits representation as a general structural vector moving average process (SVMA):

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

y_t is a vector of macroeconomic aggregates, the aggregate shocks ε_t are distributed as $\varepsilon_t \sim N(0, I)$, and the $n_y \times n_\varepsilon$-dimensional matrices Θ_ℓ denote the (square-summable) impulse response of y_t at horizon ℓ to a date-t vector of shocks ε_t. In the following, the notation $\mathbb{E}_t[\bullet]$ is reserved for expectations conditioning on the sequence of shocks $\{\varepsilon_{t-\ell}\}_{\ell=0}^{\infty}$ up to date t.

Leveraging the equivalence between linearized systems with aggregate risk and perfect-foresight transition paths, we obtain the impulse responses Θ_ℓ as solutions of a linear, perfect-foresight, infinite-horizon economy. Below boldface denotes time paths for $t = 0, 1, 2, \dots$, and all variables are expressed in deviations from the deterministic steady state. The perfect-foresight system is

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

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

Here x_t and w_t are n_x- and n_w-dimensional vectors of endogenous variables, 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 $y_t = (x_t', z_t')'$,

$\varepsilon_t = (e_t', v_t')'$.[3] The distinction between w and x is that the x's are observable to the econometrician while the w's 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 policy instrument z set as a function of x and v. In our applications to monetary policy, the instrument z is typically the short-term policy rate, and any endogenous fiscal reaction is subsumed in the "non-policy" block (2).

Given the date-0 shocks $\{e_0, v_0\}$, impulse responses are sets of bounded sequences $\{\boldsymbol{w}, \boldsymbol{x}, \boldsymbol{z}\}$ that solve (2)–(3). We assume that this solution exists and is unique, and write it as

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

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

Objects of interest. We wish to study the evolution of the economy if policy were instead, and counterfactually, set as

$$\tilde{\mathcal{A}}_x \boldsymbol{x} + \tilde{\mathcal{A}}_z \boldsymbol{z} = \boldsymbol{0} \tag{4}$$

rather than (3).[4] Assuming equilibrium existence and uniqueness for this rule (an assumption we will relax in Section 2.2), we recover the counterfactual SVMA process

$$\tilde{y}_t = \sum_{\ell=0}^{\infty} \tilde{\Theta}_\ell \varepsilon_{t-\ell}, \tag{5}$$

with the convention that $\varepsilon_t = e_t$, and where the shock impulse responses $\tilde{\Theta}_\ell$ are derived from the solution of the perfect-foresight system (2) together with (4).

We are interested in the effects of the counterfactual policy change both *unconditionally*, for the "average" business cycle, as well as *conditionally*, for particular historical episodes. The counterfactual properties of the average cycle are, in our stationary setting, fully sum-

[3] 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.

[4] Note that a counterfactual rule like (4) can, once suitably re-interpreted, also nest examples of changes in policy design that switch the policy instrument. For example, for monetary policy, if a private-sector money demand relation ties together interest rates and money supply, then a rule like (4) is general enough to capture a counterfactual switch in the instrument of monetary policy from interest rates to money supply, or *vice versa*. We elaborate on this observation, and also its limitations, in Appendix A.2.

marized by the autocovariance function of $\tilde{y}_t$, given as

$$\tilde{\Gamma}_y(\ell) = \sum_{m=0}^{\infty} \tilde{\Theta}_m \tilde{\Theta}'_{m+\ell}. \tag{6}$$

For specific historical episodes, the counterfactual evolution of the economy is instead directly given by the counterfactual process (5), with the actual history of shocks $\{\varepsilon_{t-\ell}\}_{\ell=0}^{\infty}$ instead propagating according to the counterfactual impulse responses $\{\tilde{\Theta}_\ell\}_{\ell=0}^{\infty}$.[5]

Identification result. Our objects of interest are, in this model environment, pinned down by two "sufficient statistics": policy causal effects and reduced-form projections.

Policy causal effects are the impulse responses of shocks to the policy instrument z *at all possible horizons*. To define these effects consider the following generalized policy rule:

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

While the *actual* policy rule (3) is subject only to the n_v-dimensional vector of policy shocks v_0 (which may be low-dimensional, or in fact even empty), the policy shock vector $\boldsymbol{\nu}$ in the artificial rule (3') is instead unrestricted—i.e., we allow for arbitrarily flexible wedges in the policy rule at each date $t = 0, 1, 2, \ldots$. Analogously to the discussion above, the solution of (2)–(3') given an arbitrary policy shock vector $\boldsymbol{\nu}$ alone yields

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

Θ_ν characterizes the *space* of y-allocations implementable through policy shocks, e.g., through monetary shocks at every point along the yield curve, from the short end to the long end.

The required reduced-form projections are a function of the Wold representation of y_t under the *actual* (not counterfactual) prevailing policy rule, given as

$$y_t = \sum_{\ell=0}^{\infty} \Psi_\ell u_{t-\ell}, \tag{7}$$

where u_t are orthogonalized Wold innovations, with $\text{Var}(u_t) = I$. Specifically, let $u_t^* \equiv$

[5] For conditional counterfactuals, matters are actually slightly more complicated than stated in the main text if we contemplate a policy switch that begins at some given date t^* (rather than the infinite past). If that is so, and if the economy was not at the deterministic steady state at $t^* - 1$, then the process (5) needs to be augmented with suitable initial conditions. We provide a precise discussion in Appendix A.1.

$y_t - \mathbb{E}\left(y_t \mid \{y_\tau\}_{-\infty<\tau\leq t-1}\right)$ denote one-step-ahead forecast errors under the baseline SVMA (1) with $\text{Var}(u_t^*) = \Sigma_u$, and Σ_u assumed invertible. Then $u_t \equiv \text{chol}(\Sigma_u)^{-1}u_t^*$ with $\text{Var}(u_t) = I$ and $\text{chol}(\bullet)$ giving the lower-triangular Cholesky factor.

The identification result now states that, under the additional assumption of *invertibility*, these two objects actually suffice to recover all of our policy counterfactuals of interest. As already mentioned above, the result builds closely on McKay and Wolf (2023) and Barnichon and Mesters (2023), with only the arguments for the counterfactual historical evolution of the economy new here. After stating the proposition, we briefly discuss its intuition.

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

$$\varepsilon_t \in \text{span}(\{y_\tau\}_{-\infty<\tau\leq t}), \tag{8}$$

and that the counterfactual rule (4) *delivers a unique equilibrium. Then knowledge of policy causal effects, Θ_ν, and the Wold representation of y_t,* (7)*, suffices to evaluate the counterfactual SVMA process $\tilde{y}_t$* (5) *and construct the counterfactual autocovariance function* (6)*.*

The reduced-form projections, summarized by the Wold representation of the data y_t, are shaped by both the actual in-sample policy rule (3) together with the structural shocks ε_t hitting the economy. Knowing the causal effects of policy, as encapsulated in the matrix Θ_ν, then allows us to do two things: first, *strip out* the effects of the in-sample policy; and second, *add back in* the hypothesized counterfactual policy. The formal proof of Proposition 1 leverages this intuition as follows. It begins with the reduced-form Wold innovations u_t and their impulse responses $\Psi(L)$; by invertibility, those innovations span the history of the true structural shocks $\{\varepsilon_{t-\ell}\}_{\ell=0}^{\infty}$, ensuring that Wold shocks and impulse responses capture all the relevant primitives of the economy. The second step is to use the policy causal effects Θ_ν to assess how the reduced-form innovations u_t would have propagated under a counterfactual policy rule, i.e., to map the actual reduced-form impulse responses $\Psi(L)$ into counterfactual responses $\tilde{\Psi}(L)$. Implicitly, this step removes the effects of the actual policy and adds in the effects of the new policy, turning actual into counterfactual Wold representation. With that representation in hand, all counterfactuals of interest can be evaluated.

Looking ahead. In the Sims (1980) tradition, Proposition 1 ties outcomes under counterfactual assumptions on policy conduct to two in principle measurable objects: reduced-form projections and policy shock effects. There are, however, two material challenges. First, the space of counterfactuals covered by the identification result is somewhat limited: it covers changes from one determinacy-inducing rule to another, thus implicitly restricting attention within an otherwise stable policy regime. For example, changes between two determinacy-inducing Taylor rules are covered, while stochastic switches between such rules, or changes in equilibrium selection (e.g., between monetary vs. fiscal dominance, as in Leeper, 1991), are not. Second, the informational requirements are very high: the reduced-form projections need to be formed with respect to an information set sufficiently large to ensure invertibility, and policy causal effects are an infinite-dimensional object, collecting the effects of both contemporaneous and all future expected policy instrument changes.

In the remainder of this paper we deal with those two challenges. First, in Sections 2.2 and 2.3, we extend the identification result in Proposition 1 to cover various notions of policy regime change. Second, in Sections 3 and 4, we turn to the required inputs: we first elaborate on the role of invertibility for the reduced-form projections, and then discuss how to use a mix of empirical evidence and model structure to deliver policy causal effects.

2.2 Equilibrium selection and fiscal dominance

This section presents our main novel theoretical results on policy counterfactuals involving "regime change" in equilibrium selection, with the leading practical example being changes between monetary and fiscal dominance. We will proceed in three steps. First, we generalize our preceding identification results to rule changes that allow for equilibrium multiplicity, thus now opening the door for equilibrium selection. Second, we show by example how this general discussion applies to changes between monetary and fiscal dominance in a textbook New Keynesian model. And third, we close with some supplementary discussion of the role of fiscal policy beyond equilibrium selection.

Equilibrium multiplicity and equilibrium selection. We consider the same environment as in Section 2.1, but forgo the assumption that the counterfactual rule (4) induces a unique equilibrium, and so a unique counterfactual SVMA process (5); instead, there are now, for each individual structural shock $\varepsilon_{i,t}$, many different possible impulse response paths $\tilde{\Theta}_{i,\ell}$, corresponding to different date-0 sunspots.

Proposition 2. *Suppose that the counterfactual policy rule* (4) *delivers equilibrium existence, but not necessarily uniqueness. Then, for each shock $\varepsilon_{i,t}$, knowledge of that shock's impulse responses under the baseline rule, $\Theta_{i,\ell}$, together with policy causal effects Θ_{ν}, suffices to evaluate* all possible *counterfactual impulse responses $\tilde{\Theta}_{i,\ell}$.*

Relative to prior work on the identification of policy counterfactuals, Proposition 2 delivers the full set of counterfactual equilibria, allowing equilibrium selection itself to change. To see the intuition, consider impulse responses under the initial, baseline policy rule to any shock $\varepsilon_{i,t}$. The basic logic of all policy counterfactual identification results—both Proposition 1 and also its antecedents in earlier work—is to use artificial policy shocks to the baseline rule to instead enforce the counterfactual rule. The key twist here is that, if the hypothesized alternative rule is consistent with multiple equilibria, then each such equilibrium is associated with a *different* sequence of possible artificial policy shocks, corresponding to different initial conditions, or equivalently different date-0 sunspots. In other words, knowledge of the initial impulse response $\Theta_{i,\ell}$ and policy causal effects Θ_ν now delivers a system of equations whose solution set is precisely the full set of possible counterfactual equilibria. The deeper reason is that the space of allocations implementable by policy depends on the maps $\{\mathcal{H}_w, \mathcal{H}_x, \mathcal{H}_z\}$, and not on the policy "regime," i.e., the assumed equilibrium selection, that generated the data; put succinctly, policy causal effects are portable across regimes. The proof of the proposition is constructive, explicitly mapping each possible equilibrium into an "as-if" shock sequence to the initial policy rule.

Given Proposition 2, we can now add assumptions on equilibrium selection—with fiscal dominance the leading practical example—to restore uniqueness and thus arrive at a generalized version of our main identification result.

Corollary 1. *Suppose that the SVMA(∞) process* (1) *is invertible, and that the counterfactual rule* (4) *delivers equilibrium existence. Knowledge of Θ_ν and the Wold representation* (7)*, together with an assumption on equilibrium selection, suffices to evaluate the counterfactual SVMA process $\tilde{y}_t$ and construct the counterfactual autocovariance function* (6)*.*

Echoing Proposition 1, the proof of Corollary 1 evaluates the counterfactual propagation of the Wold innovations, just now picking among the many possible counterfactual responses $\tilde{\Psi}(L)$ according to the assumed equilibrium selection criterion.

Application to fiscal dominance. We illustrate the preceding abstract discussion using the monetary policy literature's canonical example of regime change: moving between monetary and fiscal dominance. In keeping with modern incarnations of the FTPL, we use a textbook New Keynesian model (Galí, 2015; Woodford, 2003) as the model environment for our illustration, with details on model and parameterization provided in Appendix A.5. We consider an econometrician that lives in this model, and is able to estimate both the propagation of a cost-push shock as well as interest rate causal effects in a regime of *monetary*

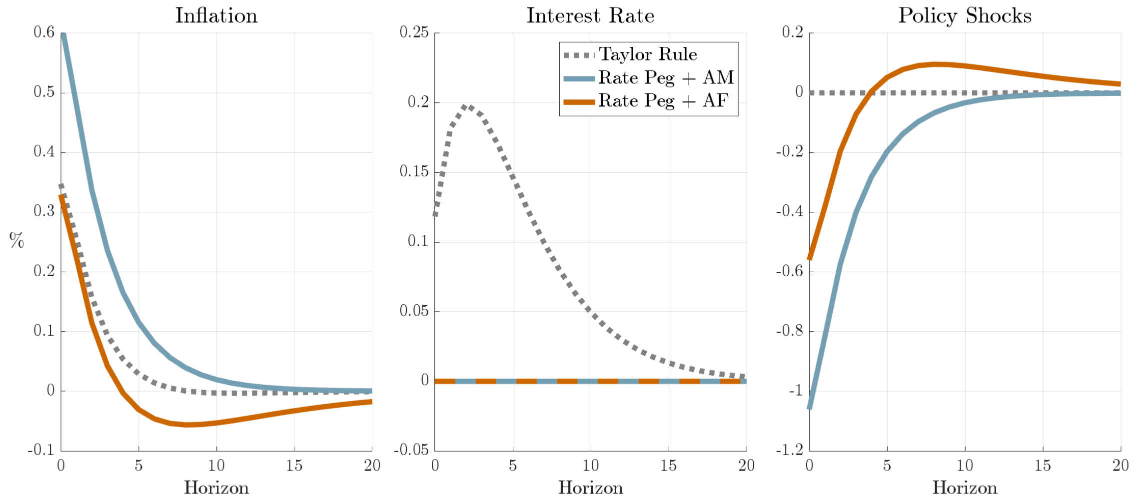

Figure 1: The left and middle panels display inflation and interest rate impulse responses to a cost-push shock in a textbook New Keynesian model, with an active Taylor rule (grey dotted) and under an interest rate peg with equilibrium selection through monetary dominance (active money, blue) or fiscal dominance (active fiscal, orange). The right panel shows the vector of policy shocks to the active Taylor rule that enforces the rate peg, for the two selections.

dominance—i.e., she has estimated the two sufficient statistics, and now seeks to implement our identification results. She then asks how that exact same cost-push shock would propagate under an interest rate peg coupled with either monetary or fiscal dominance.

Results are displayed in Figure 1. The grey dotted lines in the middle and left panels begin by showing impulse responses under an active Taylor-type rule, embedding equilibrium selection through monetary dominance. As is well-known, the hypothesized counterfactual interest rate peg is consistent with multiple equilibria (Leeper, 1991); the blue and orange lines show impulse responses in two such equilibria, with two different assumptions on equilibrium selection. In one case (blue), we select the equilibrium using monetary dominance; equivalently, this corresponds to a switch to an active Taylor rule at a far-ahead but finite horizon, to exact convergence back to steady state in finite time (Angeletos et al., 2026a), or to continuity of the solution as we cross the threshold into indeterminacy (Lubik and Schorfheide, 2004). In the other case (orange), we enforce the present-value government budget constraint under an exogenously fixed (at steady state) primary surplus process, as in the FTPL. We see that, under monetary dominance, not hiking policy rates increases inflation; under fiscal dominance, cumulative (discounted) inflation is instead zero, by the integrated government budget constraint. Our focus, however, is not on those familiar conclusions, but on the third panel: it reveals that *both* of these counterfactual equilibria, monetary and fiscal dominance alike, can be implemented using artificial monetary shocks to the baseline, and active, Taylor rule. The figure thus illustrates our first main new result—that the sufficient statistics of the semi-structural approach, even when constructed in a policy regime corresponding to one kind of equilibrium selection (here monetary dominance), at the same time also allow policy evaluation under a different regime (here fiscal dominance).

Fiscal policy beyond equilibrium selection. In the general notation of Section 2.1, if the policy instrument z is a monetary instrument (e.g., the policy rate), then the underlying fiscal policy rule is instead a part of the "non-policy" block (2). In Corollary 1, fiscal policy was allowed to matter through equilibrium selection (i.e., "fiscal dominance"); beyond that, however, our analysis throughout relied on a stable "non-policy" block (2), and thus tacitly assumed an unchanged underlying fiscal feedback rule.[6]

While the preceding discussion was thus sufficient to speak to changes between monetary and fiscal dominance in the standard equilibrium selection sense, we may alternatively, and more ambitiously, contemplate a *joint* change in monetary and fiscal rules. For example, we may contemplate the effects of a rate hike not just under either monetary or fiscal dominance, but also under different assumptions on the *timing* of any future tax response to that interest rate hike. If households are Ricardian (as in the textbook New Keynesian model) and taxes are lump-sum, then such seemingly more ambitious joint monetary-fiscal policy experiments actually turn out to be equivalent to our preceding discussion, simply because fiscal policy under Ricardian equivalence matters *only* through equilibrium selection (see Angeletos et al., 2026b), with the tax timing irrelevant. If households instead are non-Ricardian, then the timing of taxes does matter, and so now we need to treat taxes as a second argument of the policy instrument vector z; accordingly, evaluating such counterfactuals requires *two* sets of policy causal effects: those of interest rates, just as before, and now also those of changes in the timing of taxes.[7] Given that extended set of causal effects, all of the preceding results again apply; see Angeletos et al. (2025) for a recent contribution that leverages our results to evaluate such joint monetary-fiscal counterfactuals when households are non-Ricardian. We elaborate further on the connection between those joint counterfactuals and the equilibrium

[6]For example, if the initial regime was one of monetary dominance, then that fiscal rule was "passive" in the textbook sense of Leeper (1991). For the experiments in Corollary 1 and Figure 1, when switching to fiscal dominance, we do not change that rule, but instead just alter the boundary condition for equilibrium selection. This boundary selection restricts the present value of taxes (as in the FTPL), but not their path.

[7]If households are Ricardian, then the effects of fiscal shocks—i.e., of changes in the timing of taxes—are zero. This is the precise sense in which fiscal policy here matters *only* through equilibrium selection effects, and explains why in that special case monetary policy causal effects alone suffice.

selection margin of fiscal dominance discussed here in Appendix A.3.

2.3 Stochastic policy rule switching

The previous section considered policy counterfactuals that change assumptions on equilibrium selection, with that regime change, say from monetary to fiscal dominance, assumed to be permanent. Much of the literature on policy regime change over time instead allows for stochastic switches in policy, e.g., between active and passive monetary policy (see Bianchi, 2013; Bianchi and Melosi, 2017, for two well-known examples), or imperfect, stochastic commitment to forward guidance (as in Bodenstein et al., 2012; Haberis et al., 2019), as such stochastic switches are often found to be necessary to account for changes in policy conduct and macroeconomic performance over different historical episodes, notably for the U.S. This section establishes that the exact same sufficient statistics as before also allow evaluation of counterfactuals featuring such stochastic regime switches.

Modeling switching policy rules. Generalizing the counterfactual policy rule (4), we now contemplate a scenario where that counterfactual rule may stochastically switch between N possible versions $s \in \{1, \dots, N\}$, with the transition probability from rule i to j denoted Φ_{ij}. The initial policy regime s_0 is fixed, and so progressing forward in time there are now N^t policy nodes at any given date t. To extend the perfect-foresight framework of Section 2.1 to allow us to speak to such stochastic rule switches, the key conceptual step is to define the sequences $\boldsymbol{x}$ and $\boldsymbol{z}$ not just over time, but instead over all possible histories of rule realizations. In the remainder of this section we focus on the economic logic of why our results continue to apply in such a setting, with details relegated to Appendix A.4.

We will use † superscripts to indicate sequences over all possible histories of policy rule realizations, and contemplate counterfactual policy rules of the form

$$\tilde{\mathcal{A}}_x^\dagger \boldsymbol{x}^\dagger + \tilde{\mathcal{A}}_z^\dagger \boldsymbol{z}^\dagger = \boldsymbol{0}. \tag{9}$$

To aid with interpretation of this relation, consider the case of two possible policy regimes, $s \in \{1, 2\}$. The $\dagger$ sequences in that case contain two entries for date $t = 1$, one corresponding to each policy regime. For example, at $t = 1$, (9) may specify some relation between date-1 endogenous outcomes and the date-1 policy instrument if $s_1 = 1$, and a different relation if $s_1 = 2$, as, say, in a Taylor rule with stochastically switching coefficients. The $\dagger$ sequences at date $t = 2$ would then contain four entries, corresponding to each policy regime history, and so on. The private-sector block, in contrast, continues to be described by (2), still delivering a linearized relation between the current and expected future values of the policy instrument and the endogenous outcomes. The key twist, of course, is that those expectations $\{\boldsymbol{x}, \boldsymbol{z}\}$ are now linked to the corresponding $\{\boldsymbol{x}^\dagger, \boldsymbol{z}^\dagger\}$ sequences across all policy histories in (9) through the transition probabilities Φ_{ij}. Denoting the implied mapping from full histories to date-0 expectations by $\Omega_\bullet$ (e.g., $\boldsymbol{x} = \Omega_x \boldsymbol{x}^\dagger$, see Appendix A.4), we can write the non-policy block as

$$\mathcal{H}_w^\dagger \boldsymbol{w}^\dagger + \mathcal{H}_x^\dagger \boldsymbol{x}^\dagger + \mathcal{H}_z^\dagger \boldsymbol{z}^\dagger + \mathcal{H}_e^\dagger e_0^\dagger = \mathbf{0}, \tag{10}$$

where $\mathcal{H}_\bullet^\dagger = \mathcal{H}_\bullet \Omega_\bullet$. Together, equations (9)–(10) describe the new counterfactual environment with stochastic policy switches, generalized to now be in terms of full state-contingent paths, and not just expected values. This extended setting nests a large class of linearized equilibrium models with stochastic switches in policy feedback rules, e.g., as in Davig and Leeper (2007), Farmer et al. (2009), Bianchi (2013), or Bianchi and Melosi (2017).[8]

Generalized identification result. We assume that the counterfactual system (9)–(10) induces a unique equilibrium, either through the counterfactual policy rule itself, or through additional assumptions on equilibrium selection, e.g., fiscal dominance. We first, as in Proposition 2, study the counterfactual propagation of primitive structural shocks.

Proposition 3. *For each shock $\varepsilon_{i,t}$, knowledge of that shock's impulse responses under the baseline rule, $\Theta_{i,\ell}$, together with policy causal effects Θ_ν, suffices to evaluate counterfactual date-0 expectations under the counterfactual, stochastically switching rule (9).*

We see that, even though we now consider a much richer class of policy counterfactuals, the *exact same* sufficient statistics as before continue to suffice. The key is that, while the policy block is now materially richer, we maintained linearity of the remainder of the model, and so the exact same policy causal effects as before still suffice to evaluate the effects of the (now more complicated) changes in policy. While the formal result is for simplicity stated in terms of counterfactual structural shock impulse responses, we under invertibility again have a one-to-one mapping from primitive shocks to Wold innovations, so we can recover counterfactual historical evolutions and second moments largely as before, subject to some caveats on implementation discussed in Appendix A.4.

[8] Note that, suitably re-interpreted, (9) also covers canonical examples of non-linearity in policy design, e.g., a zero lower bound on the policy rate that binds for a stochastic number of periods. Our illustration in Figure 2 in fact considers such a non-linear example.

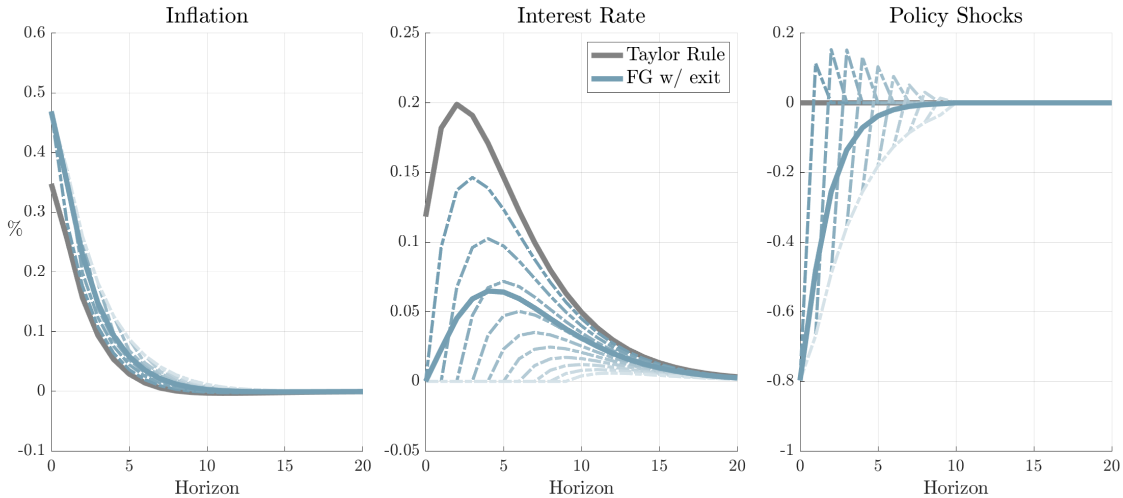

Figure 2: The model environment, shock, and grey lines are as in Figure 1. The blue lines display results for imperfect commitment to an interest rate peg (forward guidance), with stochastic exit to the active Taylor rule. The thick blue lines correspond to date-0 expectations, while the dash-dotted blue lines show expected paths conditional on different realized dates of the switch.

We provide a visual illustration in Figure 2. The environment and shock are identical to Figure 1: a New Keynesian model subject to a cost-push shock, with the grey lines still showing the impulse responses under an active Taylor-type rule. For our policy counterfactual experiment, we contemplate an imperfect, stochastic commitment to a policy rate peg: the policymaker at date 0 promises a rate peg up until date H, but at each date $t = 1, \dots, H-1$ she already abandons the peg with probability $p \in (0, 1)$, a simple reduced-form stand-in for the tension between policy design under discretion vs. commitment (Kydland and Prescott, 1977). Also exactly as before, we consider an econometrician that wishes to evaluate that counterfactual using only knowledge of the causal effects of interest rates, now leveraging our Proposition 3. Results are reported in shades of blue: the thick blue lines show date-0 expectations under the counterfactual policy, for inflation (left panel), the interest rate (middle panel), and the artificial monetary policy shocks that implement the counterfactual (right panel); the thin dotted lines are instead the paths corresponding to different *realized* dates for the switch in monetary policy design, from $t = 1$ to $t = H$, i.e., the different histories in our † notation. The results are as expected, with rates pegged until the date of the switch, and later switches corresponding to more persistently expansionary artificial policy shocks and higher inflation. For our purposes here, the key takeaway is simply the documented versatility of the sufficient statistics of the semi-structural approach: evaluation even of the rich stochastic counterfactuals in Figure 2 still just required the initial cost-push shock projections plus the monetary policy causal effects, and nothing else.

2.4 Taking stock

The upshot of our analysis in this section is that the sufficient statistics of the semi-structural approach to policy evaluation actually go much further than previously believed. The analysis continued to assume linearity as well as time invariance of the private-sector block, and so still just applies only locally around a given steady state; subject to that limitation, however, our results speak to rich, much-studied forms of policy regime change. In practice, this means that the semi-structural toolkit can now be applied to several questions of monetary policy evaluation that so far had been the purview of fully structural analyses.

The remaining hurdle, of course, is that the informational requirements are still material. We deal with that second, practical challenge in the next two sections.

3 Getting the sufficient statistics

In this section we begin with a general methodological discussion of how empirics and modeling structure can be used to recover the two sufficient statistics: reduced-form projections in Section 3.1, and policy causal effects in Section 3.2. The actual implementation will then follow in Section 4, before we turn to applications in Section 5.

3.1 Invertibility and forecasts

The assumption of invertibility underlying the identification results of Section 2 is undoubtedly strong, and has been routinely criticized in the recent literature on estimating macroeconomic shock propagation (e.g., see Plagborg-Møller and Wolf, 2022). We will argue that, while violations of that assumption can cause serious problems for shock analysis, they are much less of an issue for the purpose of policy counterfactual evaluation.

The role of invertibility. In the literature on macro shock identification, the role of invertibility is to ensure that the shock of interest can be obtained as a linear combination of reduced-form Wold innovations. In the preceding identification results, invertibility instead plays a very different role: it ensures that forecasts based on the econometrician information set—the history of observed macroeconomic aggregates—are equal to full-information, "oracle" forecasts (Barnichon and Mesters, 2023). Intuitively, once that is the case, we can use the knowledge of policy causal effects to *solely* change assumptions on policy, otherwise fixing the underlying state of the economy. We formalize this intuition in Appendix A.6, where we establish that the Wold representation of the observables y_t suffices for policy evaluation *if and only if* the forecasts it implies equal full-information forecasts.[9]

The key implication of the preceding discussion is that, in the context of the identification results of Section 2, invertibility is sufficient, but not necessary. Formally, even significant violations of invertibility may not threaten our ability to arrive at correct counterfactuals—what matters is that projections are formed with respect to an information set that does not omit any relevant predictors of the macroeconomic outcomes of interest, and thus correctly captures the date-t information relevant for the future evolution of the macro-economy. We next illustrate this theoretical observation with model-based simulations.

Illustration. We use a familiar quantitative model—that of Smets and Wouters (2007)—as a laboratory. We then consider a researcher that wishes to predict the second-moment properties of output, inflation, and nominal interest rates in this model economy under an alternative monetary rule that puts a larger weight on output stabilization. Since the true data-generating process is unknown to the researcher, she leverages the preceding identification results. Consistent with the focus of this subsection, we assume that she does know the true matrix of interest rate causal effects Θ_ν, but then relies on information sets $\{y_{t-\ell}\}_{\ell=0}^{\infty}$ that are (potentially) insufficient to deliver invertibility. Specifically, we consider two possible information sets: interest rates, output, and inflation ("small"); and then a full information set that also contains hours worked, wages, investment as well as consumption. Importantly, the latter information set satisfies invertibility, while the former one does not.

The headline result, visible in Figure 3, 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 rule, while the other lines indicate the true (solid) and predicted (dotted) counterfactual spectral densities. Consistent with our preceding theoretical results, the predicted counterfactual using the large information set equals the truth. More importantly, even the counterfactual predicted using the much smaller information set is very close to the truth. The explanation follows from our

[9]As discussed in Appendix A.7, our identification results *are* consistent with various kinds of frictions in private-sector expectation formation; however, the forecasts that the econometrician leveraging our identification results constructs still always need to be full-information forecasts.

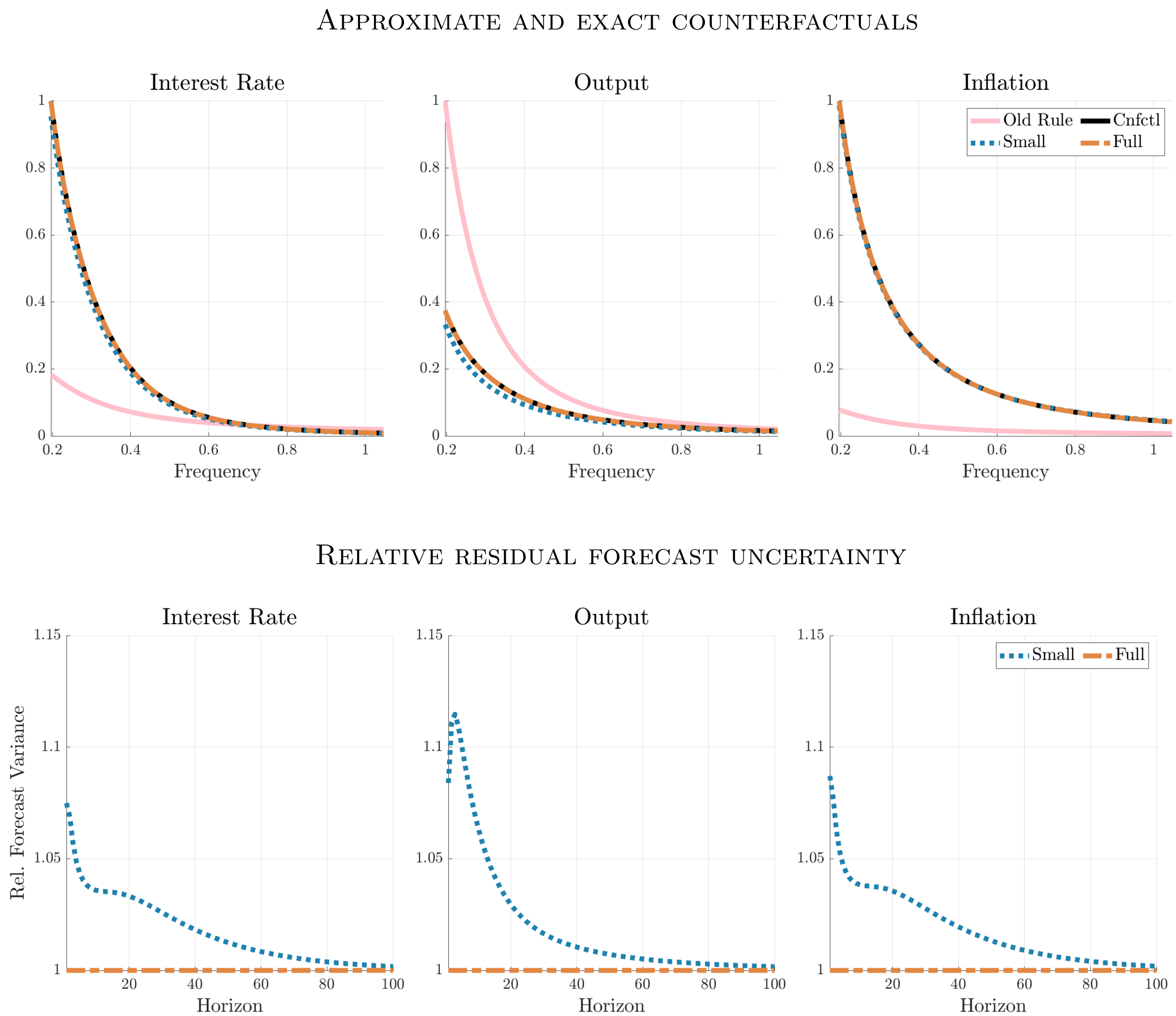

Figure 3: Top panel: business-cycle spectral densities for interest rates, output, and inflation under the old rule (solid pink) and under the counterfactual rule, true (solid black) and predicted using Proposition 1 for different information sets (dotted blue, dotted-dashed orange). We normalize the 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.

forecasting intuition given above, and is illustrated by the bottom panel. That panel shows residual forecast uncertainty for interest rates, output, and inflation at different horizons (x-axis), and for our two different information sets (different lines), with the residual forecast uncertainty under the full information set normalized to 1 at each horizon h. As $h \to \infty$, the residual forecast variances for both information sets of course limit to the same number—the unconditional variance. For intermediate h, forecasting uncertainty is instead strictly larger for the smaller information set. The differences, however, are moderate, with forecast variances that are only at most around 12 percent larger than those for the full information set. Even a much smaller information set can thus deliver accurate forecasts, and thereby also accurate counterfactuals.

Since the Smets and Wouters model features seven shocks, the previous discussion substantiates our point that the invertibility assumption plays a fundamentally different role here than in the literature on macro shock identification as in the structural VAR tradition. With seven shocks and three observables we are necessarily quite far from invertibility, and no rotation of the three-dimensional reduced-form innovations u_t will yield any of the individual structural shocks. Forecasting, however, is an easier task. Even with the much smaller information set, the forecasts are already close to full-information ones, and so, consistent with the theory discussed above, the counterfactuals are highly accurate.

Practical takeaways. The upshot of the preceding discussion is straightforward: for policy evaluation, the sole role of invertibility is to ensure a good summary of the relevant date-t state of the economy. A simple check is to ensure that any reported counterfactual conclusions are robust to reduced-form projections formed with respect to larger information sets. As long as no important predictors for any of the outcomes of interest are omitted, results even for projections with rather small information sets will be accurate.

3.2 Policy causal effects

We now turn to the second required input, policy causal effects Θ_ν.

Empirical evidence. A first natural strategy to learn about policy causal effects is regressions on policy shocks, or instrumental variables for such shocks. Interpreting the output of such regressions through the lens of the identification theory of Section 2 is straightforward: valid policy instrumental variables are correlates of contemporaneous or news policy shocks $\boldsymbol{\nu}$, and a regression on a single such instrument delivers a one-dimensional subspace of the overall policy causal effect matrix Θ_ν.[10] Note furthermore that, since we only care about the overall space of policy causal effects, whether any given empirically identified shock or

[10] More formally, a valid instrumental variable s_t is a series that correlates with *any combination* of current or news policy shocks, and not with any non-policy shocks. Projection on this instrument series then recovers the corresponding weighted average treatment effect of the particular mix of policy shocks that correlate with s_t, by standard arguments (e.g., see Plagborg-Møller and Wolf, 2022, Appendix B.3).

instrument is labeled as a "contemporaneous" or "forward guidance" disturbance is immaterial; rather, all that matters is the current and expected future path of the policy rate that this shock induces, i.e., the implied change in the yield curve. We thus simply directly focus on the implied interest rate path as the policy "treatment."

In practice, the available finite empirical evidence will never afford a complete characterization of the infinite-dimensional space of policy causal effects. Notably, as we will show in Section 4.1, the available evidence on monetary policy shocks informs us about the effects of transitory rate changes, but is largely silent on the propagation of expectations about far-ahead interest rates. To date, the semi-structural literature has responded to this challenge by restricting attention to counterfactuals that are at least approximately spanned by the available evidence (see McKay and Wolf, 2023).

Model structure. Alternatively, policy transmission can be pinned down through explicit model structure. Using the notation of Section 2, the overall space of allocations implementable by policy is completely pinned down by the non-policy model block $\{\mathcal{H}_w, \mathcal{H}_x, \mathcal{H}_z\}$. In words, the part of the model structure that governs policy propagation is the model's non-policy relations, but *not* the structural shocks to those relations $\{\mathcal{H}_e e_0\}$, nor the policy rule $\{\mathcal{A}_x, \mathcal{A}_z\}$ itself. The intuition echoes insights from the analysis of simple static models of demand and supply: to predict to first order the effect on prices of the quantity demanded changing to some hypothetical level, all we need to know is (i) the initial outcomes plus (ii) the slope of the supply function, and *not* the slope of demand, nor the actual shocks to demand or supply that generated the observed outcomes. In our setting, we similarly require initial outcomes (here the reduced-form projections) plus the non-policy block (to deliver the causal effects of any possible policy change), but neither the policy rule nor the history of true policy and non-policy shocks.[11]

A hybrid strategy. Putting the discussions on empirical evidence and model structure together suggests a simple extrapolation strategy to get Θ_ν: first specify a model of policy transmission (i.e., structure that delivers the tuple $\{\mathcal{H}_w, \mathcal{H}_x, \mathcal{H}_z\}$), discipline it by targeting the available empirical evidence on policy shocks, and finally use the model to deliver the full matrix Θ_ν. If the fit to those empirical targets is good, then extrapolation beyond the

[11]Examples of prior work that use explicit model structure to deliver the policy causal effects Θ_ν include De Groot et al. (2021) and Hebden and Winkler (2021). Our proposed hybrid strategy, described next, sits in between that prior work and the pure semi-structural approach.

support of the data is indeed the sole residual role of model structure. We note that such a hybrid strategy is a combination of empirical sufficient statistics and structural approaches in the same spirit as discussed in Chetty (2009, p. 455), just there in the context of public finance: theory is used to identify the relevant sufficient statistics, they are measured from data whenever possible, and finally the empirics are supplemented with the minimal amount of model-based extrapolation when necessary.

3.3 A discussion of the proposed approach

We close this section with a brief discussion of how the proposed hybrid strategy relates to prior work in both the semi- and fully structural traditions.

VS. THE SEMI-STRUCTURAL APPROACH. Our strategy is an extension of the pure semi-structural approach (following McKay and Wolf, 2023; Barnichon and Mesters, 2023). The two methodologies are built around the exact same set of "sufficient statistics," with our results in Section 2 further pushing the boundary of what those statistics can deliver. Turning to the statistics themselves, we do not innovate relative to prior work on the reduced-form projections; our only contribution is to argue that even projections with respect to relatively small information sets can suffice. The material difference, on the other hand, is for policy causal effects: while the pure semi-structural strategy uses *only* the empirical evidence on policy shock transmission, we instead propose to supplement it with further explicit model structure to deliver the remainder of Θ_ν. This has two implications. First, if a given policy counterfactual question can be evaluated using the already available empirical evidence alone, then the two strategies agree.[12] Second, given all of Θ_ν, we can evaluate *any* counterfactual that is covered by the identification results in Section 2. The pure semi-structural approach, in contrast, can only ever construct a *best approximation* to the contemplated counterfactual, which in any given application may or may not be accurate.

VS. THE LUCAS PROGRAM. The standard approach to policy counterfactual evaluation in the Lucas program is to rely on complete structural models estimated using full-information, likelihood-based methods (e.g., Smets and Wouters, 2007; Bocola et al., 2024, among many others). A frequent criticism of such strategies is the likelihood of model misspecification, in particular in terms of the model's primitive shocks, often labeled "dubiously structural"

[12]This statement implicitly supposes that the model used for causal effect extrapolation can match the available empirical evidence. We make sure that this fit is reasonably good in our empirical applications.

(Chari et al., 2009). An important takeaway from Section 2 is that, for a quite wide array of possible counterfactual questions, such misspecification may simply not matter: as long as the model matches reduced-form second moments of the data (and thus delivers accurate projections) and implies empirically sensible policy causal effects, the reported conclusions are necessarily correct, *even if* the model is otherwise arbitrarily misspecified.

Our proposed hybrid strategy for recovering Θ_ν, in contrast, relies on model structure not via such a full-information approach, but instead through the limited-information strategy of impulse response matching (as in Christiano et al., 2005, 2010, among many others). This more targeted use of model structure, with its sole purpose being policy causal effect extrapolation beyond the support of the data, confers three advantages. First, by construction, every component of the two sufficient statistics that can be measured directly—both the projections and at least some policy causal effects—is indeed taken from the data. Second, there is a transparency benefit: for any given policy counterfactual question, the researcher can separate the role played by the data from that of the model-based causal effect extrapolation. And third, there is a more practical concern with full-information methods: forcing any given model to match reduced-form second moments—which is what happens in standard likelihood-based estimation—may distort inference on policy causal effects, unless the *entirety* of the model is well-specified; see Appendix B.3 for further discussion.

The practical upshot for the wider Lucas program is that, once combined with the identification results presented here, limited-information strategies like impulse response matching are actually even more useful than previously believed, allowing evaluation of a large range of possible policy counterfactuals in spite of the fact that they only deliver a partial description of the economy. Intuitively, that partial description actually yields precisely the part of the sufficient statistics that the empirical evidence (so far) does not.

4 Monetary policy causal effects

We now transition to applications of the preceding methodological results. In this section, we implement the hybrid empirical-structural approach of Section 3.2 to obtain the required monetary policy causal effects Θ_ν.

4.1 Empirical evidence

The literature on monetary policy shock propagation is voluminous. In principle, the different shocks studied in that literature could lead to heterogeneous policy "treatments"—some of

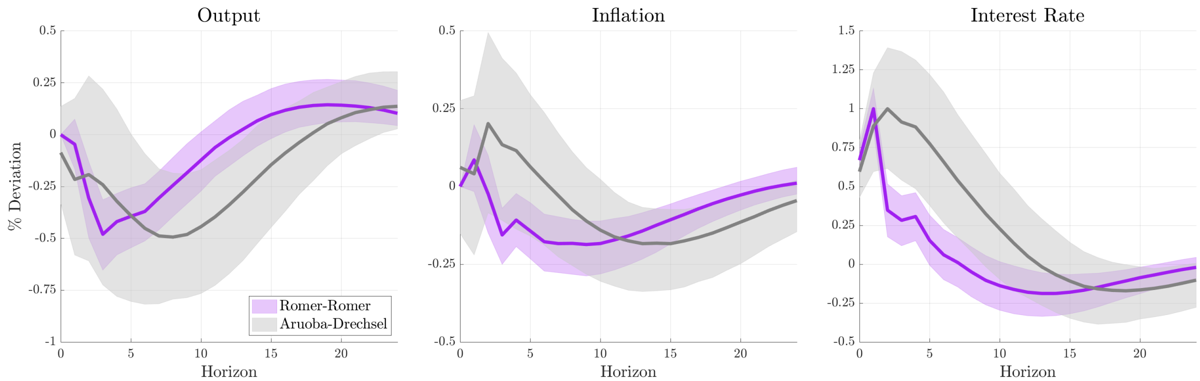

Figure 4: The purple and grey lines and shaded areas indicate impulse responses (with 16th and 84th percentile confidence bands) for the monetary policy shocks of Romer and Romer (2004) and Aruoba and Drechsel (2026). For estimation details please see Appendix C.1.

those monetary policy events may primarily revise the short end of the yield curve, i.e., lead to short-lived policy rate changes, while others may instead revise the medium or long end. Through the lens of our theory, such heterogeneity would be very useful, allowing empirical estimation of most of the policy causal effect matrix Θ_ν.

In practice, however, it turns out that most of the events we use to learn about monetary shock transmission tend to only induce rather short-lived changes of the policy rate. A paper well-suited to illustrate this observation is Swanson (2024): similarly to Gürkaynak et al. (2005), Swanson isolates high-frequency "forward guidance" shocks that alter the yield curve without affecting the short-term policy rate, and explicitly distinguishes those from "federal funds rate" shocks. In spite of this difference in the construction, however, projections of the policy rate on those two shocks only show moderate differences, with the "forward guidance" disturbances resulting in only somewhat more persistent fluctuations in policy.[13] Many of the other popular monetary shock series (including Romer and Romer, 2004; Gertler and Karadi, 2015; Jarociński and Karadi, 2020; Miranda-Agrippino and Ricco, 2021; Aruoba and Drechsel, 2026) similarly differ somewhat in the precise timing of the induced rate paths, but all broadly move the short to medium end of the yield curve.

Consistent with the previous discussion we will, for the remainder of this paper, take as our empirical reference point the policy causal effects in Figure 4, showing the propagation of

[13] In making this statement we are comparing Figures 2 and 3 in Swanson (2024). Similar conclusions on the relatively small differences between interest rate paths induced by contemporaneous and forward guidance shocks are reported in Bundick and Smith (2020). Also see D'Amico and King (2023) and Christoffel et al. (2020) for related evidence on the difficulties of credibly estimating the effects of anticipated changes in future monetary policy.

two shocks that induce relatively transitory policy rate movements. The impulse responses in the three panels are estimated through projection on the classic series of Romer and Romer (2004) as well as its more recent refinement in Aruoba and Drechsel (2026); estimation details are provided in Appendix C.1. We take these two series as representative examples because they rely on similar approaches to monetary policy shock identification, yet somewhat differ in the induced policy rate time profiles, with the series of Aruoba and Drechsel leading to a moderately more delayed interest rate change, consistent with greater policy gradualism over the past two decades.[14] The two associated policy rate "treatment" paths, as displayed in the right panel, then lead to moderately persistent falls in the level of output (left panel) as well as a more delayed fall in the inflation rate (middle panel); as expected, since the Romer and Romer treatment is more transitory, the corresponding output and inflation impulse responses are also more short-lived. Both qualitatively and quantitatively these patterns are in line with those estimated in other empirical work that uses different monetary policy shock series, including in particular the various other studies cited above.

With empirical evidence insufficient to deliver the full infinite-dimensional space of monetary policy causal effects, we next turn to model structure to extrapolate beyond the observed support of the data.

4.2 Structural models of monetary transmission

We consider multiple different models of monetary policy transmission. Doing so allows us to understand how particular model assumptions shape the extrapolation, and thus what model features are most important to substitute for missing data variation.

This subsection begins by sketching the models, with results on causal effect extrapolation following afterwards. Our first model is a relatively standard representative-agent model with nominal rigidities ("RANK"), augmented with several other frictions to allow a quantitative fit to the monetary shock 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 two models, with price- and wage-setters forming expectations with cognitive discounting, as in Gabaix (2020). Throughout

[14]This difference in the rate path presumably reflects changes in the conduct of monetary policy. Through the lens of our identification results, such a change provides no particular challenge—the key maintained assumption is stability of monetary policy causal effects, and thus of the remainder of the economy.

we provide brief descriptions, with details in Appendix C.2. We do so because the models we consider are standard; our contribution lies in how we *use* these models for extrapolation.

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 evidence on the transmission of transitory policy rate changes discussed above. Following Christiano et al. (2005), 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 (just 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 are monetary and fiscal authorities. The fiscal authority issues nominal bonds with exponential maturity structure, spends a constant amount in real terms, and then adjusts lump-sum taxes gradually to maintain long-run budget balance. Recall that fiscal policy here matters only through equilibrium selection.

  The monetary authority sets nominal interest rates. As discussed in Section 3.2 (and further in Appendix B.2), we do not need to specify any particular policy rule.

Stacking all model blocks except the behavior of the monetary authority, we then obtain $\{\mathcal{H}_w, \mathcal{H}_x, \mathcal{H}_z\}$—i.e., the "non-policy" block (2). Solving the system for every possible path of monetary policy shocks and thus of nominal interest rates, we obtain the monetary policy causal effects, Θ_ν. We estimate the model to ensure consistency between Θ_ν and the short-end estimated effects of Section 4.1—i.e., a standard impulse response matching strategy, as discussed in Section 3.2.[15] Details on the set of estimated parameters and on the choice of priors are provided in Appendix C.3.

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, 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.

Since consumers in the HANK model are non-Ricardian, the timing of taxes and transfers invariably shapes equilibrium outcomes. Thus, in joint monetary-fiscal policy counterfactuals (as discussed at the end of Section 2.2), fiscal policy matters not just via equilibrium selection, but also through the on-equilibrium effects of transfers. While the core use of model structure in our monetary policy-centric analysis is extrapolation of interest rate causal effects, we note that the model as a by-product also yields the effects of any tax path, thereby immediately allowing the evaluation of such joint counterfactuals. We will leverage this in Section 5.1.

Behavioral frictions. 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.

[15] In the standard impulse response matching literature, the researcher writes down a policy rule, and then restricts attention to contemporaneous shocks to that rule. Commitment to a rule, however, is actually not necessary—one can simply find the best fit to the empirical impulse response targets within the overall *space* implementable by policy, sidestepping the need to commit to any rule; see Appendix B.

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-weighted 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, and yet still such that the resulting behavioral models can provide an adequate fit to our empirical monetary shock estimation targets.

4.3 Causal effect extrapolation

We now turn to causal effect extrapolation. As we show in Appendix C.3, all four models are able to match the two empirical estimation targets reasonably well, and particularly so for the shock of Romer and Romer (2004). This is of course unsurprising: our baseline RANK and HANK models are very similar to Christiano et al. (2005) and Auclert et al. (2020), respectively, who conduct similar estimation exercises; and behavioral frictions matter little for the transitory policy treatments that we consider here. The key question is now how those four models extrapolate beyond the support of the data. Results are displayed in Figures 5 and 6, which show the model-implied impulse responses to "forward guidance" shocks—i.e., nominal interest rate movements at the long end of the yield curve.

Rank vs. Hank. We begin by comparing, in baseline RANK and HANK, a monetary policy intervention much farther out on the yield curve: a deviation from a standard monetary policy rule that is announced at $t = 0$ but occurs four years later. Results are in Figure 5.

The main takeaway is that the two models, which by construction closely agree on the effects of the two targeted interest rate paths, also at least approximately agree on the effects of this much more delayed monetary intervention. In particular we see that, for both output and inflation, posterior medians are close and there is heavy overlap of the posterior bands, with the peak output response of forward guidance perhaps slightly larger in HANK, echoing Auclert et al. (2020). The theoretical underpinning for this finding was provided in Werning (2015), who gives sufficient conditions for the consumption block of an incomplete-markets model to aggregate to an "as-if" Euler equation, delivering exact equivalence with representative-agent benchmarks. Intuitively, by forcing the two models to agree on the effects of transitory policy rate changes, we force heterogeneity-related channels, which may amplify or dampen monetary transmission, to largely offset, for both contemporaneous and far-ahead rate changes. This discussion in particular reveals that differences in transmission

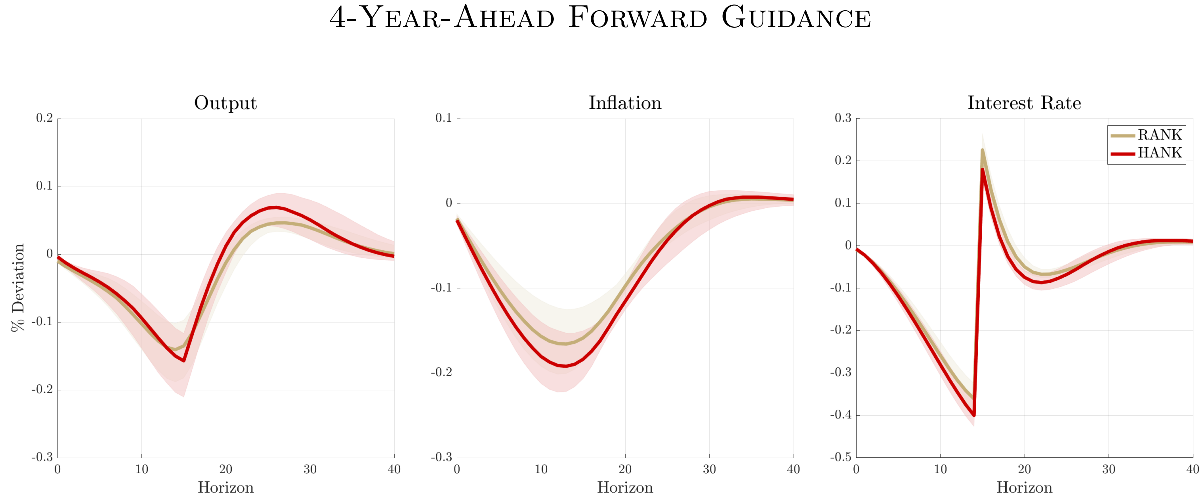

Figure 5: Policy causal effect extrapolation in the estimated RANK and HANK models. The figure shows output and inflation impulse responses to a forward guidance policy shock that leads to the interest rate movements depicted on the right. Solid lines show the posterior medians. The shaded areas show the 16th and 84th percentile posterior bands.

channels—e.g., the contrast between direct vs. indirect propagation discussed by Kaplan et al. (2018)—may not matter for counterfactual aggregate outcomes. We note, of course, that this approximate equivalence does not extend to distributional outcomes.

Baseline vs. cognitive discounting. Figure 6 repeats the same exercise, but now instead comparing the baseline and behavioral representative-agent models.

We see that now there are more material differences. In particular, and just as expected given the additional cognitive discounting among price-setters of the behavioral models, we see that the inflation responses in the behavioral model are more delayed, and the subsequent output overshooting is more pronounced. While Figure 6 shows this for the representative-agent models, we note that the same is also true for the heterogeneous-agent model variants. It follows that, for monetary policy counterfactuals that involve persistent rate changes, whether to extrapolate through models with or without behavioral frictions matters greatly.

Aside: monetary and fiscal dominance. Note that all of the above monetary policy causal effects are, for all models, conditional on the assumed passive fiscal policy rule and in particular embed the corresponding "active-money" equilibrium selection. Proceeding as discussed in Section 2.2, we can and will however use those causal effects to also evaluate counterfactual outcomes under fiscal dominance, incorporating both the equilibrium selection

4-Year-Ahead Forward Guidance

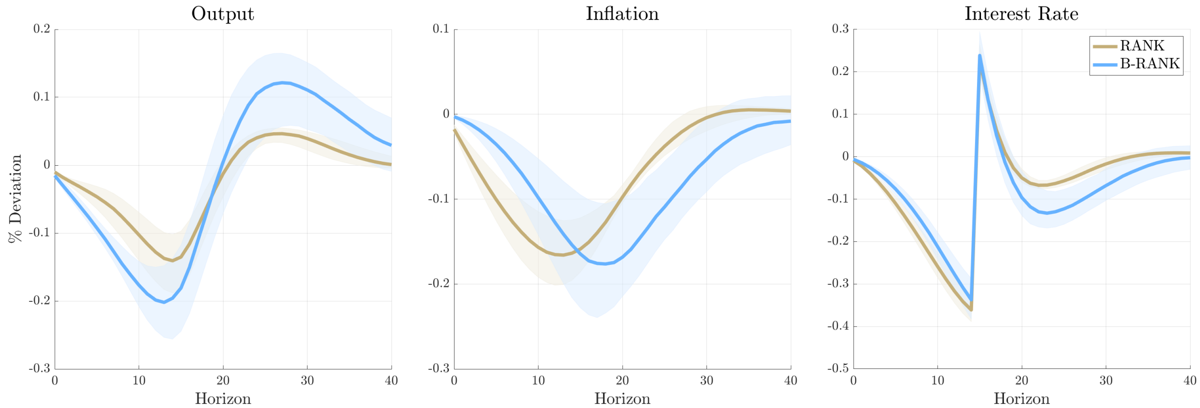

Figure 6: Policy causal effect extrapolation in the estimated RANK and behavioral-RANK (B-RANK) models. The figure shows output and inflation impulse responses to a forward guidance policy shock that leads to the interest rate movements depicted on the right. Solid lines show the posterior medians. The shaded areas show the 16th and 84th percentile posterior bands.

margin and, if applicable (for HANK), the on-equilibrium effects of taxes and transfers.

5 Applications

We now leverage our extended identification results and proposed approach to getting the required sufficient statistics to evaluate several counterfactual monetary policy scenarios.

5.1 Post-2021 inflation

For our first experiment we revisit the post-2021 U.S. inflation surge. While the Federal Reserve ultimately responded with a series of interest rate hikes, the monetary policy response was slow relative to the predictions of simple Taylor-type rules. We thus use our approach to evaluate counterfactual inflation outcomes under adherence to such a Taylor rule; in particular, leveraging our analysis in Section 2.2, we focus on the role of the underlying fiscal backing, comparing outcomes under monetary and fiscal dominance.

The policy experiment. We assume that the monetary authority follows the rule

$$i_t = \phi_\pi \pi_t, \tag{11}$$

where i_t denotes the policy rate and π_t is inflation, both expressed in deviation from steady state, and we set $\phi_\pi = 1.5$. This is a textbook active Taylor rule, and the typical assumption in monetary economics is to suppose that this active monetary rule is accommodated through "passive" fiscal adjustment. This is the policy configuration that we consider first.

Our second policy configuration is motivated by the observed fiscal policy conduct post-2021. The policy response to the Covid-19 pandemic featured significant fiscal outlays, and *ex post* these outlays were partially financed by an increase in inflation, raising the question of whether there was a switch to a fiscally-led equilibrium (see Barro and Bianchi, 2026, among others). We thus, as our second counterfactual, fix the policy rate path implied by the first Taylor rule counterfactual, but switch equilibrium selection to fiscal dominance; in other words, we contemplate a time-varying interest rate peg, with an increase in interest rates not backed by any commensurate current or future tax surpluses. To implement this switch, we follow Proposition 2 and Corollary 1 and use for equilibrium selection the requirement that, in the cumulated government budget constraint, the contemplated interest rate hike is financed not through any changes in taxes, but instead through the endogenous equilibrium paths of the other arguments of the government budget, including in particular the path of inflation—i.e., we rely on the FTPL's equilibrium selection.[16] We provide more details on how exactly we evaluate those "fiscal dominance" counterfactuals in Appendix D.2.

Getting the sufficient statistics. We start in the midst of the post-Covid inflation, in 2021:Q2, and then evaluate counterfactual projections for future policy and macroeconomic outcomes under our two different assumptions on monetary policy conduct and on the policy regime. The first required input is thus forecasts from 2021:Q2 onwards, and we construct those forecasts using a 10-variable reduced-form VAR. Further details on the VAR estimation are provided in Appendix D.1; we there document that the forecasts implied by our reduced-form VAR are competitive with other forecasting strategies.

For monetary policy causal effects, we here chiefly rely on those implied by the baseline rational-expectations RANK model. We use those causal effects to enforce the counterfactual Taylor rule (11), fix the implied path of the policy rate, and then switch to equilibrium selection through fiscal dominance. For comparison, we also briefly consider what happens

[16] Our equilibrium selection thus forces the present value of taxes to be zero; in RANK, that present value is all that matters. When we instead use HANK-implied policy causal effects, we further adjust the fiscal policy to ensure that the tax path is zero date-by-date, as the arguably more empirically relevant scenario. The model thus here provides monetary policy causal effect extrapolation both across the yield curve and for different on-equilibrium tax paths.

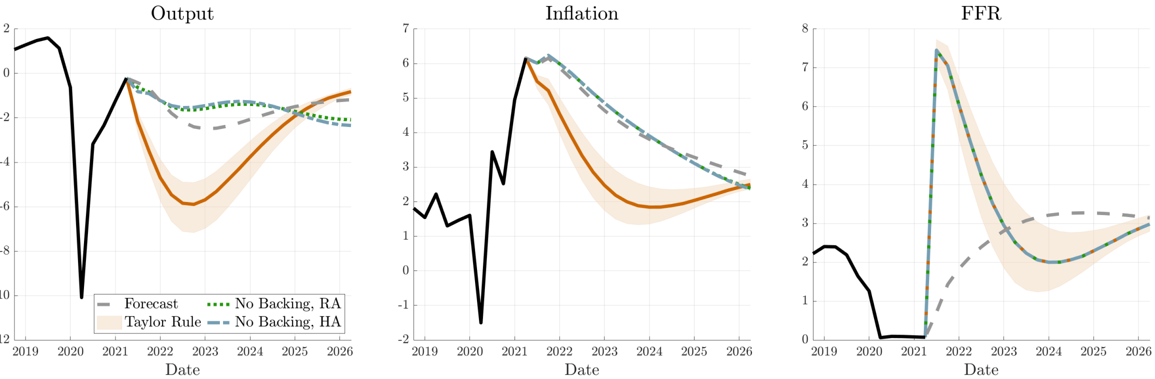

Figure 7: Counterfactual projections of output, inflation, and the federal funds rate in the post-Covid inflationary episode, under a Taylor rule-implied interest rate path and with different assumptions on fiscal backing and monetary policy causal effects. Black: data. Grey: actual forecast. Colored: posterior median (solid) and 16th and 84th percentile bands (shaded) for RANK causal effects with fiscal backing (monetary dominance, orange) and without backing (fiscal dominance, green), and for HANK without fiscal backing (blue).

if instead we use the causal effects of baseline HANK, again with fiscal dominance selection.

Results. Our main results are displayed in Figure 7. We see that inflation was very elevated in 2021:Q2, and expected to remain elevated for several years. The orange lines now show what happens under a counterfactual switch to the aggressive Taylor rule (11), assuming monetary dominance and thus necessarily a commensurate fiscal tax backing. The results are exactly as expected given our empirical monetary shock targets, which implicitly embed a relatively flat Phillips curve: aggressive policy rate hikes materially dampen inflation, but at the cost of a rather large drop in output, echoing the findings of Barnichon (2022). Results using the monetary policy causal effects implied by either HANK, or instead by either of the two behavioral models, are similar: in the former case, this is so because *all* policy rate causal effects in RANK and HANK are quite similar; in the latter case, it follows because the contemplated rate hike is relatively transitory.

We next switch to the counterfactual without fiscal backing, i.e., we change the equilibrium selection to fiscal dominance (in green for RANK, in blue for HANK). The picture flips, with the exact same sharp interest rate hikes as before now unable to materially stabilize inflation, and at the same time also not depressing output. The simple intuition is the "stepping on a rake" mechanism of Sims (2011) in action: because of their direct impact on the government budget, coupled with the absence of any kind of fiscal adjustment, policy rate hikes are now not suitable as a tool to stabilize inflation. We furthermore see that the fiscal dominance results are actually extremely similar for RANK and HANK: in the latter, the timing of taxes and transfers also matters, but ultimately the most important force shaping the displayed equilibrium paths for output and inflation is the required intertemporal government budget balance (i.e., the present-value constraint enforced by the equilibrium selection margin). These results are consistent with the discussion in Angeletos et al. (2026a), who argue that inflationary outcomes in the textbook FTPL and in HANK models with slow or absent fiscal adjustment are likely to be very similar.[17]

Discussion. Our first application illustrates sharply how any evaluation of Federal Reserve actions post-2021 is incomplete without a discussion of the fiscal framework, and in particular without the associated statements about equilibrium selection. Beyond this qualitative point, all that is required to arrive at the particular quantitative answers we provided is monetary policy causal effects—by our extended identification results, those suffice even for a researcher contemplating meaningful policy regime change in the sense of equilibrium selection through either monetary or fiscal dominance.

5.2 Average business cycle

For our second counterfactual we ask whether an optimal flexible inflation-targeting monetary policymaker could have materially reduced macroeconomic volatility over a long U.S. post-war sample. Note that this is an example of a counterfactual for average cyclical fluctuations rather than for any given particular historical episode.

The policy experiment. We will consider as our counterfactual monetary policy rule the one that minimizes the following standard "dual mandate" central bank objective:

$$\mathcal{L} = \mathbb{E}_0 \left[\sum_{t=0}^{\infty} \beta^t \left\{ \lambda_\pi \pi_t^2 + \lambda_y y_t^2 + \lambda_i (i_t - i_{t-1})^2 \right\} \right], \tag{12}$$

where π_t is inflation, y_t is the output gap, and i_t is the nominal interest rate, all in deviations from steady state. Given the monetary policy causal effects Θ_ν, we set the counterfactual monetary policy rule coefficients $\{\tilde{\mathcal{A}}_x, \tilde{\mathcal{A}}_z\}$ to yield the optimal implicit forecast targeting

[17] There is, however, a fundamental difference between RANK and HANK, as also stressed by Angeletos et al. (2026a): in the former, we obtain the monetary dominance outcomes (displayed in orange) if fiscal adjustment occurs at *any finite horizon*; in the latter, we instead smoothly converge to the blue outcomes as fiscal adjustment becomes more gradual, even if we remain in a regime of monetary dominance.

rule corresponding to (12), using the exact same expressions as in Proposition 2 of McKay and Wolf (2023).[18] We study this particular counterfactual exercise because it closely mirrors the popular strategy of flexible inflation targeting (see Svensson, 1997), and as such appears in much central bank communication (e.g., Federal Reserve Tealbook, 2016). Consistent with the discussion there we consider an equal-weights parameterization across the three objectives and over time, with $\lambda_\pi = \lambda_y = \lambda_i = 1$ and no discounting, i.e., we set $\beta \to 1$. We will throughout assume that this monetary policy counterfactual is backed by a passive fiscal authority, i.e., we are in a regime of monetary dominance.

Getting the sufficient statistics. In terms of reduced-form projections, we require the autocovariance function of the three core variables of interest, as well as other aggregates that are useful to predict them. We recover those objects similarly to how we constructed forecasts in Section 5.1, estimating a reduced-form VAR on a large vector of macroeconomic observables, and then translating this VAR to the corresponding Wold lag polynomial. To gauge the importance of the invertibility assumption and further illustrate our earlier theoretical discussion, we will furthermore also report results based on projections that instead use a much smaller information set, containing only the three variables of interest themselves. Further details on the estimation are again provided in Appendix D.1.

For monetary policy causal effects, we throughout assume our benchmark passive fiscal backing (i.e., we assume monetary dominance). We will compare results for monetary policy causal effects extrapolated using all four of our candidate models.

Results. Our headline finding is that the counterfactual policy would have achieved quite materially lower output volatilities and somewhat more stable inflation, at the cost of only moderately more volatile interest rates.[19] Results are summarized in Table 5.1, which reports actual and counterfactual volatilities of output, inflation, and the federal funds rate. The top row first of all depicts the actual in-sample volatilities. The other lines then report the counterfactuals; for now we focus on the results from our proposed "hybrid" approach, with policy causal effect extrapolation through the four models of Section 4.2. The main finding is

[18] As discussed in Section 3.3, to transparently decompose the respective roles of direct empirical evidence and model-implied causal effect extrapolation, we will also report results for the approximate counterfactual that instead *only* uses the empirical evidence of Section 4.1. In words, this counterfactual minimizes the loss (12) within the empirically estimated space of (short-end) monetary policy shock causal effects.

[19] Note that this output volatility reduction does not just reflect infeasible rate cuts during the period of a binding lower bound on nominal interest rates; see Appendix D.3 for pre-2007 results.

	Inflation	Output	Interest Rate
Actual	1.978	3.050	2.927
Counterfactual			
Hybrid			
RANK	1.661	1.687	3.792
	(1.604, 1.743)	(1.511, 1.966)	(3.450, 4.133)
HANK	1.587	1.652	3.640
	(1.554, 1.681)	(1.481, 1.958)	(3.323, 3.945)
B-RANK	1.647	1.613	3.688
	(1.540, 1.850)	(1.362, 1.859)	(3.279, 4.086)
B-HANK	1.631	1.605	3.663
	(1.551, 1.840)	(1.364, 1.859)	(3.282, 4.038)
Semi-structural	1.797	2.173	2.838
	(1.618, 1.917)	(1.910, 2.423)	(2.512, 3.172)

Table 5.1: Actual as well as counterfactual unconditional volatilities of inflation, output, and the federal funds rate, with counterfactuals for the policy rule that minimizes (12). Posterior median with 16th and 84th percentiles in parentheses. Projections via a large information set.

that, for output (and, to a lesser extent, for inflation), the predicted counterfactual volatility is far below the actual in-sample volatility.

Our second main result is that this possibility of lower output volatility is chiefly driven by the empirically estimated short-end causal effects of monetary policy, and less so by the model-based extrapolation to the long end of the yield curve. There are two ways of seeing this. First, in Table 5.1, the four hybrid method posterior volatilities are all quite close, even though we know that the behavioral and non-behavioral models extrapolate very differently. Second, in the last row (labeled "Semi-structural"), we repeat our analysis using *only* the empirically estimated short-end monetary policy causal effects, rather than the entire, model-extrapolated matrix Θ_{ν}; by construction, this approach forgoes extrapolation altogether, reducing to the existing semi-structural approach. We see that changes in monetary policy conduct at the short end of the yield curve already suffice to attain quite material reductions in the volatility of output, revealing that, for this particular application, the pure semi-structural approach already works well—and so our hybrid strategy agrees with it, as it should. We elaborate further on the mechanism underlying these findings in Appendix D.3.

Briefly, as shown by Angeletos et al. (2020), the typical U.S. post-war business-cycle fluctuation looks demand-driven, with output dropping and inflation along with it, and also tends to be relatively short-lived. Putting these facts together with the empirically estimated causal effects of transitory rate changes, it follows that such short-lived rate cuts alone would have been sufficient to achieve material output stabilization.

Our third and final result is that the preceding conclusions are largely insensitive to the precise information set used to form the reduced-form projections, illustrating our theoretical discussion of the role of the invertibility assumption. Specifically, in Appendix D.3, we repeat the analysis with reduced-form projections formed using a much smaller three-variable VAR, and yet find results very close to those in Table 5.1.

5.3 Forward guidance in the Great Recession

For our third and final application we evaluate the power of forward guidance, and in particular the role of commitment, during the Great Recession. We use this application both to illustrate the practical relevance of our results on stochastic regime switching, and to give an example of a monetary policy counterfactual where the model-driven extrapolation of causal effects beyond the data takes center stage.

The policy experiment. We seek to compare different policy options during the Great Recession, from 2008:Q4 onwards. We will evaluate three different policy options: a conventional Taylor-type rule; perfect forward guidance of zero nominal interest rates for three years, before ultimately returning to the Taylor rule; and forward guidance with imperfect commitment, where in each quarter the return to the Taylor rule happens with some probability $p = 0.1$, and at the latest after three years, as before.[20] This last experiment allows us to speak in a simple reduced-form way to the tension between policy design under discretion vs. commitment (as in Kydland and Prescott, 1977).

Getting the sufficient statistics. The sufficient statistics are constructed exactly as in the preceding sections: we use a simple reduced-form VAR to construct forecasts from

[20] We consider a Taylor rule as in (11), but now with additional interest rate smoothing to prevent excessively abrupt rate movements after the end of the forward guidance period:

$$i_t = \rho_{tr} i_{t-1} + (1 - \rho_{tr})\phi_\pi \pi_t,$$

with $\rho_{tr} = 0.9$ and $\phi_\pi = 2$.

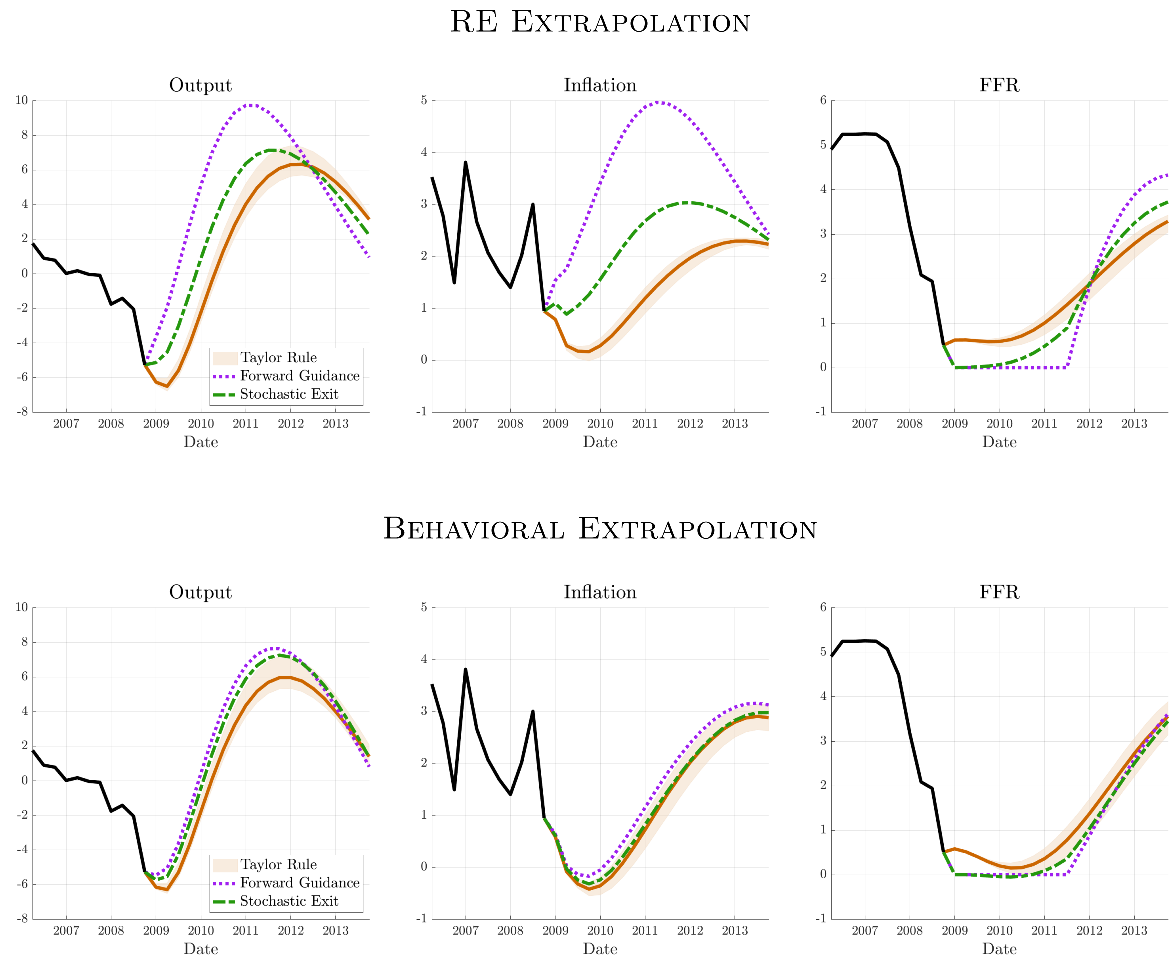

Figure 8: Counterfactual projections of output, inflation, and the federal funds rate during the Great Recession, under an active Taylor rule, forward guidance with perfect commitment, and forward guidance with stochastic exit. Policy causal effects from RANK (top) and B-RANK (bottom). Black: data. Colored: posterior median (solid) and 16th and 84th percentile bands (shaded) for the Taylor rule (orange), forward guidance with perfect commitment (purple dotted), and forward guidance with stochastic exit (green dotted-dashed).

2008:Q4 onwards, and then enforce the hypothesized counterfactual monetary policy rules using interest rate causal effects implied by our four models, under monetary dominance. We note that our estimated VAR predicts a relatively strong bounce back in economic activity; in our subsequent results we do not remark further on this bounce back, but instead focus on the *differences* between the three policy scenarios.

Results. The results are displayed in Figure 8, with the top panel corresponding to causal effects obtained from rational-expectations models (here RANK, with HANK very similar, as discussed further below), and the bottom panel showing results for the behavioral models (here B-RANK, with B-HANK almost identical). The figure shows expectations from 2008:Q4 onwards, obtained by averaging over possible exit histories from the forward guidance commitment. Beginning with the top panel we see that, in expectation from 2008:Q4 onwards, interest rates are a bit above zero under the Taylor rule (orange), perfectly pegged to zero under credible forward guidance (purple), and in between under stochastic commitment (green). Consistent with the literature on the forward guidance puzzle, we see that perfect commitment to low interest rates, when coupled with rational-expectations monetary policy causal effects, implies very sharp expansions in both output and inflation. With imperfect commitment, the effects of this policy are instead much more muted. A key comparison is with the bottom panel: if the monetary policy causal effects are provided by the behavioral models, then forward guidance—even with full commitment—is much weaker, so the three counterfactual lines are very close to each other, largely just reflecting the causal effects of interest rate movements in the near term.

Discussion. The results in Figure 8 are a very clear example of where the available empirical evidence on monetary policy shock propagation alone is not sufficient, and so assumptions on causal effect extrapolation beyond the support of the data take center stage. We already discussed in Section 4.3 that household heterogeneity together with market incompleteness does not really shape that extrapolation all too much; consistent with this observation, we in this application find broad counterfactual equivalence of the estimated RANK and HANK models, with forward guidance moderately more powerful in HANK, but the two pictures overall quite similar (as shown in Appendix D.4). The assumptions about the strength of behavioral frictions, instead, are central, as displayed prominently in Figure 8.

6 Conclusions

Our analysis in this paper has tackled two of the main limitations of the recent semi-structural approach to monetary policy counterfactual evaluation: one conceptual and one practical. *Conceptually*, we have shown that the "sufficient statistics" of that approach are also enough to evaluate policy counterfactuals that involve material policy regime change, e.g., between monetary and fiscal dominance. And *practically*, we have proposed a way of dealing with the high informational requirements. A key takeaway of our analysis is that, at least within the confines of counterfactual questions covered by our results, there is only one remaining role for explicit model structure: to extrapolate from the already estimated effects of short-lived interest rate changes to those of more persistent or delayed changes in policy rates. The main limitation of our results, and of the semi-structural approach more broadly, is the assumed linearity and time invariance of the non-policy block of the economy.

We conclude with a discussion of implications for future empirical and theoretical work. Empirically, an important task will be to estimate the dynamic causal effects of delayed or persistent changes in monetary policy. Doing so would expand further the range of counterfactual questions for which the data alone suffice and model-based extrapolation is unnecessary. Some steps in that direction are taken in Manuel and Wolf (2026). Theoretically, more attention should be paid to how different plausible models of monetary policy transmission extrapolate to the causal effects of further-ahead interest rate changes. In particular, discriminating between models with and without behavioral frictions appears to matter much more than the choice between RANK and HANK, at least for the aggregate monetary policy counterfactuals studied here.

References

Angeletos, G.-M., F. Collard, and H. Dellas (2020): "Business-cycle anatomy," *American Economic Review*, 110, 3030–70.

Angeletos, G.-M., C. Lian, and C. K. Wolf (2023): "Can Deficits Finance Themselves?" Tech. rep., National Bureau of Economic Research.

——— (2025): "Fiscal inaction as monetary support," Working Paper.

——— (2026a): "Deficits and Inflation: HANK meets FTPL," Working Paper.

Angeletos, G.-M., C. Lian, C. K. Wolf, and D. Zhang (2026b): "Monetary-Fiscal Interactions: A Reappraisal," Working Paper.

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.

Aruoba, S. B. and T. Drechsel (2026): "Identifying Monetary Policy Shocks: A Natural Language Approach," *American Economic Journal: Macroeconomics*, forthcoming.

Auclert, A. (2019): "Monetary policy and the redistribution channel," *American Economic Review*, 109, 2333–2367.

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., 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.

——— (2025): "Fiscal and monetary policy with heterogeneous agents," *Annual Review of Economics*, 17, 539–562.

Barnichon, R. (2022): "What If? Monetary Policy in Hindsight," *FRBSF Economic Letter*, 2022, 1–6.

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

——— (2026): "Policy evaluation with sufficient macro statistics: a primer," *The Econometrics Journal*, 29, 27–68.

Barro, R. J. and F. Bianchi (2026): "Fiscal Influences on Inflation in OECD Countries, 2020–2023," *The Economic Journal*, 136, 626–654.

Benati, L., R. E. Lucas Jr, J. P. Nicolini, and W. Weber (2021): "International evidence on long-run money demand," *Journal of monetary economics*, 117, 43–63.

Bianchi, F. (2013): "Regime switches, agents' beliefs, and post-World War II US macroeconomic dynamics," *Review of Economic studies*, 80, 463–490.

Bianchi, F. and L. Melosi (2017): "Escaping the great recession," *American Economic Review*, 107, 1030–1058.

Bocola, L., A. Dovis, K. Jørgensen, and R. Kirpalani (2024): "Bond market views of the fed," Tech. rep., National Bureau of Economic Research.

Bodenstein, M., J. Hebden, and R. Nunes (2012): "Imperfect credibility and the zero lower bound," *Journal of Monetary Economics*, 59, 135–149.

Brockwell, P. J. and R. A. Davis (1991): *Time Series: Theory and Methods*, Springer Series in Statistics, Springer, 2nd ed.

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

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.

Chetty, R. (2009): "Sufficient statistics for welfare analysis: A bridge between structural and reduced-form methods," *Annu. Rev. Econ.*, 1, 451–488.

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.

Christoffel, K., O. De Groot, F. Mazelis, and C. Montes-Galdón (2020): "Using forecast-augmented VAR evidence to dampen the forward guidance puzzle," Tech. rep., ECB working paper.

Cochrane, J. (2023): *The fiscal theory of the price level*, Princeton University Press.

Cogley, T., G. E. Primiceri, and T. J. Sargent (2010): "Inflation-gap persistence in the US," *American Economic Journal: Macroeconomics*, 2, 43–69.

D'Amico, S. and T. B. King (2023): "What does anticipated monetary policy do?" *Journal of Monetary Economics*, 138, 123–139.

Davig, T. and E. M. Leeper (2007): "Generalizing the Taylor principle," *American Economic Review*, 97, 607–635.

De Groot, O., F. Mazelis, R. Motto, and A. Ristiniemi (2021): "A toolkit for computing constrained optimal policy projections (copps)," ECB Working Paper.

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

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.

Farmer, R. E., D. F. Waggoner, and T. Zha (2009): "Understanding Markov-switching rational expectations models," *Journal of Economic theory*, 144, 1849–1867.

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

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.

Friedman, M. (1960): *A Program for Monetary Stability*, Fordham University Press.

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

Gagliardone, L., M. Gertler, S. Lenzu, J. Tielens, and A.-M. Ebsim (2026): "An Odd Approximation of Inflation for State Dependent Pricing Models," Working Paper.

Galí, J. (2015): *Monetary policy, inflation, and the business cycle: an introduction to the New Keynesian framework and its applications*, Princeton University Press.

Gali, J. and T. Monacelli (2005): "Monetary policy and exchange rate volatility in a small open economy," *The Review of Economic Studies*, 72, 707–734.

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.

Gürkaynak, R. S., B. Sack, and E. T. Swanson (2005): "Do actions speak louder than words? The response of asset prices to monetary policy actions and statements," *International Journal of Central Banking*, 1, 55–93.

Haberis, A., R. Harrison, and M. Waldron (2019): "Uncertain policy promises," *European Economic Review*, 111, 459–474.

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.

Jarociński, M. and P. Karadi (2020): "Deconstructing monetary policy surprises—the role of information shocks," *American Economic Journal: Macroeconomics*, 12, 1–43.

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.

Kydland, F. E. and E. C. Prescott (1977): "Rules rather than discretion: The inconsistency of optimal plans," *Journal of political economy*, 85, 473–491.

Leeper, E. M. (1991): "Equilibria under 'active'and 'passive'monetary and fiscal policies," *Journal of monetary Economics*, 27, 129–147.

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

Lubik, T. A. and F. Schorfheide (2004): "Testing for indeterminacy: An application to US monetary policy," *American Economic Review*, 94, 190–217.

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.

Manuel, E. and C. K. Wolf (2026): "Identifying Policy Causal Effects from Rule Changes," Working Paper.

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?" *Econometrica*, 91, 1695–1725.

Miranda-Agrippino, S. and G. Ricco (2021): "The Transmission of Monetary Policy Shocks," *American Economic Journal: Macroeconomics*, 13, 74–107.

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

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.

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

——— (2011): "Stepping on a rake: The role of fiscal policy in the inflation of the 1970s," *European Economic Review*, 55, 48–56.

Sims, E., J. C. Wu, and J. Zhang (2023): "The four-equation new keynesian model," *Review of Economics and Statistics*, 105, 931–947.

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.

Swanson, E. (2024): "The Macroeconomic Effects of the Federal Reserve's Conventional and Unconventional Monetary Policies," *IMF Economic Review*, 72, 1152–1184.

Werning, I. (2015): "Incomplete markets and aggregate demand," Tech. rep., National Bureau of Economic Research.

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 Monetary Policy Counterfactuals: When Do We Need Structural Models?

This appendix contains supplemental material for the article "Evaluating Monetary Policy Counterfactuals: When Do We Need Structural Models?". We provide: (i) some additional theoretical results to complement the discussion in Section 2; (ii) implementation details for our hybrid strategy; (iii) further details for the empirical and model-based analysis of monetary shock propagation in Section 4; (iv) several supplementary results to complement the applications in Section 5; and (v) all proofs.

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

A Supplementary details for our theoretical results

This appendix provides supplementary theoretical results. Appendix A.1 gives some missing details for the description of the model environment in Section 2, Appendix A.2 substantiates our claims on instrument switching, Appendix A.3 elaborates on the role of fiscal policy beyond equilibrium selection, Appendix A.4 gives details for the model with stochastic rule switching, Appendix A.5 describes the models used for our illustrations, and Appendices A.6 and A.7 elaborate on the role of invertibility and forecasting.

A.1 Environment of Section 2.1

This section provides some supplementary details on our general model set-up in Section 2.1. This is necessary for a precise definition of our (conditional) policy counterfactuals of interest, as well as for the proofs in Appendix E.1.

We first note that, under our assumptions on (1), the autocovariance function $\Gamma_y(\bullet)$ of macroeconomic observables y_t under the actual policy rule is given as

$$\Gamma_y(\ell) = \sum_{m=0}^{\infty} \Theta_m \Theta'_{m+\ell}. \tag{A.1}$$

It is straightforward to use (A.1) to derive our first sufficient statistic, the Wold representation (7); see, e.g., Brockwell and Davis (1991).

We next provide some missing details for the definition of the conditional counterfactuals, elaborating on the comment in Footnote 5. First of all note that the counterfactual SVMA (5) embeds the assumption that the counterfactual policy rule (4) is actually followed *forever*, starting in the infinite past. A more practically relevant assumption, however, is that the policymaker unexpectedly changes to the alternative rule (4) at some date t^*, having followed the original rule (3) up to $t^* - 1$. In that case we will have

$$\tilde{y}_t \;=\; \underbrace{\sum_{\ell=0}^{t-t^*} \tilde{\Theta}_\ell \varepsilon_{t-\ell}}_{\text{new shocks after } t^*} \;+\; \underbrace{\tilde{y}_t^*}_{\text{initial conditions}} \tag{A.2}$$

The first term in (A.2) 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 (4). 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

$$\mathcal{H}_w(\tilde{\boldsymbol{w}}^* - \boldsymbol{w}^*) + \mathcal{H}_x(\tilde{\boldsymbol{x}}^* - \boldsymbol{x}^*) + \mathcal{H}_z(\tilde{\boldsymbol{z}}^* - \boldsymbol{z}^*) = \mathbf{0}, \tag{A.3}$$

$$\tilde{\mathcal{A}}_x\tilde{\boldsymbol{x}}^* + \tilde{\mathcal{A}}_z\tilde{\boldsymbol{z}}^* = \mathbf{0}, \tag{A.4}$$

i.e., a system written in terms of forecast revisions.[21] These definitions allow us to introduce two particular 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 (4). We may ask how, from that point onward, the economy would be *predicted* to evolve; i.e., we would like to recover the expectation

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

where $\tilde{y}^*_{t^*+h}$ reflects how the policy change at t^* revises policy promises made and thus expectations formed before t^*.

(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 (4) from date t_1 onward; i.e., we seek to recover

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

where again the second term reflects revisions of past promises.

A.2 Changes in the policy instrument

We remarked in Footnote 4 that the results in Section 2.1 also apply, under a suitable re-interpretation, to switches in the policy instrument. We here elaborate on that observation; throughout, the leading example that we have in mind is a switch between short-term interest rates and the money supply as the instrument of monetary policy. We comment on limitations of this instrument switching logic at the end.

[21] Here, boldface denotes sequences from t^* onwards. (A.4) says the new, counterfactual policy rule holds. By (A.3), the revised forecasts remain consistent with all private-sector relationships.

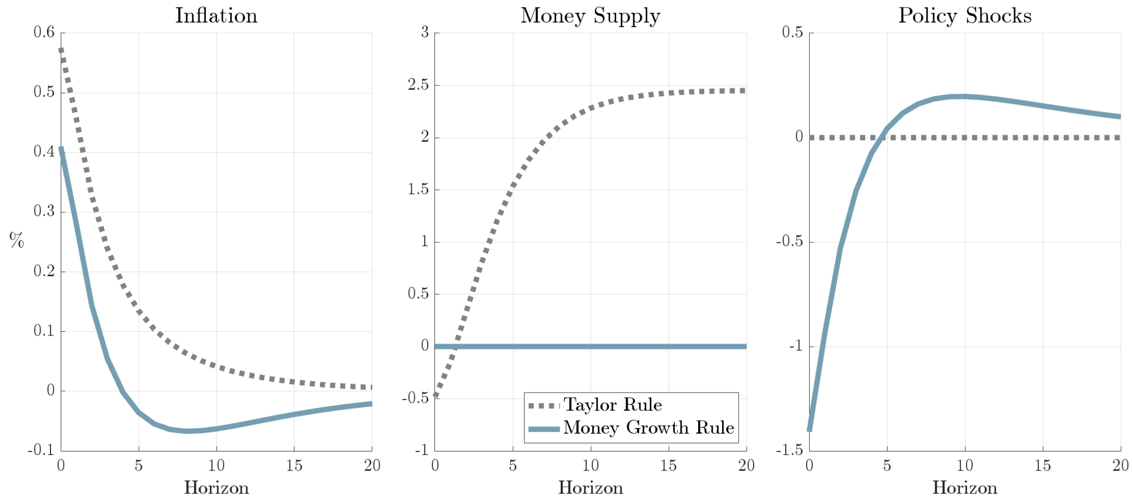

Figure A.1: The model environment and shock are as in Figure 1. The grey dotted lines correspond to outcomes under an active Taylor rule, as before; the blue lines instead show outcomes under a money growth targeting rule, with the right panel reporting the policy shock vector that enforces the counterfactual rule.

Identification result. To see the argument, suppose that now there are two (sets of) policy instruments, $\boldsymbol{z}$ and $\boldsymbol{q}$, tied together through the *private-sector* behavioral relationship

$$\mathcal{K}_w \boldsymbol{w} + \mathcal{K}_x \boldsymbol{x} + \mathcal{K}_q \boldsymbol{q} + \mathcal{K}_z \boldsymbol{z} + \mathcal{K}_e e_0 = \mathbf{0}. \tag{A.5}$$

In the example of switches between interest rates and money supply as the instrument of monetary policy, (A.5) would capture private-sector money demand as a function of interest rates and other possible determinants (e.g., income or inflation). If such a relation exists, then it is immediate that knowledge of the causal effects of shocks to a rule in z-space at the same time also allows evaluation of counterfactual rules formulated in q-space. Formally, such instrument switching is exactly nested in the environment of Section 2.1 with the alternative instrument $q \in x$ and the behavioral relation (A.5) part of the private-sector block (2), and the counterfactual policy rule (4) being formulated in terms of q. Intuitively, what matters is that Θ_ν encodes knowledge of the space of allocations implementable by the policymaker; whether that implementation happens through manipulation of the z or q instruments is, on the other hand, completely immaterial.

Illustration. We provide a visual illustration of this discussion in Figure A.1. For the experiments in that figure we yet again return to the cost-push shock example considered in Section 2, with the New Keynesian model environment now generalized to also feature a simple private-sector liquidity demand relation, linking money supply, inflation, economic activity, and interest rates. The grey dotted lines again show the propagation of the cost-push shock under the active Taylor rule. In that scenario, the short-run rate hike requires an initial contraction in the money supply; ultimately, however, the resulting increase in the price level is associated with an increase in the money supply. We now consider an econometrician that wishes to evaluate outcomes under a counterfactual that fixes the money supply, and to do so leverages our identification results, relying on the effects of the shocks to the policy rate in the baseline Taylor rule regime. The right panel shows that the contemplated money supply rule can be implemented equivalently through a sequence of shocks to the original Taylor rule, first expansionary and then contractionary. By construction, under this policy rule, the money supply does not move, leading inflation to initially increase by less and then undershoot, delivering long-run price stability.

Scope and limitations. The key assumption underlying the preceding discussion was the existence of a private-sector link between instruments, (A.5). For interest rate vs. money growth policies, a money demand equation is the natural example of such a link; as another example, for interest rate vs. exchange rate policy, the analogous relation is an interest rate parity equation (Gali and Monacelli, 2005).

Zooming out, and as already mentioned above, the core logic of these instrument-switching results is that we considered policy instruments that implement the same spaces of private-sector allocations. Our results thus naturally do *not* apply to switches between instruments that instead implement different allocations, e.g., conventional interest rate policy vs. quantitative easing as modeled in Sims et al. (2023).

A.3 Fiscal policy beyond equilibrium selection

We here elaborate further on the discussion at the end of Section 2.2, on joint monetary-fiscal counterfactuals in which the fiscal backing matters not just through equilibrium selection. Formally, we contemplate joint counterfactuals in which the fiscal authority does not respond to the hypothesized monetary change (4) as it did in the original model environment, but instead according to the alternative, counterfactual rule

$$\tilde{\mathcal{B}}_x\left(\boldsymbol{x}-\boldsymbol{x}(e_0)\right)+\tilde{\mathcal{B}}_z\left(\boldsymbol{z}-\boldsymbol{z}(e_0)\right)=\boldsymbol{0}, \tag{A.6}$$

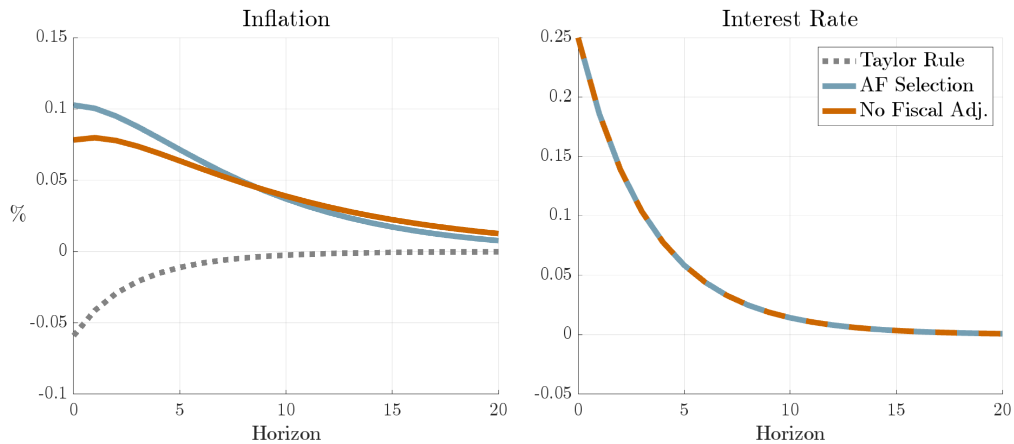

Figure A.2: Inflation impulse responses (left panel) to the rate path (right panel) in a HANK model, under different assumptions on the policy mix: a passive fiscal rule with equilibrium selection through monetary (grey dotted) or fiscal (blue) dominance; or an active fiscal rule (orange).

i.e., a different reaction to any deviation of monetary policy from the original rule. For this joint counterfactual, we require two sets of causal effects: those of changes in interest rates, as before, but now also those of changes in the time path of taxes, to enforce the counterfactual fiscal rule (A.6). More precisely, what matters is their *combination*: we need to know the causal effects of monetary policy shocks if the fiscal rule is simultaneously changed to (A.6), which we denote by $\tilde{\Theta}_\nu$.[22] We then obtain the following version of Proposition 1.

Corollary A.1. *Under the same assumptions as for Proposition 1, knowledge of $\tilde{\Theta}_\nu$ and of the Wold representation* (7) *suffices for policy evaluation under the counterfactual rule* (4) *together with the adjusted "backing"* (A.6)*.*

The intuition is simple: a change in the reaction of the fiscal authority to the counterfactual monetary policy rule switch to (4) generally changes the causal effects of that switch; conditional on those adjusted causal effects $\tilde{\Theta}_\nu$ for monetary policy shocks, however, all of the previous arguments go through without any change.

We provide a brief visual illustration in Figure A.2, showing the effects of an interest rate hike on inflation in a simple HANK model, but with three different assumptions on the fiscal backing. First, if the interest rate hike is implemented in a regime of monetary dominance and with a passive fiscal rule, then it reduces inflation (grey dotted line). Second, we suppose that the same rate hike is still backed by a passive fiscal rule, but now instead use an active

[22]Formally, those causal effects are defined as impulse responses to the shocks $\boldsymbol{\nu}$ to (3'), with the non-policy block (2) unchanged except for the fiscal rule, which is changed to (A.6).

fiscal boundary condition for equilibrium selection—i.e., we still use the baseline monetary policy effects Θ_ν, but select an equilibrium through the intertemporal government budget constraint, exactly as in Figure 1. Inflation now increases (in blue), i.e., the familiar "stepping on a rake" effect of a rate hike (see Sims, 2011). And third, we switch to an active fiscal rule not just for equilibrium selection purposes, but also along the equilibrium transition path, delivering the new interest rate causal effects $\tilde{\Theta}_\nu$, as in Corollary A.1. Inflation still rises (in orange), but now by less on impact; intuitively, following the initial burst of inflation, the government enjoys a budgetary surplus, which it pays out as demand-increasing transfers with the passive fiscal rule, but not under the active one. If households were Ricardian, then those two lines would of course be perfectly on top of each other, illustrating our discussion on equilibrium selection versus on-equilibrium tax-and-transfer path effects. We finally note that, even in HANK, the two lines are still relatively close to each other, reflecting the fact that, quantitatively, the net present value relation on interest rates and inflation enforced by the boundary equilibrium selection is the most important role of the fiscal backing. Our application in Section 5.1 provides another illustration of this, with the equilibrium selection margin, and not assumptions on the precise details of the tax timing, taking center stage.

A.4 Environment for Section 2.3

Recall that the initial policy regime s_0 is fixed, so progressing forward in time there are N^t policy nodes for date t. We write histories of policy regimes as

$$h_t = (s_0, s_1, \ldots, s_t)$$

and now formally define the † notation, described only heuristically in the main text, as follows. For any date t, we write

$$z_t^\dagger = \begin{pmatrix} z_{t,h_t^{(1)}} \\ \vdots \\ z_{t,h_t^{(N^t)}} \end{pmatrix}$$

as the set of date-t policy values z for each possible history up to date t. We then let

$$\boldsymbol{z}^\dagger \equiv \begin{pmatrix} z_0^\dagger \\ z_1^\dagger \\ \vdots \\ z_T^\dagger \end{pmatrix}$$

denote the set of all possible values of z across all histories, and analogously for x, with T possibly infinite. For the remainder of this appendix, to simplify notation, boldface vectors will stack by date, $\boldsymbol{x} = (x_0', x_1', \ldots)'$, with the maps $\mathcal{H}_\bullet$ re-ordered accordingly.

Now write Markov-switching counterfactual policy rules as

$$\tilde{\mathcal{A}}_x^\dagger \boldsymbol{x}^\dagger + \tilde{\mathcal{A}}_z^\dagger \boldsymbol{z}^\dagger = \mathbf{0}. \tag{A.7}$$

In words, (A.7) specifies a relation between endogenous outcomes and the policy instruments for each date and each possible policy history.[23] Next note that the date-0 expectations $\boldsymbol{x}$ and $\boldsymbol{z}$, which enter the unchanged non-policy block (2), satisfy

$$\boldsymbol{x} = \Omega_x \boldsymbol{x}^\dagger, \quad \boldsymbol{z} = \Omega_z \boldsymbol{z}^\dagger$$

where

$$\Omega_x = \text{blkdiag}\left(\omega_0' \otimes I_{n_x}, \omega_1' \otimes I_{n_x}, \ldots, \omega_T' \otimes I_{n_x}\right)$$

and analogously for Ω_z, and $\omega_t \in \mathbb{R}^{N^t}$ is a date-t probability vector giving the probabilities of reaching the respective date-t histories, with

$$\omega_t(h_t) = \prod_{r=0}^{t-1} \Phi_{s_r, s_{r+1}}, \quad \omega_0 = 1.$$

Finally recall that, in this extended model environment, the private-sector block (2) actually applies unchanged, because of certainty equivalence; we can then use the $\Omega_\bullet$ maps to translate the private-sector block to instead be in † space, as done in (10).

To evaluate counterfactual outcomes under the rule (A.7) conditional on any given shock sequence $\boldsymbol{\varepsilon}$ (as required for Proposition 3), we need to translate both the baseline equilibrium outcomes in the original system (2)–(3), as well as policy causal effects Θ_ν, to † space. For the baseline equilibrium outcomes we just repeat the baseline outcomes for all possible policy branches under the new rule, i.e., we set

$$x_t^\dagger(\boldsymbol{\varepsilon}) = \begin{pmatrix} x_t(\boldsymbol{\varepsilon}) \\ \vdots \\ x_t(\boldsymbol{\varepsilon}) \end{pmatrix}, \quad z_t^\dagger(\boldsymbol{\varepsilon}) = \begin{pmatrix} z_t(\boldsymbol{\varepsilon}) \\ \vdots \\ z_t(\boldsymbol{\varepsilon}) \end{pmatrix}$$

[23] Note that this is a construction exactly as in Appendix A.8 of McKay and Wolf (2023), with the aggregate states now the possible realizations of the history of policy rules.

and

$$\boldsymbol{x}^\dagger(\boldsymbol{\varepsilon}) \equiv \begin{pmatrix} x_0^\dagger(\boldsymbol{\varepsilon}) \\ x_1^\dagger(\boldsymbol{\varepsilon}) \\ \vdots \\ x_T^\dagger(\boldsymbol{\varepsilon}) \end{pmatrix}, \quad \boldsymbol{z}^\dagger(\boldsymbol{\varepsilon}) \equiv \begin{pmatrix} z_0^\dagger(\boldsymbol{\varepsilon}) \\ z_1^\dagger(\boldsymbol{\varepsilon}) \\ \vdots \\ z_T^\dagger(\boldsymbol{\varepsilon}) \end{pmatrix}.$$

For policy causal effects define

$$\left[\Theta^\dagger_{x,\nu}\right]_{(t,h_t),(\ell,g_\ell)} = \Theta_{x,\nu}(t,\ell)A_{t,\ell}(h_t,g_\ell), \quad \left[\Theta^\dagger_{z,\nu}\right]_{(t,h_t),(\ell,g_\ell)} = \Theta_{z,\nu}(t,\ell)A_{t,\ell}(h_t,g_\ell)$$

where $h_t = (s_0, \ldots, s_t)$ and $g_\ell = (q_0, \ldots, q_\ell)$ are two possible histories and $A_{t,\ell}(h_t, g_\ell)$ is such that

$$\mathbb{E}_t\left[\nu_\ell\right] = (A_{t,\ell} \otimes I_{n_z})\nu_\ell^\dagger,$$

i.e., $A_{t,\ell}$ is the node-to-node weight matrix for dates (t,ℓ), and can be obtained as a function purely of Φ, with

$$A_{t,\ell}(h_t, g_\ell) = \begin{cases} 1 & \text{if } g_\ell = (s_0, \ldots, s_\ell) \\ 0 & \text{otherwise} \end{cases}$$

if $\ell < t$,

$$A_{t,t} = I_{N^t}$$

if $\ell = t$, and

$$A_{t,\ell} = \begin{cases} \prod_{r=t}^{\ell-1} \Phi_{q_r, q_{r+1}} & \text{if } g_\ell \text{ descends from } h_t \\ 0 & \text{otherwise} \end{cases}$$

if $\ell > t$. Given these expanded policy causal effects, policy counterfactual evaluation can proceed exactly as in our baseline perfect-foresight economies; the final steps are then provided in the proof of Proposition 3.

Historical evolutions and second moments. Proposition 3 delivers counterfactual policy state-dependent paths and date-t expectations conditional on a vector of date-t shocks. It is immediate that, under invertibility, that result also delivers counterfactual forecasts from any given date onwards, i.e., the "conditional forecast" counterfactual. For the historical evolution counterfactual, the researcher would of course need to specify the actual realized policy regime for each date t along the path of interest; conditional on that choice, the arguments of Proposition 1 apply without change. Given an assumption on the realized history of policy regimes for any given sample, second moments can then also straightforwardly be computed.

A.5 Models for illustrative simulations

We provide supplementary details describing the models for our illustrative simulations in Figures 1 to 3, A.1 and A.2.

Simple Rank. The simulations in Figures 1, 2 and A.1 rely on a simple textbook RANK model. The core three equations of the model under active monetary policy are

$$y_t = \mathbb{E}_t\left[y_{t+1}\right] - \frac{1}{\gamma}\left(r_t^n - \mathbb{E}_t\left(\pi_{t+1}\right)\right), \tag{A.8}$$

$$\pi_t = \kappa y_t + \beta \mathbb{E}_t\left(\pi_{t+1}\right) + e_t^s, \tag{A.9}$$

$$r_t^n = \rho_{tr} r_{t-1}^n + (1-\rho_{tr})\phi_\pi \pi_t, \tag{A.10}$$

where the notation follows textbook conventions, with e_t^s as the cost-push shock. The government budget constraint, relevant for the equilibrium selection counterfactual in Figure 1, is

$$b_t = \frac{1}{\beta} b_{t-1} - \tau_t + \frac{\bar{b}}{\bar{y}}(r_{t-1}^n - \pi_t), \tag{A.11}$$

where b_t denotes government debt, τ_t is taxes, and bars denote steady-state values. Finally, for the money growth rule counterfactual in Figure A.1, we also add the private-sector money demand rule

$$m_t - p_t = y_t - \eta r_t^n, \tag{A.12}$$

where p_t is the price level, normalized so that $p_{-1} = 0$.

For all experiments we assume that the cost-push shock follows an ARMA(1,1) process, with autoregressive coefficient 0.7, and moving average coefficient 0.4. We also throughout parameterize the private-sector block with $\beta = 0.99^{1/4}$, $\gamma = 1$, and $\kappa = 0.02$, all conventional values in the literature. The baseline active Taylor rule experiments then correspond to impulse responses with $\rho_{tr} = 0.8$ and $\phi_\pi = 1.7$, delivering equilibrium existence and uniqueness under a passive fiscal backing. Our different counterfactual experiments then make the following further particular assumptions:

- Figure 1: We enforce monetary dominance by switching back to the baseline rule at a large but finite future horizon (following Angeletos et al., 2026a), while we enforce fiscal dominance by requiring that the intertemporal version of the government budget constraint (A.11) holds, with $\bar{b} = \bar{y} = 1$. The latter strategy is equivalent to closing the model with a fiscal rule that sets $\tau_t = 0$ for all t.

- Figure 2: The counterfactual policy rule is again an interest rate peg, but now enforced only stochastically: after date $H = 10$, the policy rule is guaranteed to have returned to the baseline active Taylor rule; between dates 0 and H, there is in each period a constant probability $p = 0.24$ of prematurely switching back to the Taylor rule.

- Figure A.1: We set $\eta = 0.3$, and then close the model with the strict money targeting rule $m_t = 0$ in lieu of (A.10).

Simple Hank. Following Angeletos et al. (2023), we use a simple perpetual-youth OLG structure as a minimal, yet quantitatively relevant, perturbation of the simple RANK model in the direction of HANK. Specifically, the only change is that the IS curve becomes

$$y_t = \mathbb{E}_t\left[y_{t+1}\right] - \frac{1}{\gamma}\left(r_t^n - \mathbb{E}_t\left(\pi_{t+1}\right)\right) + \alpha b_t, \tag{A.13}$$

where $\alpha = \frac{(1-\beta\omega)(1-\omega)}{\omega}$, and $\omega \in [0, 1]$ is the survival probability of households from period to period, with $\omega = 1$ corresponding to the permanent-income limit. As discussed in the main text, in HANK, the fiscal rule matters beyond its role in equilibrium selection, so we now need to also specify the specific tax adjustment rule; we set

$$\tau_t = \frac{\tau_b}{\beta} b_{t-1}. \tag{A.14}$$

For the experiment in Figure A.2 we set $\omega = 0.75$, and begin by computing impulse responses to a one-off monetary policy shock to (A.10) with the same parameterization as above for RANK, and with $\tau_b = 0.4$. We then fix that rate path, and consider two "active-fiscal" counterfactuals: for the first, equilibrium selection counterfactual, we keep $\tau_b = 0.4$, but select among equilibria by requiring that the present value of taxes is zero; for the second counterfactual, we set $\tau_b = 0$.

Smets and Wouters (2007). The laboratory for Figure 3 is the well-known structural model of Smets and Wouters (2007), but with one minor change—we assume that the monetary authority follows rules of the form

$$r_t^n = \rho_{tr} r_{t-1}^n + (1 - \rho_{tr})\left(\phi_\pi \pi_t + \phi_y y_t\right) + v_t, \tag{A.15}$$

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.15) with $\phi_\pi = 1.5$ and $\rho_{tr} = \phi_y = 0$. She then wishes to predict the counterfactual second-moment properties of interest if instead the monetary authority followed the rule (A.15) but with $\rho_{tr} = 0.8$, $\phi_\pi = 1.5$ 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 3.

A.6 Role of invertibility

We begin by establishing, consistent with the intuition given in Sections 2.1 and 3.1, 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.

- *Conditional counterfactuals.*

    (i) 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

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

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

    (ii) We now need to construct $\tilde{y}^1_t$ as well as $\sum_{\ell=0}^{t-t_1} \tilde{\Theta}_\ell \varepsilon_{t-\ell}$. As in the previous item, $\tilde{y}^1_t$ 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 $\tilde{y}_{t_1} = \tilde{\Theta}_0 \varepsilon_{t_1} + \tilde{y}^1_{t_1}$. 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

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

completing the argument.

- *Counterfactual autocovariance function.* 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 $\tilde{\Gamma}_y(\ell)$.

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 3.1 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.7 Behavioral models

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 admissible (and which ones 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

$$\mathcal{H}_{i,w}^R \boldsymbol{w} + \mathcal{H}_{i,x}^R \boldsymbol{x} + \mathcal{H}_{i,z}^R \boldsymbol{z} + \mathcal{H}_{i,e}^R e_0 = \mathbf{0}. \tag{A.16}$$

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.16) is replaced by an alternative of the form

$$\mathcal{H}_{i,w}^B \boldsymbol{w} + \mathcal{H}_{i,x}^B \boldsymbol{x} + \mathcal{H}_{i,z}^B \boldsymbol{z} + \mathcal{H}_{i,e}^B e_0 = \mathbf{0}. \tag{A.17}$$

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 identification result, 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.

[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.

B The hybrid strategy

Appendices B.1 and B.2 provide implementation details for our proposed hybrid strategy: first on impulse response estimation, then on model estimation. Appendix B.3 elaborates on some of the well-known fragilities of standard full-information likelihood-based approaches.

B.1 Impulse response estimation

The first step of the impulse response matching approach is the estimation of the monetary shock impulse response targets. We assume access to n_ν distinct policy shocks, and that the subsequent model estimation targets the impulse responses of n_m outcome variables over H impulse response horizons to the identified shocks. We stack these impulse responses in the vector θ_ν, of dimension $n_\nu \cdot n_m \cdot H$.

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

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

As in much prior work on impulse response matching, we propose to estimate θ_ν using standard Bayesian VAR methods. Estimation delivers draws $i = 1, 2, \ldots, N$ of the policy shock causal effect vector, denoted $\hat{\theta}_{i,\nu}$. We obtain $\hat{\theta}_\nu$ as the posterior mode of the estimated policy shock causal effects. For V_{θ_ν} we first construct:

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

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

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

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

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

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]

B.2 Model estimation

We consider the estimation of a list $\mathcal{M}$ of 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 $\{\mathcal{H}_w, \mathcal{H}_x, \mathcal{H}_z\}$. Each model then has a parameter vector ψ_j mapping into $\{\mathcal{H}_w, \mathcal{H}_x, \mathcal{H}_z\}$, and a prior distribution $p(\psi_j \mid \mathcal{M}_j)$ for the model parameters. We use impulse response matching strategies to arrive at posterior distributions for parameters given models, proceeding in two steps. First, for a given model $\mathcal{M}_j$ and parameter vector ψ_j, we explain how to obtain $\theta_\nu(\psi_j, \mathcal{M}_j)$, the model-implied analogue of the empirically observed policy shock causal effect vector. This step is non-standard; in particular, we will explain why we actually do not need to 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.

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

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

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 conditional on the policy regime, which we for all of our applications take to be monetary dominance. 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 $(n_m\bar{T}) \times \bar{T}$ matrix, where $\bar{T}$ 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 $n = 1, \ldots, n_\nu$, define a $\bar{T} \times 1$ vector of news shocks $\tilde{\boldsymbol{\nu}}_n$. Vertically stack all these vectors of policy news shocks in the $(n_\nu\bar{T}) \times 1$ vector $\tilde{\nu} = [\tilde{\boldsymbol{\nu}}'_1, \ldots, \tilde{\boldsymbol{\nu}}'_{n_\nu}]'$. Define also for convenience the $(n_m n_\nu \bar{T}) \times (n_\nu \bar{T})$ 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

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

   where $\tilde{p}(\hat{\theta}_\nu, \tilde{\theta}_\nu, V_{\theta_\nu})$ is the assumed density for "data" $\hat{\theta}_\nu$ with mean $\tilde{\theta}_\nu$ and covariance matrix V_{θ_ν}, and $\tilde{\nu}_{H+1:\bar{T},n}$ denotes elements $H+1, H+2, \ldots, \bar{T}$ of vector $\tilde{\boldsymbol{\nu}}_n$.[27] Given that $\tilde{p}$ 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).

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

[26] We set a truncation horizon of $\bar{T} = 300$. 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.

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 (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, conditional on fixed equilibrium selection.[28]

Posterior distribution & marginal likelihood. Given the above strategy to evaluate $\theta_\nu(\psi_j, \mathcal{M}_j)$, we can now estimate posteriors for model parameters using a standard limited-information Bayesian estimation approach. We can define an approximate likelihood of the "data," $\hat{\theta}_\nu$, as a function of ψ_j given $\mathcal{M}_j$:

$$p(\hat{\theta}_\nu \mid \psi_j, \mathcal{M}_j) \propto \exp\left[-0.5\left(\hat{\theta}_\nu - \theta_\nu(\psi_j, \mathcal{M}_j)\right)' V_{\theta_\nu}^{-1}\left(\hat{\theta}_\nu - \theta_\nu(\psi_j, \mathcal{M}_j)\right)\right]. \tag{B.3}$$

Combining the prior together with the likelihood (B.3), 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 $\mathcal{M}_j$.

To actually compute these objects 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

[28] This claim works exactly in population as $\bar{T}, H \to \infty$. 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 responses 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.2), to avoid getting too many draws outside of the parameter support.

the proposal distribution, updating the proposal variance-covariance matrix to the observed variance-covariance matrix of parameters in the first N_a draws (scaled 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 $c_2 = 0.6$ for all models. Our acceptance rates for all of the models considered range between 20 and 30 percent. We store $N_d = 1000$ draws, selected as 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]

Note that a parameterized model implies a policy transmission map

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

In order to actually implement our applications, we need to store these large transmission maps for each draw from the posterior of each model; i.e., we need to store 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 $\bar{T} \times \bar{T}$ matrices is very expensive in terms of memory, we store only the $T_u \times T_u$ top left elements, with $T_u = 200$.

B.3 Vulnerabilities of full-information structural approaches

We provide some further details supplementing the discussion in Section 3.3, briefly elaborating on some well-known vulnerabilities of the standard full-information approach.

Model misspecification and inference. Under standard full-information approaches to model estimation (like, e.g., Smets and Wouters, 2007), misspecification 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 misspecification in shock processes, as such misspecification is particularly likely in practice (Chari et al., 2009). Analogous arguments apply to misspecification in policy rules.

Suppose the true data-generating process is

$$y_t = \Theta^*(L)\xi_t, \tag{B.4}$$

$$\xi_t = B^*(L)\varepsilon_t, \tag{B.5}$$

[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.

where $\varepsilon_t \sim N(0, I)$. Relative to (1), the system (B.4)–(B.5) is written to explicitly separate the exogenous process (i.e., equation (B.5)) from the endogenous model propagation (i.e., equation (B.4)). 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 $\Psi^*(L) = \Theta^*(L)B^*(L)$.

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

$$y_t = \Theta_{\psi_1}(L)\xi_t \tag{B.6}$$

$$\xi_t = B_{\psi_2}(L)\varepsilon_t \tag{B.7}$$

where again $\varepsilon_t \sim N(0, I)$. We assume that there is no misspecification in the endogenous propagation part of the model: there is a (in fact unique) ψ_1^* such that $\Theta_{\psi_1^*}(L) = \Theta^*(L)$. The shock process, however, is misspecified; 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.4)–(B.5) and (B.6)–(B.7) 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., misspecification in the endogenous shock propagation part, including in particular the policy space Θ_ν. Since our proposed approach to policy evaluation 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 proposed limited-information approach is arguably less subject to this concern, simply because it only requires the researcher to partially specify the structural 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

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

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

C Monetary policy causal effects

This appendix provides supplementary details on our analysis of monetary shock causal effects along the yield curve in Section 4. We elaborate first on the empirical analysis in Section 4.1, and then on model-based extrapolation in Sections 4.2 and 4.3.

C.1 Empirical evidence

We provide further details on how we construct our empirical estimates of monetary policy shock transmission at the short end of the yield curve.

Data. We are interested in impulse responses of three outcome variables: the output gap, inflation, and the policy rate. These series are constructed as follows. Unless indicated otherwise, each series is transformed to stationarity following Hamilton (2018), and series names refer to FRED mnemonics.

- *Output gap.* We take log output per capita from FRED (`A939RX0Q048SBEA`). We interpret the stationarity-transformed series as a measure of the output gap.
- *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, and we consider the sample period 1969:Q1–2006:Q4, stopping before the Great Recession to capture a relatively stable monetary regime. Our monetary policy shock series are obtained from the replication files of Aruoba and Drechsel (2026) (for their shock) and Ramey (2016) (for the Romer and Romer shock). We aggregate by averaging the monthly series, and we set all missing values of the monetary shock IVs to zero, as in Känzig (2021).

Econometric implementation. We estimate a VAR in the two shock series together with our three outcome variables of interest. Following the recommendations of Plagborg-Møller and Wolf (2021), we order the Aruoba and Drechsel (2026) shock first in a recursive identification of our VAR, thus delivering invertibility-robust estimates of the desired dynamic causal effects. The Romer and Romer (2004) shock we instead order *after* output and inflation (but before the policy instrument), consistent with the original implementation in that paper. We include two lags, a linear time trend, and use a uniform-normal-inverse-Wishart distribution over the orthogonal reduced-form parameterization (as in Arias et al., 2018). Our estimation results are robust to these particular choices. This procedure yields draws of the monetary policy shock causal effect vector $\hat{\theta}_\nu$, which are then used to construct V_{θ_ν} following the steps outlined in Appendix B.1.

C.2 Structural models of monetary policy transmission

This section provides some supplementary details for our structural models of monetary transmission sketched in Section 4.2. 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.2.1 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

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

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 $\tilde{r}_{t+1} = \tilde{r}_t^n - \pi_{t+1}$, $\frac{P_{t+1}}{P_t} = \exp(\pi_{t+1})$, 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.2.3, 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:[32]

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

where $a(u_t)$ is a utility cost of adjusting capacity, and $q_t k_t$ is the total cost of capital purchases for next period.[33] The first-order conditions are:

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

Log-linearizing around the steady state:

$$\begin{aligned}
\tilde{y}_t &= \alpha(\tilde{u}_t + \tilde{k}_{t-1}) + (1-\alpha)\tilde{\ell}_t \\
\tilde{w}_t &= \tilde{p}_t^I + \alpha(\tilde{u}_t + \tilde{k}_{t-1}) - \alpha\tilde{\ell}_t \\
\frac{\psi}{1-\psi}\tilde{u}_t &= \tilde{p}_t^I + (\alpha-1)(\tilde{u}_t + \tilde{k}_{t-1}) + (1-\alpha)\tilde{\ell}_t
\end{aligned}$$

[32] 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.

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

$$\tilde{q}_t = \mathbb{E}_t \left[-\tilde{r}_{t+1} + \left(1 - \frac{1-\delta}{1+\bar{r}}\right) (\tilde{p}^I_{t+1} + (\alpha - 1)(\tilde{k}_t + \tilde{u}_{t+1}) + (1-\alpha)\tilde{\ell}_{t+1}) + \frac{1-\delta}{1+\bar{r}} \tilde{q}_{t+1} \right]$$

where $\frac{\psi}{1-\psi} = a''(1)/a'(1)$ is the curvature parameter of the capacity utilization cost function, which we parametrize following Smets and Wouters (2007) 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.2.3.

Capital good producers solve

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

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:

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

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

$$\tilde{q}_t = \kappa(\tilde{i}_t - \tilde{i}_{t-1}) - \frac{\kappa}{1+\bar{r}} \mathbb{E}_t \left[\tilde{i}_{t+1} - \tilde{i}_t\right]$$

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:

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

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:

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

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:

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

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

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

Policy rule for computation. For our numerical analysis, we close the model with a determinacy-inducing (active) Taylor rule, ensuring monetary dominance:

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

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 5 and 6, we subject this rule to four-year-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.3), and pick the steady state level of government consumption such that the intertemporal government budget constraint holds.

C.2.2 HANK

The only two differences relative to the 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.[34] Households save in government bonds, while firm capital and equity are 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

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

[34] 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

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

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

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

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.2.3 Adding behavioral frictions

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

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

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

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

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

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

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

or

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

and so

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

or

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

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

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

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

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

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

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

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

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

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

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

Substituting the definition of $\check{\pi}_t$, imposing full price indexation ($\zeta = 1$), and noting that $\tilde{\mu}_t = \tilde{p}_t^I$ yields the price-NKPC used in our main analysis:

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

where $\beta^p \equiv \bar{\beta}\theta_p m \left[1 + \frac{\kappa_p}{1-\bar{\beta}\theta_p m} \right]$.

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

$$\mathbb{E}_t \sum_{\tau=t}^{\infty} \left(\bar{\beta}\theta_w\right)^{\tau-t} \left[\lambda_\tau \left(\frac{W_\tau}{P_\tau}\right) \left(\frac{W_{\tau|t}}{W_\tau}\right)^{1-\epsilon_w} - v'\left(\ell_\tau\right) \left(\frac{W_{\tau|t}}{W_\tau}\right)^{-\epsilon_w}\right] L_\tau$$

where λ_τ 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.[35]
The first-order condition is

$$\mathbb{E}_t \sum_{\tau=t}^{\infty} \left(\bar{\beta}\theta_w\right)^{\tau-t} v'\left(\ell_\tau\right) \ell_\tau \epsilon_w W_\tau^{\epsilon_w} \prod_{s=t+1}^{\tau} \exp(-\epsilon_w \zeta_w \pi_{s-1})$$
$$= \mathbb{E}_t \sum_{\tau=t}^{\infty} \left(\bar{\beta}\theta_w\right)^{\tau-t} \lambda_\tau(\epsilon_w - 1) \frac{W_{\tau|t}}{P_\tau} W_\tau^{\epsilon_w} \ell_\tau \prod_{s=t+1}^{\tau} \exp(-\epsilon_w \zeta_w \pi_{s-1}).$$

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

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

or

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

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

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

[35] In the case of RANK, λ_t is as discussed in Appendix C.2.1, 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

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

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

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

Applying cognitive discounting:

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

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

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

With full wage indexation ($\zeta_w = 1$) this gives the wage-NKPC used in our main analysis:

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

where $\beta^w \equiv \bar{\beta}\theta_w m \left[1 + \frac{\kappa_w}{1-\bar{\beta}\theta_w m}\right]$.

C.3 Model calibration and estimation

This section provides details on the parameterization of our estimated models of monetary transmission. We proceed in two steps—first the calibration part, and then the estimation.

Calibration. For all four 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. This yields $\beta = 0.9536$. 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.[36]

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 26 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 lump-sum taxes, with $\tau_b = 0.045$. 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.1.[37]

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,

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

[37] The baseline determinacy-inducing 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 5 and 6.

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 (annual)	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)
$1/\eta$	Gov. debt duration (quarters)	5	Kaplan et al. (2018)
τ_b	Speed of fiscal adjustment	0.045	Gradual fiscal adjustment

Table C.1: Calibrated model parameters.

we fix $m = 0.65$, at the lower end of the range considered by Gabaix (2020). We make this choice because our data are only weakly informative about m; posterior model odds after joint estimation of all models, as reported in previous versions of this paper, illustrate this point.

Figure C.1 repeats the two empirical monetary shock estimation targets from Figure 4 and then overlays the matched impulses at the posterior mode for each of our four models (beige and red, solid and dotted). We see that all four of the models are able to match the two empirical estimation targets reasonably well: both interest rate impulses lead to a hump-shaped decline in output as well as a delayed decline in inflation, with the responses more back-loaded for the more delayed interest rate innovation of Aruoba and Drechsel (2026).[38]

Table C.2 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

[38]While the tightly estimated interest rate path of Romer and Romer (2004) is matched almost perfectly, the fit to the somewhat noisier delayed interest rate path of Aruoba and Drechsel (2026) is worse, with the estimated policy treatment path less persistent than the data target. Forcing closer alignment on that second rate path does not, however, change any of this section's takeaways: behavioral and non-behavioral model variants still largely agree on the effects of this now slightly more gradual treatment, while clearly differing for the more delayed rate changes (see Figure 6). These additional results are available upon request.

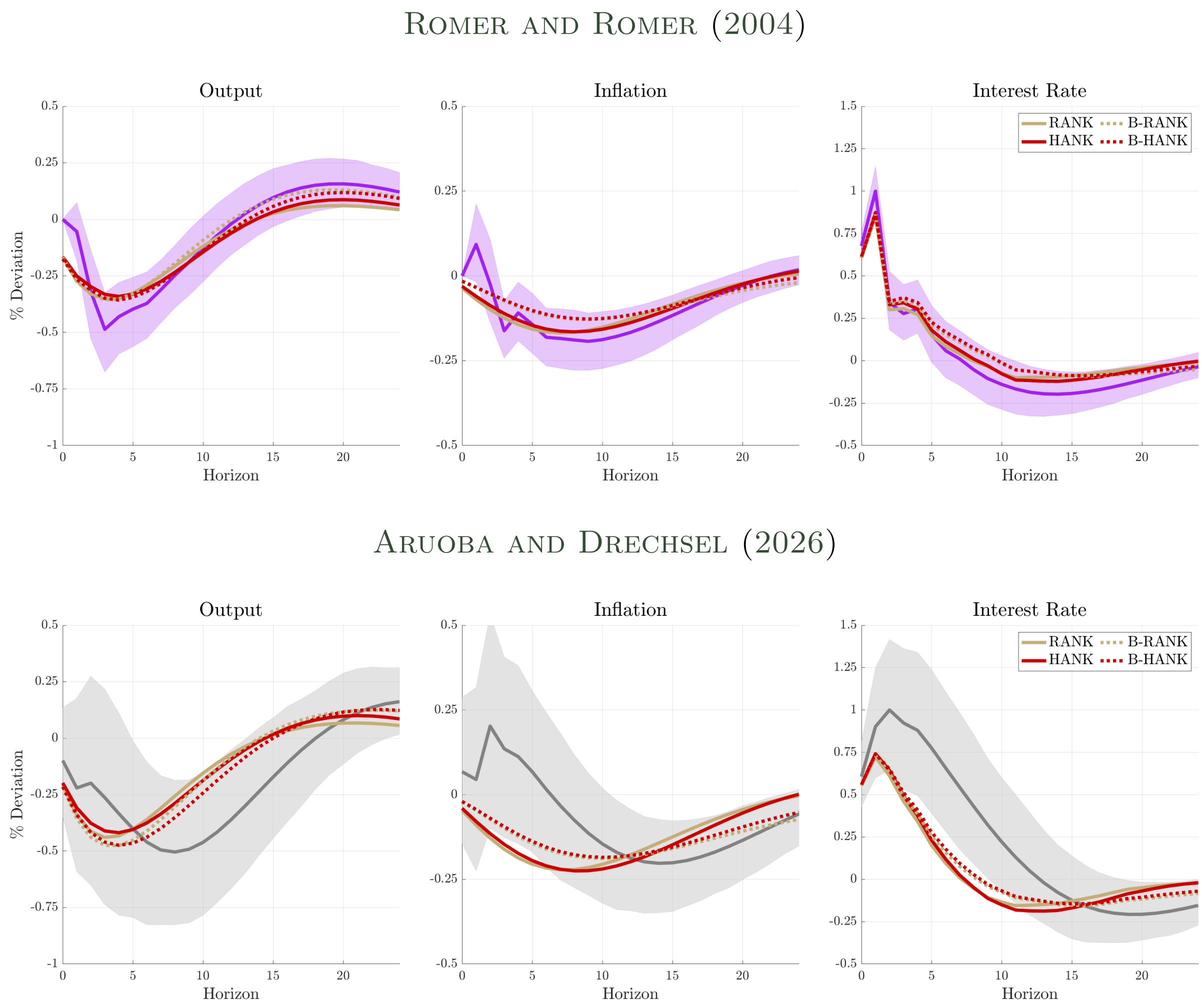

Figure C.1: The purple and grey lines and areas indicate the two empirically estimated monetary shock impulse responses (see Section 4.1), with 16th and 84th percentile confidence bands. The remaining lines indicate the model-implied impulse responses at the estimated posterior modes. Beige: representative-agent consumer block. Red: heterogeneous-agent consumer block. Solid: no cognitive discounting. Dotted: cognitive discounting with $m = 0.65$.

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		0.7636	0.7505	0.7580	0.6019	0.8761
	θ_p	Beta	0.67	0.20		0.9312	0.9137	0.9337	0.7859	0.9725
	θ_w	Beta	0.67	0.20		0.8635	0.7553	0.8161	0.3575	0.9627
	κ	Normal	5.00	1.50		5.3440	5.7793	5.7465	3.7110	7.9987
	ψ	Beta	0.50	0.15		0.3997	0.4541	0.4502	0.2166	0.7066
HANK - RE	θ	Beta	0.70	0.20		0.9710	0.8662	0.8966	0.6393	0.9873
	θ_p	Beta	0.67	0.20		0.9423	0.9123	0.9377	0.7646	0.9769
	θ_w	Beta	0.67	0.20		0.8617	0.7721	0.8302	0.3983	0.9642
	κ	Normal	5.00	1.50		5.6933	5.8925	5.8571	3.8067	8.0830
	ψ	Beta	0.50	0.15		0.4709	0.4459	0.4410	0.2103	0.6921
RANK - CD	h	Beta	0.70	0.10		0.7722	0.7511	0.7583	0.5979	0.8800
	θ_p	Beta	0.67	0.20		0.8596	0.8749	0.9047	0.7023	0.9638
	θ_w	Beta	0.67	0.20		0.9460	0.7607	0.8304	0.3634	0.9635
	κ	Normal	5.00	1.50		5.3338	5.8195	5.7980	3.7185	7.9928
	ψ	Beta	0.50	0.15		0.5478	0.4691	0.4668	0.2302	0.7145
HANK - CD	θ	Beta	0.70	0.20		0.9598	0.8536	0.8843	0.6111	0.9850
	θ_p	Beta	0.67	0.20		0.8498	0.8480	0.8779	0.6328	0.9626
	θ_w	Beta	0.67	0.20		0.9500	0.8090	0.8900	0.4188	0.9646
	κ	Normal	5.00	1.50		5.5923	5.9980	5.9640	3.9139	8.1532
	ψ	Beta	0.50	0.15		0.5335	0.4435	0.4379	0.2135	0.6917

Table C.2: 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

Appendix D.1 provides further details for how we construct the reduced-form projections required for our applications. Supplementary results for our three main applications then follow in Appendices D.2 to D.4. We finally provide a fourth application, to money supply growth rules in the 1970s, in Appendix D.5.

D.1 Reduced-form projections

We provide supplementary details on how we construct the reduced-form projections for our applications. 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`, 2023.03.07 vintage).

- *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 application in Section 5.2, we consider a sample from 1960:Q1–2019:Q4. For Section 5.3, we stop the sample in 2008:Q4, so that our forecasts do not use any future data. Finally, for the post-2021 inflation counterfactual in Section 5.1, we extend the sample to 2021:Q2, the contemplated forecasting date. Note that the output gap, inflation, and federal funds rate series are all constructed exactly as in Appendix C.1 for our empirical analysis of monetary policy shock propagation.

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 5.1 and 5.3, we include two lags. Our use of a reduced-form VAR for this purpose agrees with the findings in Li et al. (2023), who show that VARs tend to dominate other estimation methods in terms of mean-squared error and thus for point estimation.

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).[39] Table D.1 shows the mean squared

[39]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

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

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 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 D.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.

	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	0.716	0.732
Interest rate	0.945	0.953	0.759	0.781

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

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.

D.2 Supplementary results for Section 5.1

We provide further details on the definition and the computation of the "fiscal dominance" counterfactuals reported in Figure 7. For monetary policy itself, the counterfactual is that we peg the path of interest rates at the level implied by the active Taylor rule under monetary dominance—a rate peg that on its own delivers equilibrium indeterminacy. For fiscal policy, we then assume that the fiscal authority responds to this change in monetary policy design not according to the prevailing in-sample rule, but according to the rule described in Appendix C.2, just now with $\tau_b = 0$, i.e., a textbook active fiscal rule, featuring a perpetual absence of fiscal adjustment. Note that, in the language of Section 2.2, this is a joint monetary-fiscal counterfactual: we change fiscal policy in terms of both equilibrium selection and the on-equilibrium tax-and-transfer path.

To evaluate this counterfactual, for RANK, it suffices to implement the interest rate peg and then resolve equilibrium indeterminacy by forcing the government budget constraint to hold in net present value terms in the absence of any tax adjustment. For HANK, matters are slightly more complicated, in line with our discussion at the end of Section 2.2 and in Appendix A.3: our contemplated counterfactual changes the fiscal rule, too, so we use the HANK model to get interest rate causal effects in the absence of fiscal adjustment, and then enforce the counterfactual interest rate peg.[40]

D.3 Supplementary results for Section 5.2

We provide three sets of supplementary results: using the main business-cycle shock of Angeletos et al. (2020) to give further intuition; relying on reduced-form projections using a much smaller information set; and repeating the analysis for a sample excluding the Great Recession, and thereby the binding zero lower bound.

The main business-cycle shock. We zoom in further and study the counterfactual propagation, under our hypothesized optimal targeting rule, of one particular linear combination of Wold residuals that Angeletos et al. (2020) have labeled the "main business-cycle shock"—that is, the reduced-form shock that "explains" the largest share of short-term volatility in real outcomes. The black-dotted lines in Figure D.1 show the propagation of

[40]We could of course also evaluate a pure equilibrium selection counterfactual, as in Figure A.2. This counterfactual is however arguably less empirically relevant than the one we consider: ours has zero tax adjustment forever, the alternative only features zero tax adjustment in present value terms.

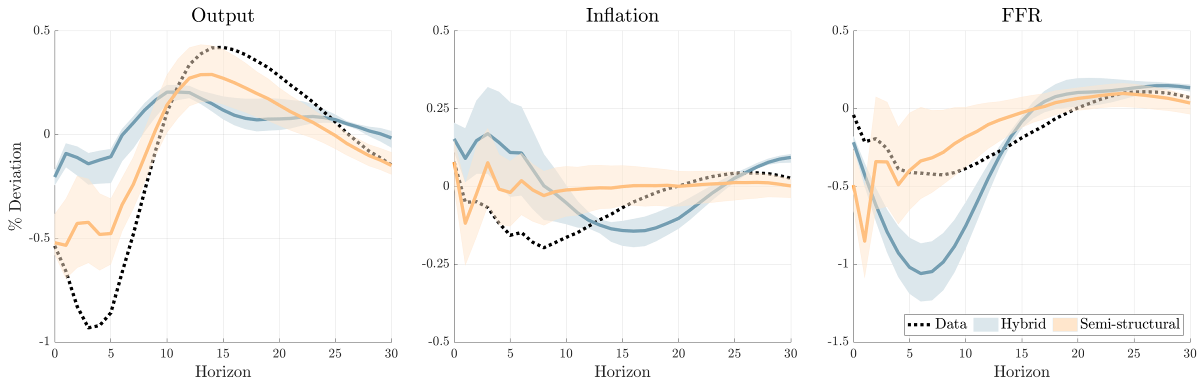

Figure D.1: Counterfactual impulse responses of output, inflation, and the federal funds rate to the main business-cycle shock (see Angeletos et al., 2020), under the policy rule that minimizes (12). Black dotted: data point estimate under observed policy. Blue: posterior 16th and 84th percentile bands with causal effect extrapolation through baseline RANK. Orange: posterior 16th and 84th percentile bands using empirical causal effect estimates only.

this shock under the in-sample monetary policy reaction: inflation drops just a little, output drops quite materially, and monetary policy only somewhat leans against this contraction. The blue and orange lines then report the propagation of this particular reduced-form innovation under our counterfactual monetary policy, showing posterior medians and bands with policy causal effect extrapolation using RANK (blue) and relying on empirical policy causal effect estimates alone (orange).[41] The key observation is that, in both cases, the contemplated counterfactual monetary policy leans against the reduced-form shock much more than the in-sample policy, stabilizing output at the cost of moderately higher inflation and larger policy rate movements. Model-based causal effect extrapolation simply allows the volatilities to be reduced even further by more finely tailoring the path of the interest rate response to the output fluctuations that it is designed to offset.

The takeaway from Figure D.1 is that our headline conclusions in Table 5.1 are essentially driven by two moments of the data: first, that "typical" output movements in the historical time series 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 main results in Table 5.1.

[41]Consistent with our previous discussion, counterfactuals using causal effect extrapolation based on any of the three other models look similar.

	Inflation	Output	Interest Rate
Actual	1.971	2.974	2.907
Counterfactual			
Hybrid			
RANK	1.603	1.675	4.192
	(1.511, 1.786)	(1.523, 1.823)	(3.754, 4.610)
HANK	1.525	1.677	4.047
	(1.449, 1.721)	(1.559, 1.822)	(3.616, 4.432)
B-RANK	1.661	1.704	4.117
	(1.508, 1.900)	(1.423, 1.928)	(3.625, 4.536)
B-HANK	1.657	1.683	4.118
	(1.529, 1.878)	(1.439, 1.894)	(3.639, 4.544)
Semi-structural	1.753	2.197	2.686
	(1.510, 1.923)	(1.882, 2.471)	(2.099, 3.176)

Table D.3: Actual as well as counterfactual unconditional volatilities of inflation, output, and the federal funds rate, with counterfactuals for the policy rule that minimizes (12). Posterior median with 16th and 84th percentiles in parentheses. Projections via a small information set.

A smaller information set. We next repeat the analysis of Section 5.2 using reduced-form projections not based on our large ten-variable VAR, but instead on a very small three-variable VAR, containing only the minimal set of observables needed to implement our hypothesized counterfactual rule: output, inflation, and interest rates. Comparing Table D.3 and Table 5.1 we see that the results are not just qualitatively similar, but instead quite close even quantitatively. This is a practical illustration of the discussion in Section 3.1: the purpose of the invertibility assumption is merely to ensure good forecasting, and so even projections based on small information sets can deliver accurate counterfactuals.

Restricted sample. We finally repeat our analysis on a shorter subsample, now stopping in 2007:Q1, before the Great Recession and thus before the binding zero lower bound. Results are reported in Table D.4; we see that all conclusions extend without change. 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 same logic from the discussion above continues to apply.

	Inflation	Output	Interest Rate
Actual	2.194	3.152	3.037
Counterfactual			
Hybrid			
RANK	1.738	1.894	3.923
	(1.658, 1.858)	(1.719, 2.173)	(3.536, 4.273)
HANK	1.656	1.824	3.750
	(1.602, 1.778)	(1.657, 2.132)	(3.416, 4.032)
B-RANK	1.644	1.672	3.977
	(1.484, 1.988)	(1.458, 1.861)	(3.544, 4.376)
B-HANK	1.605	1.677	3.874
	(1.478, 1.965)	(1.497, 1.896)	(3.475, 4.257)
Semi-structural	1.959	2.349	2.802
	(1.778, 2.097)	(2.052, 2.604)	(2.576, 3.136)

Table D.4: Actual as well as counterfactual unconditional volatilities of inflation, output, and the federal funds rate, with counterfactuals for the policy rule that minimizes (12). Posterior median with 16th and 84th percentiles in parentheses. Estimation sample: 1960:Q1–2007:Q1.

D.4 Supplementary results for Section 5.3

Recall that, in Figure 8, we used monetary policy causal effects extrapolated through the rational-expectations and behavioral RANK models. Figure D.2 shows results for the same exercise using instead the two HANK models. We see that, consistent with the discussion in Section 4.3, the differences between RANK and HANK are materially smaller than the differences between the behavioral and non-behavioral model variants. In particular, moving from RANK to HANK somewhat increases the strength of forward guidance (as in Auclert et al., 2020), but overall delivers a picture very similar to that of Figure 8.

D.5 Further application: a money growth rule in the 1970s

For a practical illustration of how our identification results apply to policy counterfactuals that involve changes in the policy instrument, we here evaluate the counterfactual evolution of the U.S. economy if the Federal Reserve had followed a $k\%$ money growth rule in the spirit of Friedman (1960). Figure D.3 shows the growth of the money supply (M1 plus money market

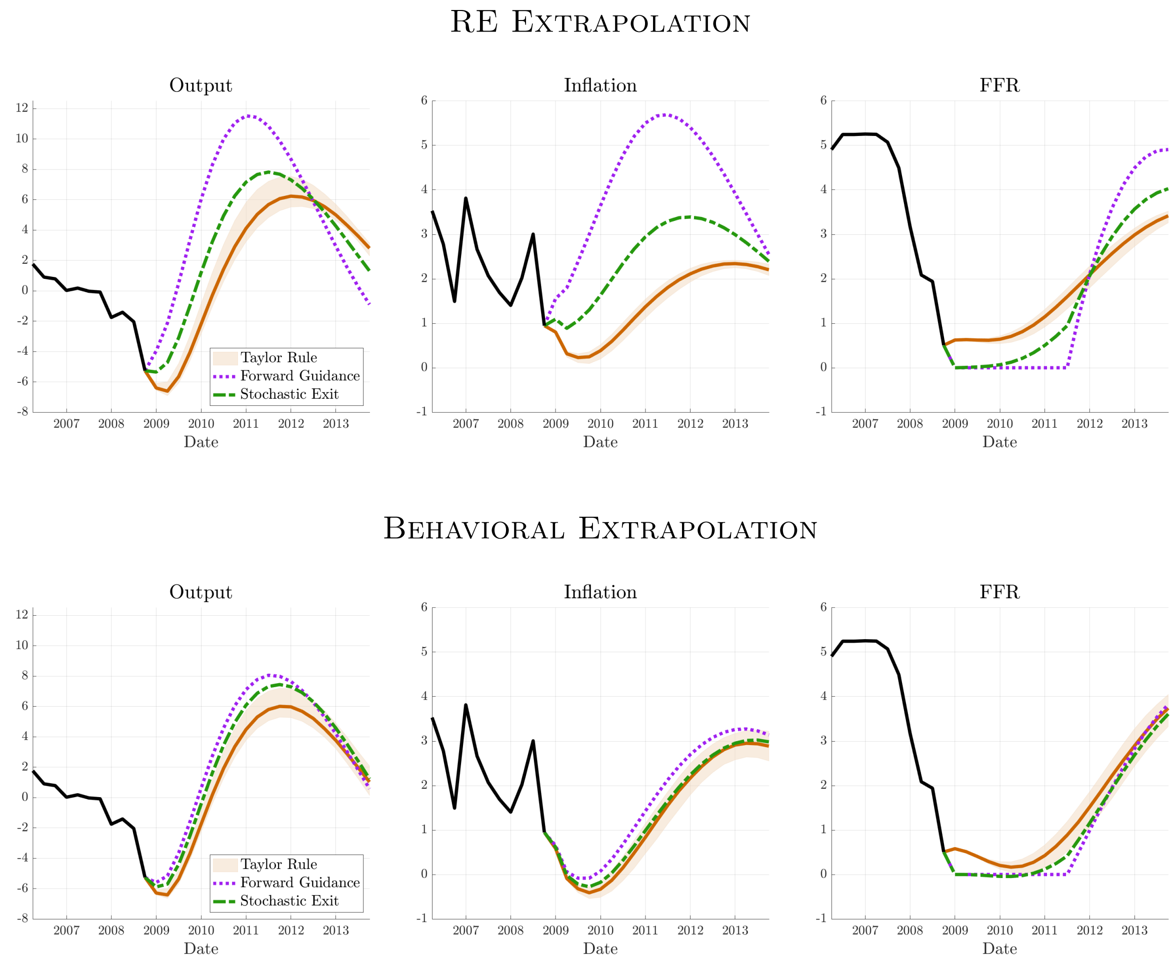

Figure D.2: Counterfactual projections of output, inflation, and the federal funds rate during the Great Recession, under an active Taylor rule, forward guidance with perfect commitment, and forward guidance with stochastic exit. Policy causal effects from HANK (top) and B-HANK (bottom). Black: data. Colored: posterior median (solid) and 16th and 84th percentile bands (shaded) for the Taylor rule (orange), forward guidance with perfect commitment (purple dotted), and forward guidance with stochastic exit (green dotted-dashed).

fund assets) during the 1970s (black) together with the counterfactual that we consider—a constant 5% growth rate from 1972:Q1 onwards. We choose 5% as it corresponds to the trend growth rate of real GDP (3%) plus 2% inflation. We seek to evaluate how the entire decade would have evolved with such a counterfactual rule, i.e., an example of the "historical evolution" counterfactual covered by our identification results.

Getting the sufficient statistics. To implement our identification results, we now require our two sufficient statistics—forecasts and causal effects—not just for output, infla-

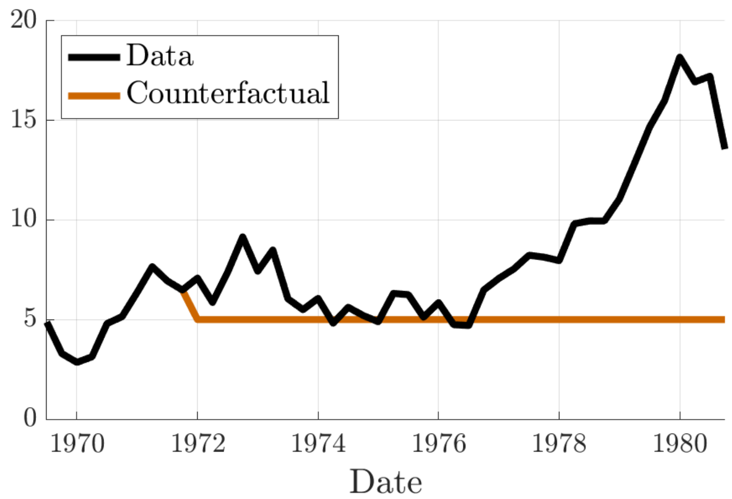

Figure D.3: Actual (black) and counterfactual (orange) growth rate of the money supply (measured as M1 plus money market fund assets) during the 1970s.

tion, and the policy rate, but also for the money supply.

Beginning with policy causal effects, we restrict attention to our rational-expectations RANK model, and now add a money demand equation to that model's non-policy block, exactly as in Appendix A.2. We do so following Benati et al. (2021), who estimate a money demand equation of the form

$$\log M_t = \log(P_t Y_t) - \alpha \log i_t,$$

where M_t is nominal money, $P_t Y_t$ is nominal GDP, and i_t is the short-term nominal rate. They estimate $\alpha = 0.3$; differencing this equation yields $d \log M_t = \pi_t + d \log y_t - \alpha d \log i_t$, which is the relation we add to our RANK model. Given this extension of the model, we can extrapolate from the causal effects of monetary policy on output, inflation and interest rates to the causal effects on money growth.

Turning to forecasts, to align forecasts and causal effects, we perform a change of variables that converts the money supply into a stationary residual in the co-integrating relationship implied by our money demand equation. Formally, we define $z_t \equiv \log M_t - \log(P_t Y_t) + \alpha \log i_t$, and then add this series to our original ten-variable VAR system. Under this change of variables, the policy rule $d \log M_t = k$ becomes

$$dz_t + \pi_t + d \log y_t - \alpha d \log i_t = d \log M_t = k.$$

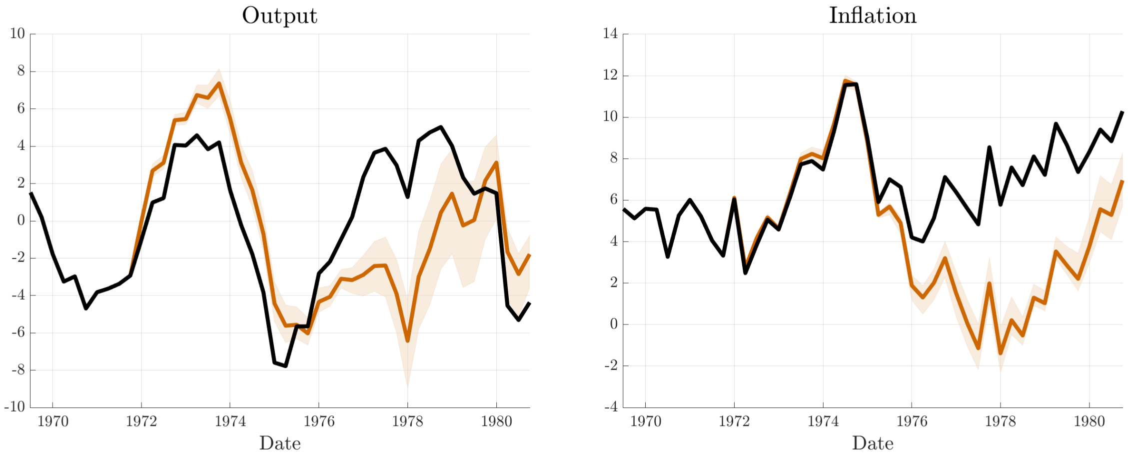

Figure D.4: Counterfactual evolution of output and inflation in the 1970s, under a policy of $k\%$ money growth. Black: data. Orange: counterfactual posterior median (solid) and 16th and 84th percentile bands (shaded). Extended RANK model used for causal effect extrapolation.

Note that the causal effects of policy on dz_t are, by construction, zero. Conceptually, with this approach, we are forecasting the time-varying intercept in our money demand equation, and then assume that monetary policy moves the economy along this demand curve.

Results. Figure D.4 shows the observed data and the counterfactual evolution of the 1970s under our money growth targeting rule. The main impact of the counterfactual is seen in the later part of the decade. Note in Figure D.3 that money growth is only somewhat above 5% until 1977, at which point it begins growing very strongly. Figure D.4 then shows the implications of a lower money growth rate, with output and in particular inflation materially lower from 1976 onwards.

E Proofs

E.1 Proof of Proposition 1

Consider using the policy transmission map Θ_ν to predict the counterfactual propagation of the Wold innovations u_t under the counterfactual policy rule (4), 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

$$\tilde{\mathcal{A}}_x \left(\Psi_{\bullet,x,j} + \Theta_{x,\nu}\tilde{\boldsymbol{\nu}}_j\right) + \tilde{\mathcal{A}}_z \left(\Psi_{\bullet,z,j} + \Theta_{z,\nu}\tilde{\boldsymbol{\nu}}_j\right) = \mathbf{0}. \tag{E.1}$$

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

$$\tilde{y}_t = \sum_{\ell=0}^{\infty} \tilde{\Psi}_\ell u_{t-\ell}. \tag{E.2}$$

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

$$\tilde{\Psi}(L) = \tilde{\Theta}(L)P'. \tag{E.3}$$

We now recover each of the desired counterfactuals.

- *Counterfactual autocovariance function.* Consider using the counterfactual process (E.2) to recover the desired counterfactual second-moment properties. Its implied autocovariance function is

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

where the first equality uses (E.3), and the second one follows since P is an orthogonal matrix.

- *Conditional counterfactuals.* We show how to recover the two conditional counterfactuals defined in Appendix A.1; setting $t^* = -\infty$, this delivers the counterfactual process $\tilde{y}_t$ as defined in Section 2.1. Applying Proposition 1 of McKay and Wolf (2023) to the system (A.3)–(A.4) that defines initial conditions $\tilde{\boldsymbol{y}}^*$, we see that we can recover initial conditions at t^* as

$$\tilde{\boldsymbol{y}}^* = \boldsymbol{y}^* + \Theta_\nu \tilde{\boldsymbol{\nu}}^*, \tag{E.4}$$

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

$$\tilde{\mathcal{A}}_x \left(\boldsymbol{x}^* + \Theta_{x,\nu}\tilde{\boldsymbol{\nu}}^*\right) + \tilde{\mathcal{A}}_z \left(\boldsymbol{z}^* + \Theta_{z,\nu}\tilde{\boldsymbol{\nu}}^*\right) = \mathbf{0}. \tag{E.5}$$

Note that our informational requirements suffice to construct $\tilde{\boldsymbol{\nu}}^*$ and thus allow us to also evaluate the initial conditions term $\tilde{\boldsymbol{y}}^*$. In particular, invertibility ensures that $\boldsymbol{x}^*$ and $\boldsymbol{z}^*$ are equal to date-$t^* - 1$ forecasts based on the Wold representation (7), 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 (E.2) and (E.4) 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 (E.2) and (E.4) to recover the historical counterfactual $\tilde{y}_t$. We have

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

E.2 Proof of Proposition 2

Fix an arbitrary date-0 shock vector e_0, and let $\{\boldsymbol{x}(e_0), \boldsymbol{z}(e_0)\}$ denote the (unique) impulse responses under the baseline rule (3). We will show that the set of possible counterfactual impulse responses under the rule (4) is exactly the set of paths

$$\left\{\boldsymbol{x}(e_0) + \Theta_{x,\nu}\tilde{\boldsymbol{\nu}}, \;\; \boldsymbol{z}(e_0) + \Theta_{z,\nu}\tilde{\boldsymbol{\nu}}\right\}, \tag{E.6}$$

where $\tilde{\boldsymbol{\nu}}$ is any bounded solution of the system

$$\tilde{\mathcal{A}}_x\left(\boldsymbol{x}(e_0)+\Theta_{x,\nu}\tilde{\boldsymbol{\nu}}\right)+\tilde{\mathcal{A}}_z\left(\boldsymbol{z}(e_0)+\Theta_{z,\nu}\tilde{\boldsymbol{\nu}}\right)=\mathbf{0}. \tag{E.7}$$

Since (E.6)–(E.7) involve only the baseline impulse responses and the policy causal effects Θ_ν, this characterization proves the proposition.

First, let $\tilde{\boldsymbol{\nu}}$ be any bounded solution of (E.7). By linearity, the paths (E.6) satisfy the non-policy block (2) given e_0; and by (E.7) they satisfy the counterfactual rule (4). They thus constitute a counterfactual equilibrium.

Conversely, let $\{\tilde{\boldsymbol{x}}(e_0),\tilde{\boldsymbol{z}}(e_0)\}$ be any counterfactual equilibrium, and define the (bounded) artificial policy shock vector

$$\tilde{\boldsymbol{\nu}}\equiv-\left(\mathcal{A}_x\tilde{\boldsymbol{x}}(e_0)+\mathcal{A}_z\tilde{\boldsymbol{z}}(e_0)\right).$$

Then $\{\tilde{\boldsymbol{x}}(e_0),\tilde{\boldsymbol{z}}(e_0)\}$ satisfies the (unchanged) non-policy block (2) as well as, by construction of $\tilde{\boldsymbol{\nu}}$, the generalized baseline rule (3') with policy shocks $\tilde{\boldsymbol{\nu}}$. By linearity and the assumed uniqueness of impulse responses under the baseline rule (Section 2.1), the unique such solution is $\{\boldsymbol{x}(e_0)+\Theta_{x,\nu}\tilde{\boldsymbol{\nu}},\boldsymbol{z}(e_0)+\Theta_{z,\nu}\tilde{\boldsymbol{\nu}}\}$, so the counterfactual equilibrium is of the form (E.6); and since it satisfies the counterfactual rule (4), $\tilde{\boldsymbol{\nu}}$ solves (E.7).

E.3 Proof of Corollary 1

By Proposition 2, under invertibility, knowledge of Θ_ν together with the Wold representation (7) suffices to recover all possible counterfactual Wold representations $\tilde{\Psi}(L)$. As in the proof of Proposition 1, those counterfactual Wold representations are given as $\tilde{\Psi}(L)=\tilde{\Theta}(L)P'$, just now for each possible counterfactual structural impulse responses $\tilde{\Theta}(L)$. Any given assumption on equilibrium selection picks out a unique Wold polynomial among that candidate set; given the unique $\tilde{\Psi}(L)$, the proof of Proposition 1 applies unchanged.

E.4 Proof of Proposition 3

The counterfactual system is

$$\begin{aligned}\mathcal{H}_w^\dagger\boldsymbol{w}^\dagger+\mathcal{H}_x^\dagger\boldsymbol{x}^\dagger+\mathcal{H}_z^\dagger\boldsymbol{z}^\dagger+\mathcal{H}_e^\dagger e_0^\dagger &= \mathbf{0},\\ \tilde{\mathcal{A}}_x^\dagger\boldsymbol{x}^\dagger+\tilde{\mathcal{A}}_z^\dagger\boldsymbol{z}^\dagger &= \mathbf{0}.\end{aligned}$$

By McKay and Wolf (2023), we can evaluate the counterfactual propagation of any shock vector e_0 (in † space) by recovering $\tilde{\boldsymbol{\nu}}^\dagger$ as the unique solution of

$$\tilde{\mathcal{A}}_x^\dagger \left(\boldsymbol{x}^\dagger(e_0) + \Theta_{x,\nu}^\dagger \tilde{\boldsymbol{\nu}}^\dagger\right) + \tilde{\mathcal{A}}_z^\dagger \left(\boldsymbol{z}^\dagger(e_0) + \Theta_{z,\nu}^\dagger \tilde{\boldsymbol{\nu}}^\dagger\right) = \mathbf{0},$$

and then computing

$$\tilde{\boldsymbol{x}}^\dagger(e_0) = \boldsymbol{x}^\dagger(e_0) + \Theta_{x,\nu}^\dagger \tilde{\boldsymbol{\nu}}^\dagger, \quad \tilde{\boldsymbol{z}}^\dagger(e_0) = \boldsymbol{z}^\dagger(e_0) + \Theta_{z,\nu}^\dagger \tilde{\boldsymbol{\nu}}^\dagger.$$

Now note that $\boldsymbol{x}^\dagger(e_0), \boldsymbol{z}^\dagger(e_0)$ can be recovered from the known $\boldsymbol{x}(e_0), \boldsymbol{z}(e_0)$ by just repeating entries across policy histories for a given date, while $\Theta_{x,\nu}^\dagger, \Theta_{z,\nu}^\dagger$ can be recovered from $\Theta_{x,\nu}, \Theta_{z,\nu}$, both as described in Appendix A.4. We can thus evaluate $\tilde{\boldsymbol{\nu}}^\dagger$, from here $\tilde{\boldsymbol{x}}^\dagger(e_0), \tilde{\boldsymbol{z}}^\dagger(e_0)$, and finally

$$\tilde{\boldsymbol{x}}(e_0) = \Omega_x \tilde{\boldsymbol{x}}^\dagger(e_0), \quad \tilde{\boldsymbol{z}}(e_0) = \Omega_z \tilde{\boldsymbol{z}}^\dagger(e_0),$$

completing the argument.

E.5 Proof of Corollary A.1

With slight abuse of notation, we re-use the $\mathcal{H}_\bullet$ notation to refer to the private-sector block exclusive of the fiscal rule, which previously was part of (2). With this convention the counterfactual system corresponding to the joint policy experiment is

$$\begin{aligned} \mathcal{H}_w \boldsymbol{w} + \mathcal{H}_x \boldsymbol{x} + \mathcal{H}_z \boldsymbol{z} + \mathcal{H}_e e_0 &= \mathbf{0}, & \text{(E.8)} \\ \tilde{\mathcal{A}}_x \boldsymbol{x} + \tilde{\mathcal{A}}_z \boldsymbol{z} &= \mathbf{0}, & \text{(E.9)} \\ \tilde{\mathcal{B}}_x \left(\boldsymbol{x} - \boldsymbol{x}(e_0)\right) + \tilde{\mathcal{B}}_z \left(\boldsymbol{z} - \boldsymbol{z}(e_0)\right) &= \mathbf{0}, & \text{(E.10)} \end{aligned}$$

where $\boldsymbol{x}(e_0), \boldsymbol{z}(e_0)$ denote the impulse responses to the shock vector e_0 in the original system (2)–(3). The desired counterfactual impulse responses, $\tilde{\boldsymbol{x}}(e_0), \tilde{\boldsymbol{z}}(e_0)$, solve the system (E.8)–(E.10). Proceeding exactly as in McKay and Wolf (2023), define $\tilde{\boldsymbol{\nu}}$ as the unique solution to the system

$$\tilde{\mathcal{A}}_x \left(\boldsymbol{x}(e_0) + \tilde{\Theta}_{x,\nu} \tilde{\boldsymbol{\nu}}\right) + \tilde{\mathcal{A}}_z \left(\boldsymbol{z}(e_0) + \tilde{\Theta}_{z,\nu} \tilde{\boldsymbol{\nu}}\right) = \mathbf{0},$$

and then recover the desired counterfactual outcomes as

$$\tilde{\boldsymbol{x}}(e_0) = \boldsymbol{x}(e_0) + \tilde{\Theta}_{x,\nu} \tilde{\boldsymbol{\nu}}, \quad \tilde{\boldsymbol{z}}(e_0) = \boldsymbol{z}(e_0) + \tilde{\Theta}_{z,\nu} \tilde{\boldsymbol{\nu}},$$

which by construction satisfy all of (E.8)–(E.10). In particular, the use of the adjusted policy shock causal effects $\tilde{\Theta}_{\nu}$ ensures that the adjusted backing in (E.10) is also satisfied.