What Can Time-Series Regressions Tell Us About Policy Counterfactuals?[†]

Alisdair McKay
FRB Minneapolis

Christian K. Wolf
MIT & NBER

Preliminary and incomplete – comments welcome!

This version: August 2021
First version: July 2021

Abstract: We show that, in a general family of linearized structural macroeconomic models, knowledge of the dynamic causal effects of contemporaneous and news shocks to the prevailing policy rule can be used to: (a) construct counterfactuals under alternative policy rules; and (b) recover the optimal policy rule corresponding to a given loss function. Under our assumptions, the derived counterfactuals and optimal policies are robust to the Lucas critique. We then discuss strategies for applying these insights when only a limited amount of causal evidence on policy shock transmission is available.

[†]Email: `alisdair.mckay@mpls.frb.org` and `ckwolf@mit.edu`. We received helpful comments from Isaiah Andrews, Marios Angeletos, Martin Beraja, Kilian Huber, Geert Mesters, Ben Moll, Emi Nakamura, Mikkel Plagborg-Møller and Jón Steinsson. The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Minneapolis or the Federal Reserve System.

1 Introduction

Conventional wisdom among macroeconomists says that valid policy counterfactuals require a fully-specified structural model. This methodological rule dates back to Lucas (1976), who argued that reduced-form estimates are unlikely to be invariant to changes in policy and so will invariably lead to invalid counterfactuals.

In this paper we revisit the Lucas critique in the context of the linearized structural models that are often used to analyze macroeconomic stabilization policies. We argue that, through the lens of such models, the information contained in empirical time-series regressions is in principle sufficient to sidestep the Lucas critique. Estimates of the dynamic causal effects of policy shocks—typically obtained from vector autoregressions or local projections—allow researchers to construct counterfactuals for a general family of alternative policy rules and to solve for the policy rule that is optimal for a given policymaker loss function. The first part of the paper offers a precise characterization of the information required to construct these counterfactuals: the econometrician needs to know the dynamic causal effects of all possible contemporaneous and news shocks to the prevailing policy rule. An important challenge for our approach is that existing evidence falls short of these high informational requirements. In the second part of the paper we present several strategies that allow researchers to construct counterfactuals using the limited empirical evidence that is actually available in practice. We demonstrate that, at least for monetary policy, the existing evidence is already sufficient to yield meaningful conclusions about policy counterfactuals and optimal policy calculations.

Our results apply to structural models in which private sector behavior depends on policy only through the expected path of the policy instrument, and not on the coefficients of the policy rule *per se*.[1] This property is a feature shared by essentially all linearized business-cycle models, from the analysis of fiscal policy in the real business-cycle model (e.g., Baxter & King, 1993), to monetary policy in the canonical New Keynesian model (e.g., Woodford, 2011; Galí, 2015), to richer models with many frictions and shocks (e.g., Christiano et al., 2005; Smets & Wouters, 2007) and to even the more recent models with rich consumer and firm heterogeneity (e.g., Kaplan et al., 2018; Ottonello & Winberry, 2020). The intuition is simple: linearized models imply certainty equivalence, so policy rules affect the economy only through

[1]More precisely, the policy rule is allowed to matter only through the path of the instrument and through equilibrium selection. Our assumptions on equilibrium existence and uniqueness for the different rules that we consider address equilibrium selection.

the expected value of the policy instrument. For example, in standard models of monetary transmission, households do not care whether nominal interest rates are expected to be high because the systematic monetary rule is hawkish, or because a dovish rule was subject to hawkish "shocks." The key implication of this model property is that we can always re-interpret any arbitrary policy rule as the prevailing rule—the rule generating the observed data—together with a suitable set of contemporaneous and anticipated policy shocks.

We then connect this property of models to empirical methods. Consider an econometrician living in an economy that is consistent with our assumptions and closed with some fixed policy rule. If that policy rule is subject to random disturbances, then the econometrician can use standard semi-structural time-series methods to estimate the dynamic causal effects of policy shocks on outcomes of interest (see Ramey, 2016, for a survey). If she is able to recover the dynamic causal effects of *all* possible contemporaneous and news shocks, then—by our identification results—she can construct policy counterfactuals that are robust to the Lucas critique. First, if she knows the causal effects of policy shocks onto all variables entering some new (counterfactual) policy rule, then she can recover the sequence of shocks that maps the observed rule into that new rule. Second, if she is given a loss function, then she can use her knowledge of the causal effects of the policy instrument onto the policy targets to derive an optimal rule in the form of a forecast targeting criterion.[2] Intuitively, observing the effects of policy shocks tells the econometrician how movements in the policy instrument affect the target variables that enter the loss function. That is, she knows the space of target variables that is implementable through manipulation of the policy instrument. She can then find the point in that space that minimizes the loss function.

The main challenge to implementing these insights is that existing empirical evidence on policy shocks is limited. Formally, the identification result requires that the econometrician can estimate the causal effects of the full menu of possible contemporaneous and news shocks to the prevailing policy rule. For example, in the context of monetary policy, she would need to know the effects of shocks to interest rates at every single point along the yield curve. Such fine-grained, maturity-by-maturity evidence is not available. In the second part of the paper we present two complementary approaches to dealing with this lack of data.

The first approach realizes that, in the face of incomplete empirical evidence, our results apply without any change to the *subspace* of dynamic causal effects spanned by that available

[2]To be clear, our identification results are silent on the shape of the objective function. Explicit, fully specified structural models are one way to arrive at such objective functions. However, given that objective functions in practice are often derived from a legislated mandate rather than economic theory, we believe it is useful to have a method of calculating optimal policy for an objective function that is taken as given.

evidence—that is, we can construct counterfactuals for alternative policy rules that deviate from the prevailing rule in a way consistent with the empirically identified shocks. The more shocks we observe, the richer the deviations from the prevailing rule that we can entertain. By the same token, we can find the optimal policy rule within this spanned set. The more evidence is available, the closer is this restricted optimal rule to the full optimal rule.

We provide a practical illustration of these insights with an application to monetary policy counterfactuals. Our starting point is the causal effect of a contractionary investment-specific technology shock under the actually observed monetary policy reaction, estimated using the shock series of Ben Zeev & Khan (2015). The two objects of interest are then the counterfactual propagation of this shock under (a) an alternative policy rule that aggressively stabilizes output and (b) the optimal policy rule corresponding to a "dual mandate"-type loss function.[3] To construct the two counterfactuals, we follow Christiano et al. (1999) and Gertler & Karadi (2015) to learn about the dynamic causal effects of persistent and short-lived changes in the federal funds rate, respectively. We then leverage our theoretical results to explore the counterfactuals (a) and (b) in this identified subspace.

Our second approach uses additional prior information to extrapolate evidence on the policy shocks that we did observe into the required impulse responses for the ones that we did not. Mathematically, we face a matrix completion problem: we require the full set of policy causal effects (i.e., an infinite-dimensional linear map), yet only have evidence on some specific identified shocks (i.e. some specific weighted averages of the columns of the map). Our strategy is to impose a prior on the required full matrix of causal effects, and then estimate this restricted matrix using the available evidence. To demonstrate this approach, we impose the prior restriction that output and inflation are linked through a Phillips curve. We then combine this (dogmatic) prior on the full inflation and output impulse response maps with empirical evidence obtained from identified monetary policy shocks. This procedure gives an estimate of the full implementable space of output-inflation paths.

Returning to our application, we can now use our estimates of the implementable output-inflation space to characterize output and inflation counterfactuals for (a) an alternative rule that completely stabilizes output and (b) an optimal dual-mandate policy. Now, however, our results are exact, rather than being limited to the identified subspace of causal effects. As expected, the results are similar to the identified subspace analysis, if somewhat smoother and more accurately estimated, reflecting the use of additional prior information.

[3]That is, the loss function is quadratic in output and inflation deviations from trend, with equal weights.

LITERATURE. Our analysis connects with two main strands of literature.

First, we study the relationship between fully structural policy counterfactuals and semi-structural estimates of policy shock transmission (Lucas, 1976; Sims, 1980). Sims (1987) points out that, if "observed data contains exogenous variation in policy along the lines we are currently contemplating," then purely statistical estimates should be informative about policy counterfactuals. The key question is how to implement this idea in a way that accounts for the effects of the counterfactual policy rule on private sector expectations. Estimates of the effects of a single policy shock are generally insufficient for this purpose. For example, the conventional VAR approach to policy counterfactuals—proposed in Sims & Zha (2006) and used in Bernanke et al. (1997) and Eberly et al. (2019)—strings together a sequence of *unanticipated* policy shocks that enforces the alternative rule along the equilibrium path. Such a sequence of unanticipated shocks is equivalent to a rule change only under the restrictive assumption that the private sector today does not respond to policy changes that will materialize tomorrow (see Section 4.1 for a full discussion). By instead combining contemporaneous *and news* shocks, our approach remains valid for a forward-looking private sector, but at the cost of much higher informational requirements.[4] Our work is also related to Barnichon & Mesters (2020b), who use the dynamic causal effects of policy shocks to evaluate the change in a policymaker's objective function following a small change in policy. They argue that such a perturbation is not subject to the Lucas critique because it is small and therefore would not be considered a change in policy regime by the private sector. The logic underlying our results is quite different: we argue that the Lucas critique can be sidestepped not because expectations do not change, but because suitably chosen (news) shocks change expectations just like a counterfactual change in the policy rule.

Second, our identification results build on recent advances in solution methods for dynamic general equilibrium models. As in Auclert et al. (2021), we characterize model equilibria using sequence-space methods; and as in Guren et al. (2021) and Wolf (2020), we connect this sequence-space representation to empirically estimable objects. In contemporaneous and independent work, Groot et al. (2021) and Hebden & Winker (2021) show how to use the same arguments as in our identification results to efficiently *compute* policy counterfactuals in structural models with occasionally binding constraints. Our focus is not computational—we aim to calculate policy counterfactuals directly from empirical evidence, forcing us to confront the fact that such evidence is limited.

[4]The Lucas critique is also sidestepped in Beraja (2020). His analysis has weaker informational requirements than ours, but relies on stronger theoretical restrictions. The two approaches are thus complementary.

Outline. The remainder of the paper proceeds as follows. Section 2 presents our main identification results. Sections 3 and 4 then discuss two approaches to dealing with realistic data limitations, and apply our results to construct monetary policy counterfactuals to investment-specific technology shocks. Section 5 concludes.

2 Dynamic causal effects & policy counterfactuals

This section contains our core identification results. We begin by presenting a simple static version of our argument in the context of a standard small-scale New Keynesian model, and then extend the argument to a general class of linearized dynamic models.

Throughout the paper, we formulate our analysis in linearized perfect-foresight economies. Due to certainty equivalence, the equilibrium dynamics of a linear model with uncertainty will coincide with the solution to such a linearized perfect-foresight economy. We thus emphasize that all results presented below extend without any change to models with aggregate risk solved using conventional first-order perturbation techniques.[5]

2.1 A simple example

We begin with a simple illustration of our core identification argument in the context of an explicit, familiar structural model—the three-equation New Keynesian model (see Woodford, 2011; Galí, 2015).

The perfect-foresight model economy is described by the following equations:

$$y_t = -\sigma(i_t - \pi_{t+1}) + y_{t+1} \tag{1}$$

$$\pi_t = \kappa y_t + \beta \pi_{t+1} + \varepsilon_t^s \tag{2}$$

$$i_t = \phi_\pi \pi_t + \nu_t, \tag{3}$$

where y_t is output, π_t is inflation, i_t is the nominal interest rate, ε_t^s is a cost-push shock, and ν_t is a monetary policy shock. The first two equations describe the behavior of the private sector, while the last equation is the monetary policy rule. Underlying this linear model is a set of structural assumptions that micro-found the parameters of the linearized economy. For our purposes, the crucial property of these micro-foundations is that they do not imply

[5] For example see Fernández-Villaverde et al. (2016), Boppart et al. (2018) or Auclert et al. (2021) for a detailed discussion of this point.

any cross-equation restrictions between the policy rule and the private-sector behavior—σ, κ and β as well as the cost-push shock process are all unaffected by changes in the policy rule (i.e., ϕ_π). To simplify the analysis as much as possible, we assume that shocks are perfectly transitory, so the system (1) - (3) becomes static, with $y_t = \pi_t = i_t = 0$ for $t \geq 2$.[6]

Objects of interest. We wish to characterize the behavior of the economy in response to the cost-push shock ε_t^s under policy rules different from (3). In particular we are interested in the following two counterfactuals. First, we would like to know the behavior of $\{y_t, \pi_t, i_t\}$ in response to the cost-push shock under the alternative policy rule

$$i_t = \tilde{\phi}_\pi \pi_t \tag{4}$$

where $\tilde{\phi}_\pi \neq \phi_\pi$. Second, for a policymaker with a known loss function of the form

$$\lambda_\pi \pi_t^2 + \lambda_y y_t^2, \tag{5}$$

we would like to recover the optimal policy rule and in particular characterize optimal output, inflation and interest rate responses to the cost-push shock ε_t^s.

Empirical evidence. Consider an econometrician who observes data generated from the model (1) - (3) under the baseline monetary policy rule (3). Using conventional semi-structural methods (Ramey, 2016), and with enough data, she can perfectly recover the impulse response matrices $\Theta_{x,\varepsilon,\phi_\pi}$ and Θ_{x,ν,ϕ_π}—the impulse responses of $x = \{y, \pi, i\}$ to shocks ε and ν under the baseline rule ϕ_π—given as

$$\begin{pmatrix} y_t \\ \pi_t \\ i_t \end{pmatrix} = \underbrace{\begin{pmatrix} -\frac{\sigma\phi_\pi}{1+\kappa\sigma\phi_\pi} \\ \frac{1}{1+\kappa\sigma\phi_\pi} \\ \frac{\phi_\pi}{1+\kappa\sigma\phi_\pi} \end{pmatrix}}_{\equiv\Theta_{x,\varepsilon,\phi_\pi}} \times \varepsilon_t^s + \underbrace{\begin{pmatrix} -\frac{\sigma}{1+\kappa\sigma\phi_\pi} \\ -\frac{\kappa\sigma}{1+\kappa\sigma\phi_\pi} \\ \frac{1}{1+\kappa\sigma\phi_\pi} \end{pmatrix}}_{\equiv\Theta_{x,\nu,\phi_\pi}} \times \nu_t \tag{6}$$

Our main result is that knowledge of $\{\Theta_{x,\varepsilon,\phi_\pi}, \Theta_{x,\nu,\phi_\pi}\}$ is in fact sufficient to construct the two desired counterfactuals. That is, knowledge of the causal effects of exogenous and policy shocks ε_t^s and ν_t under some baseline policy rule is actually enough to construct

[6]We furthermore as usual assume that the Taylor principle holds, so the system has a unique bounded solution with the claimed properties.

counterfactual impulse responses to ε_t^s under either the alternative rule (4) or optimal policy for the dual-mandate loss function (5).

ALTERNATIVE POLICY RULES. We begin with our first counterfactual. To construct the counterfactual, we are going to design a monetary shock ν_t that maps the baseline rule (3) into the alternative rule (4). By definition of $\Theta_{x,\varepsilon,\phi_\pi}$ and Θ_{x,ν,ϕ_π}, such a shock—together with the equilibrium aggregates $\{\tilde{y}_t, \tilde{\pi}_t, \tilde{i}_t\}$ that it implies—must satisfy the following system:

$$\begin{aligned}
\tilde{y}_t &= \Theta_{y,\varepsilon,\phi_\pi} \times \varepsilon_t^s + \Theta_{y,\nu,\phi_\pi} \times \nu_t && (7)\\
\tilde{\pi}_t &= \Theta_{\pi,\varepsilon,\phi_\pi} \times \varepsilon_t^s + \Theta_{\pi,\nu,\phi_\pi} \times \nu_t && (8)\\
\tilde{i}_t &= \Theta_{i,\varepsilon,\phi_\pi} \times \varepsilon_t^s + \Theta_{i,\nu,\phi_\pi} \times \nu_t && (9)\\
\tilde{i}_t &= \tilde{\phi}_\pi \tilde{\pi}_t && (10)
\end{aligned}$$

(7) - (9) are the impulse responses to the shock tuple $\{\varepsilon_t^s, \nu_t\}$, and (10) states that the new policy rule (4) holds. In words, we set the shock ν_t to enforce the new policy rule, imposing that the mapping from shocks to equilibrium aggregates is consistent with the impulse response coefficients in $\Theta_{x,\varepsilon,\phi_\pi}$ and Θ_{x,ν,ϕ_π}.

The key point is that the solution to (7) - (10) is identical to the solution we obtain using the structural equations to solve for the equilibrium under the new rule (4). This claim is easily verified. First, the structural solution is $\tilde{\pi}_t = (1 + \kappa\sigma\tilde{\phi}_\pi)^{-1}\varepsilon_t^s$, which we obtain by replacing ϕ_π with $\tilde{\phi}_\pi$ in (6). Alternatively, solving (7) - (10) for ν_t gives

$$\nu_t = -\frac{\tilde{\phi}_\pi \Theta_{\pi,\varepsilon,\phi_\pi} - \Theta_{i,\varepsilon,\phi_\pi}}{\tilde{\phi}_\pi \Theta_{\pi,\nu,\phi_\pi} - \Theta_{i,\nu,\phi_\pi}} \times \varepsilon_t^s = -\frac{\left(\tilde{\phi}_\pi - \phi_\pi\right)\Theta_{\pi,\varepsilon,\phi_\pi}}{\left(\tilde{\phi}_\pi - \phi_\pi\right)\Theta_{\pi,\nu,\phi_\pi} - 1} \times \varepsilon_t^s$$

and so

$$\begin{aligned}
\tilde{\pi}_t &= -\frac{\Theta_{\pi,\varepsilon,\phi_\pi}}{\left(\tilde{\phi}_\pi - \phi_\pi\right)\Theta_{\pi,\nu,\phi_\pi} - 1} \times \varepsilon_t^s = \frac{(1 + \kappa\sigma\phi_\pi)^{-1}}{\left(\tilde{\phi}_\pi - \phi_\pi\right)(1 + \kappa\sigma\phi_\pi)^{-1}\kappa\sigma + 1} \times \varepsilon_t^s \\
&= \frac{1}{1 + \kappa\sigma\tilde{\phi}_\pi} \times \varepsilon_t^s,
\end{aligned}$$

exactly as claimed. The intuition is simple. By construction, the shock ν_t is such that the new policy rule (4) holds. Private sector behavior, however, depends on the policy rule only to the extent that it affects the value of the policy instrument i_t. With i_t set as in the

equilibrium under the new policy rule, all other equilibrium aggregates are also exactly as in that counterfactual equilibrium.

Optimal policy. Next we consider optimal dual mandate policy for a policymaker with preferences described by (5). The conventional, fully structural approach of treating the behavioral relations (1) - (2) as constraints yields the optimal implicit policy rule

$$\pi_t + \frac{\lambda_y}{\kappa \lambda_\pi} y_t = 0. \tag{11}$$

Equation (11) together with the Phillips curve (2) pins down optimal inflation-output pairs. To derive the same rule and optimal outcomes using the measured causal effects, consider the alternative problem of choosing the best *deviation* ν_t from the baseline rule to minimize the policymaker loss. The solution to this problem is characterized by the first-order condition

$$\lambda_\pi \pi_t \Theta_{\pi,\nu,\phi_\pi} + \lambda_y y_t \Theta_{y,\nu,\phi_\pi} = 0 \tag{12}$$

Substituting the definitions of $\Theta_{\pi,\nu,\phi_\pi}$ and Θ_{y,ν,ϕ_π}, we find that (12) reduces to (11), as claimed. Finally, combining this rule with the shock impulse responses $\Theta_{\pi,\varepsilon,\phi_\pi}$ and $\Theta_{y,\varepsilon,\phi_\pi}$, it is straightforward to verify that we obtain the same optimal output-inflation pairs as those computed using the proper optimal policy problem.

As above, the key insight is that—because of linearity—we can think of counterfactual policies as shocks to the baseline rule. By observing these shocks, the econometrician recovers the implementable space of targets (here y-π pairs). Given a loss function, knowledge of this implementable space is enough to recover the optimal policy rule and response to ε_t^s.

Summary & outlook. The simple example of this section is special in two respects. First, the analysis relied on a particular, explicit model—the three-equation New Keynesian model. Second, we have restricted attention to an essentially static version of that economy. The next three subsections show that our identification results apply to a general *family* of linearized *dynamic* models. The analysis will proceed exactly in parallel to the simple example of this section: we first define the environment and the objects of interest, and then present the information required to construct the desired counterfactuals. The key addition relative to the intuition from the simple model will be that, in a general dynamic environment, our identification arguments will require information on impulse response *paths* for both contemporaneous policy shocks as well as policy *news* shocks.

2.2 Model & objects of interest

We consider a linearized perfect foresight model economy. Throughout, boldface denotes time paths for $t = 0, 1, 2, \ldots$, and all variables are expressed in deviations from the model's deterministic steady state.

The model economy is summarized by the equilibrium system

$$\mathcal{H}_w \boldsymbol{w} + \mathcal{H}_x \boldsymbol{x} + \mathcal{H}_z \boldsymbol{z} + \mathcal{H}_\varepsilon \boldsymbol{\varepsilon} = \mathbf{0} \tag{13}$$

$$\mathcal{A}_x \boldsymbol{x} + \mathcal{A}_z \boldsymbol{z} + \boldsymbol{\nu} = \mathbf{0} \tag{14}$$

w and x are n_w- and n_x-dimensional vectors of endogenous variables, z is a n_z-dimensional vector of policy instruments, ε is a n_ε-dimensional vector of exogenous structural shocks, and ν is an n_z-dimensional vector of policy shocks. The distinction between w and x is that all variables in x are observable to the policymaker and econometrician alike, while the variables in w are not. The infinite-dimensional linear maps $\{\mathcal{H}_w, \mathcal{H}_x, \mathcal{H}_z, \mathcal{H}_\varepsilon\}$ summarize the non-policy block of the economy, yielding $n_w + n_x$ restrictions for each t. Our key assumption is that these linear maps do not depend in any way on the coefficients of the policy rule $\{\mathcal{A}_x, \mathcal{A}_z\}$; instead, policy only matters through the path of the instrument z, with the rule (14) giving n_z restrictions on the policy instruments for each t. Note in particular that the policy shock *sequence* $\boldsymbol{\nu}$ contains the full menu of possible contemporaneous ($t = 0$) and news ($t > 0$) shocks to the policy rule (14).

Given $\{\boldsymbol{\varepsilon}, \boldsymbol{\nu}\}$, an equilibrium is a set $\{\boldsymbol{w}, \boldsymbol{x}, \boldsymbol{z}\}$ that solves (13) - (14). We assume that the baseline policy rule $\{\mathcal{A}_x, \mathcal{A}_z\}$ is such that an equilibrium exists and is unique for any $\{\boldsymbol{\varepsilon}, \boldsymbol{\nu}\}$.

Assumption 1. *The policy rule in* (14) *induces a unique and determinate equilibrium. That is, the infinite-dimensional linear map*

$$\mathcal{B} \equiv \begin{pmatrix} \mathcal{H}_w & \mathcal{H}_x & \mathcal{H}_z \\ \mathbf{0} & \mathcal{A}_x & \mathcal{A}_z \end{pmatrix}$$

is invertible.

Given $\{\boldsymbol{\varepsilon}, \boldsymbol{\nu}\}$, we write that unique solution as $\{\boldsymbol{w}_\mathcal{A}(\boldsymbol{\varepsilon}, \boldsymbol{\nu}), \boldsymbol{x}_\mathcal{A}(\boldsymbol{\varepsilon}, \boldsymbol{\nu}), \boldsymbol{z}_\mathcal{A}(\boldsymbol{\varepsilon}, \boldsymbol{\nu})\}$. Most interest will center on impulse responses to exogenous shocks $\boldsymbol{\varepsilon}$ when the policy rule is followed perfectly ($\boldsymbol{\nu} = \mathbf{0}$); with some slight abuse of notation we will simply write those impulse responses as $\{\boldsymbol{w}_\mathcal{A}(\boldsymbol{\varepsilon}), \boldsymbol{x}_\mathcal{A}(\boldsymbol{\varepsilon}), \boldsymbol{z}_\mathcal{A}(\boldsymbol{\varepsilon})\}$.

DISCUSSION & SCOPE. Our results in the remainder of this paper will apply to *any* structural model that can be written in the general form (13) - (14). As emphasized before, the key property of the model is that policy matters for the non-policy block *only* through the realized path of the policy variables $\boldsymbol{z}$; equivalently, in the linearized economy with aggregate risk, policy matters only through its effects on the expected future path of the instrument z. While restrictive, the separation between policy rule and non-policy model blocks that our results require is a feature of many structural models that are used for counterfactual and optimal policy analysis.

To illustrate this point we consider several examples of well-known, explicit structural models that fit into the framework (13) - (14). The static three-equation model of Section 2.1 is an obvious case: Euler equation (1) and Phillips curve (2) are the policy-invariant private-sector block (13), and (3) is the policy rule (14). Appendix A.2 shows the specific linear maps that translate the model into the form of (13)-(14). For a slightly richer example, consider the heterogeneous-agent New Keynesian (HANK) model of Wolf (2021). That model consists of New Keynesian Phillips Curve (NKPC),

$$\boldsymbol{\pi} = \kappa \boldsymbol{y} + \beta \boldsymbol{\pi}_{+1} + \boldsymbol{\varepsilon}^s, \tag{15}$$

a consumer demand block (or IS relation),

$$\boldsymbol{y} = \mathcal{C}_y \boldsymbol{y} + \mathcal{C}_\pi \boldsymbol{\pi} + \mathcal{C}_i \boldsymbol{i} + \mathcal{C}_\tau \boldsymbol{\tau} + \boldsymbol{\varepsilon}^d, \tag{16}$$

a government budget constraint,

$$\mathbf{0} = \bar{\tau} \boldsymbol{\tau} + \bar{b}(1 + \bar{i})(\boldsymbol{i}_{-1} - \boldsymbol{\pi}) \tag{17}$$

and a monetary policy rule,

$$\boldsymbol{i} = \phi_i \boldsymbol{i}_{-1} + (1 - \phi_i)(\phi_\pi \boldsymbol{\pi} + \phi_y \boldsymbol{y}) + \boldsymbol{\nu}, \tag{18}$$

where π is inflation, y is output, i is the nominal rate of interest, τ denotes transfers (which by (17) adjust to balance the government budget), $(\varepsilon^s, \varepsilon^d)$ are supply and demand shocks, and ν is the monetary policy shock.[7] The NKPC is as in the three-equation model, while the coefficient matrices in (16) are derived from aggregating the partial equilibrium household

[7]The subscripts $+1$ and -1 denote time paths shifted forward or backward one period, respectively.

consumption decisions and thus again do not depend on policy rules. This HANK model fits into our structure with $w = \tau$, $x = (\pi, y)$, $z = i$ and $\varepsilon = (\varepsilon^s, \varepsilon^d)$, (15) - (17) as the block (13), and (18) as the policy rule (14).[8] Finally, with slightly more elaborate versions of the same line of reasoning, it is straightforward to see that, once linearized, even workhorse estimated business-cycle models—such as Christiano et al. (2005) or Smets & Wouters (2007)—as well as recent quantitative HANK models—such as Auclert et al. (2020) or McKay & Wieland (2021)—fit into our structure.

While thus clearly quite general, our framework also has important limitations. First, since we leverage certainty equivalence of the linearized model economy, our identification results will generally not yield *globally* valid policy counterfactuals. Second, the policy invariance assumption embedded in the equilibrium system (13) - (14) is not plausible for all kinds of policy rules: it generally holds for rules that only respond to aggregate perturbations of the macro-economy (such as Taylor rules), but will be violated by policies that change the model's steady state. For example, in the HANK model of Wolf (2021) sketched above, changes in the long-run tax-and-transfer system will affect the coefficient matrices in (16), so such policies are necessarily outside the purview of our analysis.[9]

Objects of interest. In the context of the general structural model represented by (13) - (14) we now want to learn about two sets of policy counterfactuals.

1. Arbitrary alternative rules. Consider an alternative policy rule

$$\tilde{\mathcal{A}}_x \boldsymbol{x} + \tilde{\mathcal{A}}_z \boldsymbol{z} = \mathbf{0} \tag{19}$$

Just like the baseline rule, this alternative policy rule is also assumed to induce a unique, determinate equilibrium.

Assumption 2. *The policy rule in* (19) *induces a unique and determinate equilibrium. That is, the infinite-dimensional linear map*

$$\tilde{\mathcal{B}} \equiv \begin{pmatrix} \mathcal{H}_w & \mathcal{H}_x & \mathcal{H}_z \\ \mathbf{0} & \tilde{\mathcal{A}}_x & \tilde{\mathcal{A}}_z \end{pmatrix}$$

[8] The only actual policy choice here is the nominal rate i. Lump-sum taxes—which passively adjust to balance the budget—are thus part of the policy-invariant block (13).

[9] Formally, the coefficient matrices in (16) are derivative matrices for an aggregate consumption function evaluated at the model's steady state. Changes in the tax-and-transfer function change the steady state and so also the coefficient matrices.

is invertible.

Given this alternative rule $\tilde{\mathcal{A}}$, we ask: what are the dynamic response paths $\boldsymbol{x}_{\tilde{\mathcal{A}}}(\boldsymbol{\varepsilon})$ and $\boldsymbol{z}_{\tilde{\mathcal{A}}}(\boldsymbol{\varepsilon})$ to the exogenous shock path $\boldsymbol{\varepsilon}$?

2. Optimal policy. Consider a policymaker with a quadratic loss function of the form

$$\mathcal{L} = \sum_{i=1}^{n_x} \lambda_i \boldsymbol{x}_i' W \boldsymbol{x}_i \tag{20}$$

where i indexes the n_x distinct (observable) macroeconomic aggregates collected in x, λ_i denotes policy weights, and W is a symmetric positive-definite discounting matrix.[10] We assume that the optimal policy problem has a unique solution.

Assumption 3. *Given any vector of exogenous shocks $\boldsymbol{\varepsilon}$, the problem of choosing the policy variable $\boldsymbol{z}$ to minimize the loss function (20) subject to the non-policy constraint (13) has a unique solution.*

We are interested in two questions. First, what policy rule is optimal for such a policymaker? Second, given that optimal rule $(\mathcal{A}_x^*, \mathcal{A}_z^*)$, what are the corresponding dynamic response paths $\boldsymbol{x}_{\mathcal{A}^*}(\boldsymbol{\varepsilon})$ and $\boldsymbol{z}_{\mathcal{A}^*}(\boldsymbol{\varepsilon})$?

The objective of the remainder of this section is to characterize the information required to recover these desired policy counterfactuals. The key insight is that all of the required information can in principle be recovered from data generated under the baseline policy rule.

2.3 Identification

We begin by defining the dynamic causal effects that lie at the heart of our identification results. By Assumption 1, we can write the solution to the system (13) - (14) as

$$\begin{pmatrix} \boldsymbol{w} \\ \boldsymbol{x} \\ \boldsymbol{z} \end{pmatrix} = \underbrace{-\mathcal{B}^{-1} \times \begin{pmatrix} \mathcal{H}_\varepsilon & \mathbf{0} \\ \mathbf{0} & I \end{pmatrix}}_{\equiv \Theta_{\mathcal{A}}} \times \begin{pmatrix} \boldsymbol{\varepsilon} \\ \boldsymbol{\nu} \end{pmatrix}$$

[10] Our focus on a separable quadratic loss functions is in line with standard optimal policy analysis, but not essential. As shown in Appendix A.1, our results extend to the non-separable quadratic case, where the loss is now given by $\boldsymbol{x}'Q\boldsymbol{x}$ for a weighting matrix Q. While our approach in principle also applies to even richer loss functions, the resulting optimal policy rule will generally not fit into the form (14).

The linear map $\Theta_{\mathcal{A}}$ collects the impulse responses of $\boldsymbol{w}$, $\boldsymbol{x}$ and $\boldsymbol{z}$ to the non-policy and policy shocks $(\boldsymbol{\varepsilon}, \boldsymbol{\nu})$ under the prevailing, baseline policy rule (14) with parameters $\mathcal{A}$. We will partition it as

$$\Theta_{\mathcal{A}} \equiv \begin{pmatrix} \Theta_{w,\varepsilon,\mathcal{A}} & \Theta_{w,\nu,\mathcal{A}} \\ \Theta_{x,\varepsilon,\mathcal{A}} & \Theta_{x,\nu,\mathcal{A}} \\ \Theta_{z,\varepsilon,\mathcal{A}} & \Theta_{z,\nu,\mathcal{A}} \end{pmatrix}. \tag{21}$$

All of our identification results will require knowledge of $\{\Theta_{x,\nu,\mathcal{A}}, \Theta_{z,\nu,\mathcal{A}}\}$—the *full* sets of dynamic causal effects of the policy shocks $\boldsymbol{\nu}$. That is, the econometrician needs to know the effects of every possible current and future (announced) deviation from the prevailing policy rule onto the policy instruments z as well as the (observable) endogenous variables x (i.e., all of the arguments of the policy function and the policymaker loss). Furthermore, to construct counterfactual paths that correspond to a given shock sequence $\boldsymbol{\varepsilon}$, the researcher also needs to know the dynamic causal effects of that shock sequence under the baseline policy rule, $\{\boldsymbol{x}_{\mathcal{A}}(\boldsymbol{\varepsilon}), \boldsymbol{z}_{\mathcal{A}}(\boldsymbol{\varepsilon})\}$.

These informational requirements are the natural dynamic generalization of those for the simple model in Section 2.1. First, since the model is now dynamic, a given policy shock now generates entire *paths* of impulse responses, corresponding to the rows of the Θ's. Second, rather than a single shock, we now need to know causal effects corresponding to the full menu of possible contemporaneous and news shocks $\boldsymbol{\nu}$—the columns of the Θ's.

Alternative policy rules. We begin with the first object of interest—policy counterfactuals after a shock sequence $\boldsymbol{\varepsilon}$ under an arbitrary alternative policy rule.

Proposition 1. *Suppose that* $\{\Theta_{x,\nu,\mathcal{A}}, \Theta_{z,\nu,\mathcal{A}}\}$ *and* $\{\boldsymbol{x}_{\mathcal{A}}(\boldsymbol{\varepsilon}), \boldsymbol{z}_{\mathcal{A}}(\boldsymbol{\varepsilon})\}$ *are known. Then, for any alternative policy rule* $\{\tilde{\mathcal{A}}_x, \tilde{\mathcal{A}}_z\}$ *that induces a unique, determinate equilibrium, we can recover the policy counterfactuals* $\boldsymbol{x}_{\tilde{\mathcal{A}}}(\boldsymbol{\varepsilon})$ *and* $\boldsymbol{z}_{\tilde{\mathcal{A}}}(\boldsymbol{\varepsilon})$ *as the unique solution to the system*

$$\begin{pmatrix} I & \mathbf{0} & -\Theta_{x,\nu,\mathcal{A}} \\ \mathbf{0} & I & -\Theta_{z,\nu,\mathcal{A}} \\ \tilde{\mathcal{A}}_x & \tilde{\mathcal{A}}_z & \mathbf{0} \end{pmatrix} \begin{pmatrix} \boldsymbol{x} \\ \boldsymbol{z} \\ \boldsymbol{\nu} \end{pmatrix} = \begin{pmatrix} \boldsymbol{x}_{\mathcal{A}}(\boldsymbol{\varepsilon}) \\ \boldsymbol{z}_{\mathcal{A}}(\boldsymbol{\varepsilon}) \\ \mathbf{0} \end{pmatrix}. \tag{22}$$

Proof. The equilibrium system under the new policy rule can be written as

$$\begin{pmatrix} \mathcal{H}_w & \mathcal{H}_x & \mathcal{H}_z \\ \mathbf{0} & \tilde{\mathcal{A}}_x & \tilde{\mathcal{A}}_z \end{pmatrix} \begin{pmatrix} \boldsymbol{w} \\ \boldsymbol{x} \\ \boldsymbol{z} \end{pmatrix} = \begin{pmatrix} -\mathcal{H}_\varepsilon \\ \mathbf{0} \end{pmatrix} \boldsymbol{\varepsilon} \tag{23}$$

By Assumption 2 we know that (23) has a unique solution $\{\boldsymbol{x}_{\tilde{\mathcal{A}}}(\boldsymbol{\varepsilon}), \boldsymbol{z}_{\tilde{\mathcal{A}}}(\boldsymbol{\varepsilon})\}$. To characterize this solution as a function of observables, consider instead the alternative system (22). Since (13) also holds under the initial policy rule, and since the last line of (22) imposes the new policy rule, it follows that any $(\boldsymbol{x}, \boldsymbol{z})$ that are part of a solution of (22) are also part of a solution of (23). Since by assumption (23) has a unique solution, it follows that the system (22) is solved by at most one set of paths $(\boldsymbol{x}, \boldsymbol{z})$.

It remains to establish that the system (22) has a solution. For this consider the candidate tuple $\{\boldsymbol{x}_{\tilde{\mathcal{A}}}(\boldsymbol{\varepsilon}), \boldsymbol{z}_{\tilde{\mathcal{A}}}(\boldsymbol{\varepsilon}), \boldsymbol{\nu} = (\tilde{\mathcal{A}}_x - \mathcal{A}_x)\boldsymbol{x}_{\tilde{\mathcal{A}}}(\boldsymbol{\varepsilon}) + (\tilde{\mathcal{A}}_z - \mathcal{A}_z)\boldsymbol{z}_{\tilde{\mathcal{A}}}(\boldsymbol{\varepsilon})\}$. Since the paths $\{\boldsymbol{w}_{\tilde{\mathcal{A}}}(\boldsymbol{\varepsilon}), \boldsymbol{x}_{\tilde{\mathcal{A}}}(\boldsymbol{\varepsilon}), \boldsymbol{z}_{\tilde{\mathcal{A}}}(\boldsymbol{\varepsilon})\}$ solve (23), it follows that they are also a solution to the system

$$\begin{pmatrix} \mathcal{H}_w & \mathcal{H}_x & \mathcal{H}_z \\ \mathbf{0} & \mathcal{A}_x & \mathcal{A}_z \end{pmatrix} \begin{pmatrix} \boldsymbol{w} \\ \boldsymbol{x} \\ \boldsymbol{z} \end{pmatrix} = - \begin{pmatrix} \mathcal{H}_\varepsilon \boldsymbol{\varepsilon} \\ (\tilde{\mathcal{A}}_x - \mathcal{A}_x)\boldsymbol{x}_{\tilde{\mathcal{A}}}(\boldsymbol{\varepsilon}) + (\tilde{\mathcal{A}}_z - \mathcal{A}_z)\boldsymbol{z}_{\tilde{\mathcal{A}}}(\boldsymbol{\varepsilon}) \end{pmatrix} \tag{24}$$

But by Assumption 1 we know that the system (24) has a unique solution, so indeed the paths $\{\boldsymbol{w}_{\tilde{\mathcal{A}}}(\boldsymbol{\varepsilon}), \boldsymbol{x}_{\tilde{\mathcal{A}}}(\boldsymbol{\varepsilon}), \boldsymbol{z}_{\tilde{\mathcal{A}}}(\boldsymbol{\varepsilon})\}$ are that solution. It then follows from the definition of $\Theta_{\mathcal{A}}$ in (21) that the candidate tuple also solves (22), completing the argument. □

(22) is the dynamic generalization of the system (7) - (10). The intuition is exactly the same: since we know the effects of all possible perturbations $\boldsymbol{\nu}$ of the baseline rule, we can always construct a perturbation that mimics the equilibrium instrument path under the new instrument rule. But since the first model block (13) depends on the policy rule *only* via the expected instrument path, the equilibrium allocations under the new counterfactual rule and the perturbed prevailing rule are the same. The only difference relative to the simple model is that, because the full system is dynamic, we need contemporaneous *and* news shocks to the baseline rule to mimic an arbitrary alternative rule.

Optimal policy. The second identification result concerns optimal policy.

Proposition 2. *Consider a policymaker loss function (20), and suppose that $\{\Theta_{x,\nu,\mathcal{A}}, \Theta_{z,\nu,\mathcal{A}}\}$ are known. Then we can recover the optimal policy rule $\{\mathcal{A}_x^*, \mathcal{A}_z^*\}$ as*

$$\mathcal{A}_x^* = \left(\lambda_1 \Theta'_{x_1,\nu,\mathcal{A}} W, \lambda_2 \Theta'_{x_2,\nu,\mathcal{A}} W, \ldots, \lambda_{n_x} \Theta'_{x_{n_x},\nu,\mathcal{A}} W\right) \tag{25}$$

$$\mathcal{A}_z^* = \mathbf{0}. \tag{26}$$

If $\{\boldsymbol{x}_{\mathcal{A}}(\varepsilon), z_{\mathcal{A}}(\varepsilon)\}$ are also known, then we can furthermore recover counterfactuals for the shock path $\boldsymbol{\varepsilon}$ under the optimal policy rule, $\boldsymbol{x}_{\mathcal{A}^*}(\boldsymbol{\varepsilon})$ and $\boldsymbol{z}_{\mathcal{A}^*}(\boldsymbol{\varepsilon})$, through Proposition 1.

Proof. The solution to the optimal policy problem is characterized by the following first-order conditions:

$$\begin{aligned} \mathcal{H}'_w(I \otimes W)\boldsymbol{\varphi} &= \mathbf{0} && (27)\\ (\Lambda \otimes W)\boldsymbol{x} + \mathcal{H}'_x(I \otimes W)\boldsymbol{\varphi} &= \mathbf{0} && (28)\\ \mathcal{H}'_z W\boldsymbol{\varphi} &= \mathbf{0} && (29) \end{aligned}$$

where $\Lambda = \text{diag}(\lambda_1, \lambda_2, \dots)$ and φ is the multiplier on (13). By Assumption 3 we know that the system (27) - (29) together with (13) has a unique solution $\{\boldsymbol{x}^*(\boldsymbol{\varepsilon}), \boldsymbol{z}^*(\boldsymbol{\varepsilon}), \boldsymbol{\varphi}^*(\boldsymbol{\varepsilon})\}$.

Now consider the alternative problem of choosing deviations $\boldsymbol{\nu}$ from the prevailing rule to minimize (20) subject to the system (13) - (14). This second problem gives the FOCs

$$\begin{aligned} \mathcal{H}'_w(I \otimes W)\boldsymbol{\varphi} &= \mathbf{0} && (30)\\ (\Lambda \otimes W)\boldsymbol{x} + \mathcal{H}'_x(I \otimes W)\boldsymbol{\varphi} + \mathcal{A}'_x W\boldsymbol{\varphi}_z &= \mathbf{0} && (31)\\ \mathcal{H}'_z(I \otimes W)\boldsymbol{\varphi} + \mathcal{A}'_z W\boldsymbol{\varphi}_z &= \mathbf{0} && (32)\\ W\boldsymbol{\varphi}_z &= \mathbf{0} && (33) \end{aligned}$$

where φ_z is the multiplier on (14). It follows from (33) that $\boldsymbol{\varphi}_z = \mathbf{0}$. But then (30) - (32) together with (13) determine the same unique solution as before, and $\boldsymbol{\nu}$ adjusts residually to satisfy (14). The original problem and the alternative problem are thus equivalent.

Next note that, by Assumption 1, we can re-write the alternative problem's constraint set as

$$\begin{pmatrix} \boldsymbol{w} \\ \boldsymbol{x} \\ \boldsymbol{z} \end{pmatrix} = \Theta_{\mathcal{A}} \times \begin{pmatrix} \boldsymbol{\varepsilon} \\ \boldsymbol{\nu} \end{pmatrix} \tag{34}$$

The problem of minimizing (20) subject to (34) gives the optimality condition

$$\sum_{i=1}^{n_x} \lambda_i \Theta'_{x_i,\nu,\mathcal{A}} W \boldsymbol{x}_i = 0 \tag{35}$$

By the equivalence of the policy problems, it follows that (35) is an optimal policy *rule*, taking the form (25) - (26). Finally, the second part of the result follows from Proposition 1 since (35) is just a special example of a policy rule $\{\tilde{\mathcal{A}}_x, \tilde{\mathcal{A}}_z\}$.

□

The optimal policy rule in (35) is the dynamic analogue of the static rule in (12).[11] The intuition is as before: since we know the (now dynamic) causal effects of every possible policy perturbation $\boldsymbol{\nu}$ on the policymaker targets $\boldsymbol{x}$, we in fact know the space of those targets that is implementable through policy actions. At an optimum, we must be at the point of this space that minimizes the policymaker loss (20).

Aside: relative impulse responses. Our statements of Propositions 1 and 2 use *absolute* impulse responses $\{\Theta_{x,\nu,\mathcal{A}}, \Theta_{z,\nu,\mathcal{A}}\}$. Both results, however, actually only require information on *relative* dynamic causal effects: if, for example, the first impulse response map $\Theta_{x_1,\nu,\mathcal{A}}$ is invertible, then the proofs of both results apply without any change to the maps $\{\tilde{\Theta}_{x,\nu,\mathcal{A}}, \tilde{\Theta}_{z,\nu,\mathcal{A}}\}$, where $\tilde{\Theta}_{x_i,\nu,\mathcal{A}} \equiv \Theta_{x_i,\nu,\mathcal{A}} \times \Theta^{-1}_{x_1,\nu,\mathcal{A}}$ and $\tilde{\Theta}_{z_i,\nu,\mathcal{A}} \equiv \Theta_{z_i,\nu,\mathcal{A}} \times \Theta^{-1}_{x_1,\nu,\mathcal{A}}$. Intuitively, both for counterfactual rules of the form (19) as well as for optimal policy, the only information required by the econometrician are the relative (or normalized) implementable spaces of policy targets and instruments x and z. Our connection of theory and measurement in Section 4.2 will heavily leverage this observation.

2.4 Quantitative illustration

We complement our theoretical discussion of identification with a numerical illustration in the context of a quantitative HANK model. The purpose of this section is to provide a visual representation of our results in the context of a model that is—unlike the simple case of Section 2.1—neither static nor solvable in closed-form.

We use the HANK model of Wolf (2021), sketched in Section 2.2 and with details of the model parameterization relegated to Appendix A.3. We first of all solve the model with a baseline policy rule of

$$i_t = \phi_\pi \pi_t \tag{37}$$

for $\phi_\pi = 1.5$. Using this solution we compute the policy causal effects $\{\Theta_{x,\nu,\mathcal{A}}, \Theta_{z,\nu,\mathcal{A}}\}$ and

[11] Note that, by mapping our perfect foresight economy to a linearized economy with aggregate risk, we can re-write that optimal policy rule as a forecasting targeting rule (Svensson, 1997):

$$\sum_{i=1}^{n_x} \lambda_i \Theta'_{x_i,\nu,\mathcal{A}} W \mathbb{E}_t [\boldsymbol{x}_i] = 0 \tag{36}$$

where now $\boldsymbol{x}_i = (x_{it}, x_{it+1}, \dots)'$. In words, expectations of future targets must always minimize the policymaker loss within the space of (expected) allocations that are implementable via changes in the policy stance. For a timeless perspective, (36) must apply to *revisions* of policymaker expectations at each t.

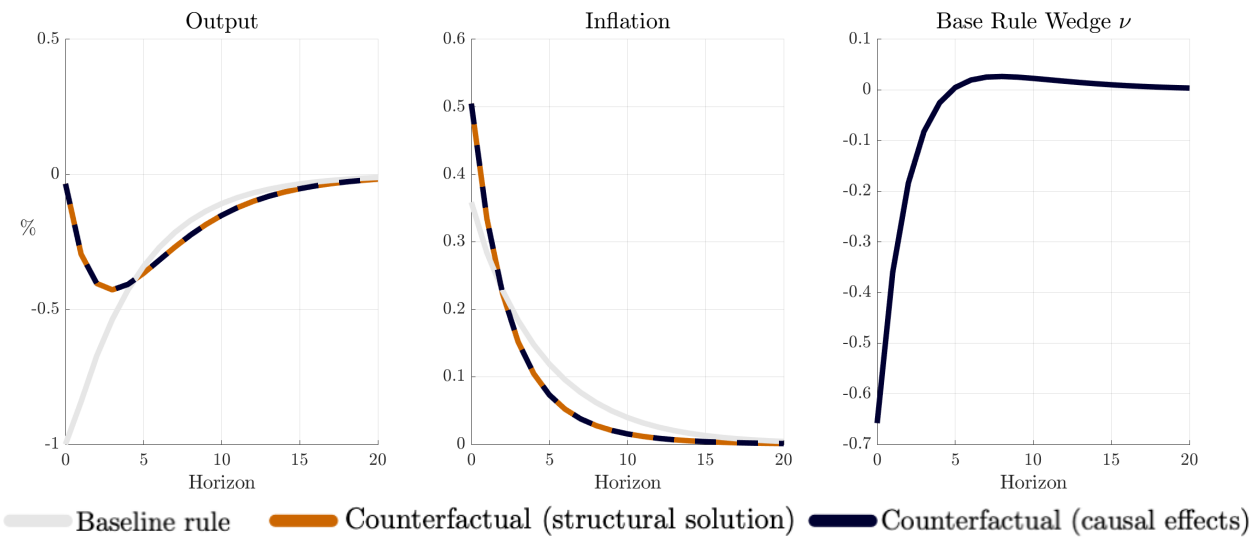


Figure 1: Output and inflation impulse responses together with the equivalence shock wedge $\boldsymbol{\nu}$ (see (22)) for the HANK model with policy rules (37) and (38). The impact output contraction under the prevailing baseline rule is normalized to -1%.

the impulse responses $\{\boldsymbol{x}_{\mathcal{A}}(\boldsymbol{\varepsilon}), \boldsymbol{z}_{\mathcal{A}}(\boldsymbol{\varepsilon})\}$ to a contractionary cost-push shock ε^s. Then, following Propositions 1 and 2, we use those impulse responses to construct policy counterfactuals.

Alternative policy rules. For our first experiment, we would like to learn about the behavior of output and inflation under an alternative policy rule

$$i_t = \phi_i i_{t-1} + (1-\phi_i)(\phi_\pi \pi_t + \phi_y y_t) \tag{38}$$

for $\phi_i = 0.9$, $\phi_\pi = 2$ and $\phi_y = 0.5$.

We will perform this calculation in two ways. First, we make use of the structural equations of the model: we simply replace the baseline policy rule with the alternative rule and then re-solve the model. The cost-push shock impulse responses under the baseline rule and the counterfactual new rule are displayed as the grey and orange lines in Figure 1.

Next, we use Proposition 1 to equivalently construct the desired counterfactual without actually re-solving the model. We do so using $\{\boldsymbol{x}_{\mathcal{A}}(\boldsymbol{\varepsilon}), \boldsymbol{z}_{\mathcal{A}}(\boldsymbol{\varepsilon})\}$ and $\{\Theta_{x,\nu,\mathcal{A}}, \Theta_{z,\nu,\mathcal{A}}\}$—the dynamic causal effects of the fundamental shock and of policy shocks generated under the prevailing baseline rule (37). We feed these inputs into (22) to solve for $\boldsymbol{x}$, $\boldsymbol{z}$ and $\boldsymbol{\nu}$. The dark

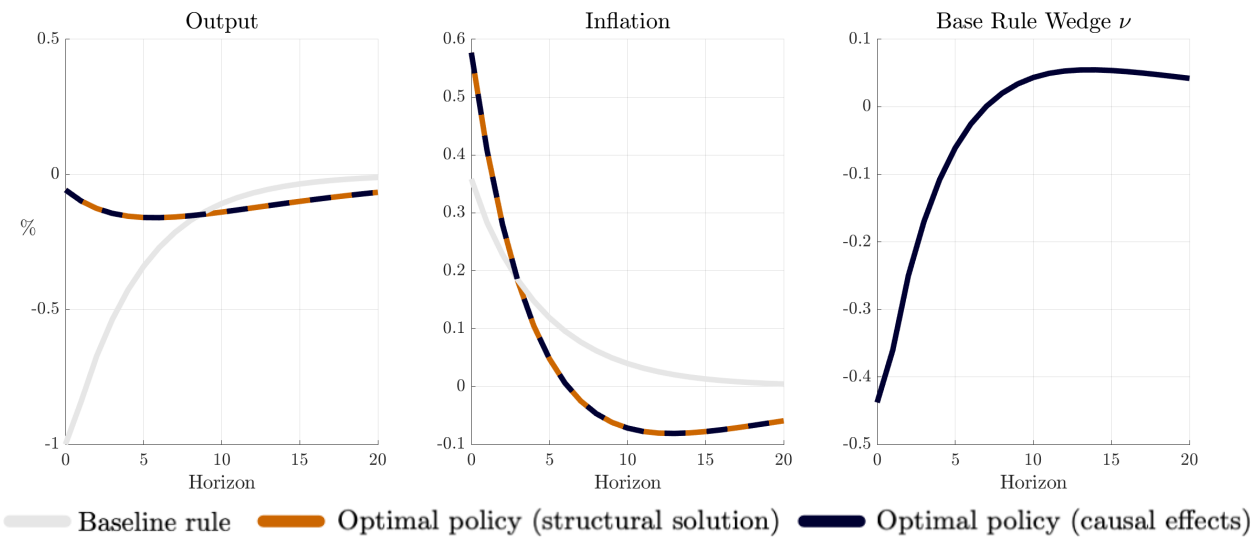


Figure 2: Output and inflation impulse responses together with the equivalence shock wedge $\boldsymbol{\nu}$ (see (22)) for the HANK model with policy rules (37) and the optimal policy given by (39). The impact output contraction under the prevailing baseline rule is normalized to -1%.

blue lines in the left and middle panels of Figure 1 show that, as expected, the solution is identical to that implied by re-solving the model. The right panel then shows the sequence of shocks $\boldsymbol{\nu}$ that maps the baseline prevailing rule into the new rule. Since the new rule is more accommodating, the sequence of shocks is persistently negative (expansionary).

Optimal policy. Our second experiment studies optimal policy under a dual mandate loss function

$$\mathcal{L} = \lambda_\pi \boldsymbol{\pi}'\boldsymbol{\pi} + \lambda_y \boldsymbol{y}'\boldsymbol{y} \tag{39}$$

with $\lambda_\pi = \lambda_y = 1$. We again start by solving for the optimal policy using conventional methods: we derive the policy rule corresponding to the first-order conditions (27) - (29), solve the model given that policy rule, and report the result as the orange lines in the left and middle panels of Figure 2. We see that, at the optimum, the cost-push shock moves inflation by much more than output, consistent with the assumed policy weights and the relatively flat Phillips curve. Compared to this optimal policy, the simple baseline rule of the form (37) tightens too much.

We then instead use Proposition 2 to equivalently recover the optimal policy rule and the corresponding cost-push shock impulse responses. We begin with the optimal rule itself.

By (35), the optimal rule is given as

$$\lambda_\pi \Theta'_{\pi,\nu,\mathcal{A}} \boldsymbol{\pi} + \lambda_y \Theta'_{y,\nu,\mathcal{A}} \boldsymbol{y} = 0$$

A researcher with knowledge of the dynamic causal effects of monetary policy shocks on inflation and output, $\{\Theta_{\pi,\nu,\mathcal{A}}, \Theta_{y,\nu,\mathcal{A}}\}$, is able to construct this optimal policy rule. We can then create a counterfactual response to the cost-push shock using (22), again requiring only knowledge of the causal effects of policy. As expected, the resulting impulse responses—the dark blue lines—are identical to those obtained by explicitly solving the optimal policy problem. Finally, the right panel of Figure 2 shows the optimal policy as a deviation $\boldsymbol{\nu}$ from the prevailing baseline rule.

2.5 Discussion

We have demonstrated that, in a quite general family of linearized macroeconomic models, the Lucas critique can in principle be circumvented through empirical measurement: if the econometrician can measure the (dynamic) causal effects of fundamental and policy shocks under the baseline rule, then she can construct a rich menu of policy counterfactuals. The simple intuition is that an econometrician observing impulse responses for all possible policy shocks in essence has access to valid instrumental variables to determine the effects of all possible paths for the policy instrument, thus allowing her to construct arbitrary counterfactuals or recover optimal policy rules.

In the remainder of this paper we present ways to operationalize this insight. The essential hurdle faced by our approach is that its informational requirements are extremely high: we would need evidence on the dynamic causal effects of the full menu of all possible contemporaneous and news policy shocks—evidence that is clearly not available, for any policy instrument. We will present two ways of dealing with this challenge. First, in Section 3, we show that, if a researcher was able to estimate the dynamic causal effects of a limited number of policy shocks, then she can still construct counterfactuals and find optimal rules in the subspace of changes to the policy rule spanned by those observed shocks. Second, in Section 4, we discuss strategies that use restrictions coming from economic theory to map the available partial evidence into the required full menu of dynamic causal effects. Throughout, we apply our results to analyze the aggregate effects of investment-specific technology shocks under counterfactual monetary policy rules.

3 Counterfactuals in identified subspaces

We consider the empirically relevant case of a researcher that is only able to estimate the dynamic causal effects of a finite (small) number of particular policy shocks. Formally, letting $\mathcal{V}$ be a linear map whose columns collect the responses of the policy shock $\boldsymbol{\nu}$ to $n_{\tilde{\nu}}$ identified sources of variation in policy, we assume that the researcher knows the dynamic causal effects of shocks to the baseline rule $\boldsymbol{\nu}$ that take the form $\boldsymbol{\nu} = \mathcal{V} \times \tilde{\boldsymbol{\nu}}$, where $\tilde{\boldsymbol{\nu}}$ is $n_{\tilde{\nu}}$-dimensional. We write those causal effects as $\{\Theta_{x,\tilde{\nu},\mathcal{A}}, \Theta_{z,\tilde{\nu},\mathcal{A}}\}$, where now both maps have a finite number of columns $n_{\tilde{\nu}}$, equal to weighted averages of the columns of the original maps $\{\Theta_{x,\nu,\mathcal{A}}, \Theta_{z,\nu,\mathcal{A}}\}$.[12] For example, a researcher may be able to estimate the effects of a couple of *particular* types of monetary shock (e.g., a transitory and a persistent interest rate change), but not those of *any possible* pattern of rate adjustments.

The remainder of this section proceeds as follows. First, in Section 3.1, we show that the estimated causal effects can still be used to construct counterfactuals for *some* alternative policy rules—rules that can be written as the prevailing baseline rule plus linear combinations of the identified shocks. Second, in Section 3.2, we find the *optimal* rule within the policy space spanned by the identified shocks. We illustrate both sets of counterfactuals with applications to monetary policy transmission, leveraging dynamic causal effects for two distinct monetary shocks: those of Christiano et al. (1999) and Gertler & Karadi (2015).

3.1 Counterfactual rules in identified subspaces

With the researcher observing $\{\Theta_{x,\tilde{\nu},\mathcal{A}}, \Theta_{z,\tilde{\nu},\mathcal{A}}\}$, the proof of Proposition 1 now only works for *particular* alternative policy rules that satisfy the restriction

$$\tilde{\mathcal{A}}_x(\boldsymbol{x}_{\mathcal{A}}(\boldsymbol{\varepsilon}) + \Theta_{x,\tilde{\nu},\mathcal{A}} \times \tilde{\boldsymbol{\nu}}) + \tilde{\mathcal{A}}_z(\boldsymbol{z}_{\mathcal{A}}(\boldsymbol{\varepsilon}) + \Theta_{z,\tilde{\nu},\mathcal{A}} \times \tilde{\boldsymbol{\nu}}) = 0$$

for some $\tilde{\boldsymbol{\nu}} \in \mathbb{R}^{n_{\tilde{\nu}}}$. That is, we must be able to replicate the alternative rule as the prevailing baseline rule plus some linear combination of the particular observed shocks, with weights given by $\tilde{\boldsymbol{\nu}}$. The larger $n_{\tilde{\nu}}$, the larger this identified subspace. If a researcher is interested in a rule outside of the spanned subspace, then one way forward is to find the best possible

[12] Our discussion in this section focusses on the finite-shock case, so $\{\Theta_{x,\tilde{\nu},\mathcal{A}}, \Theta_{z,\tilde{\nu},\mathcal{A}}\}$ have a small number of columns. In any empirical application, those linear maps of course also have a finite number T of rows. We do not pay much attention to this limitation since we consider shocks and counterfactual policies with sufficiently short-lived dynamics, making the maximal truncation horizon immaterial.

fit using the actually empirically observed shocks. For example, under a simple quadratic loss function for deviations from the (unattainable) target counterfactual policy rule, the best-fitting shock vector would be given as

$$\tilde{\boldsymbol{\nu}} = -\left[\left(\tilde{\mathcal{A}}_q \Theta_{q,\tilde{\nu},\mathcal{A}}\right)' \times \left(\tilde{\mathcal{A}}_q \Theta_{q,\tilde{\nu},\mathcal{A}}\right)\right]^{-1} \times \left[\left(\tilde{\mathcal{A}}_q \Theta_{q,\tilde{\nu},\mathcal{A}}\right)' \times \tilde{\mathcal{A}}_q \times \boldsymbol{q}_{\mathcal{A}}(\varepsilon)\right] \tag{40}$$

where $\boldsymbol{q} = (\boldsymbol{x}', \boldsymbol{z}')'$. Whether or not any given desired counterfactual rule is (at least approximately) contained within the space spanned by the empirically observed shocks is an inherently application-dependent question.

Application. We illustrate this discussion with an application to investment-specific technology shocks. Our object of interest is the behavior of output and inflation following such a technology shock and under a counterfactual monetary policy rule that aggressively stabilizes output fluctuations. We present the main results here, and relegate empirical implementation details to Appendix B.1.

Our approach requires two inputs. First, we need the dynamic causal effects of the shock of interest under the baseline policy rule. For this we use the investment-specific technology shock of Ben Zeev & Khan (2015). Second, we need the dynamic causal effects of some (ideally rich) menu of different monetary policy shocks. We consider two of the most popular examples of such shocks: the recursively identified shock of Christiano et al. (1999), and the high-frequency identification of Gertler & Karadi (2015). The dynamic response of nominal interest rates differs quite substantially across those two identifications schemes, as shown in Figure B.1: gradual and long-lived for Christiano et al., and relatively transitory for Gertler & Karadi. With these two estimates in hand, we can then follow (40) to construct the best possible approximation to a rule that aims to perfectly stabilize aggregate output.

Figure 3 presents our results. Under the prevailing baseline rule (grey), the policymaker leans against the inflationary effects of the negative technology shock, further pushing down aggregate real activity. Under our counterfactual rule (orange), on the other hand, interest rates are cut aggressively, keeping output relatively close to trend throughout, but of course at the cost of persistently elevated inflation. By our identification results, any structural model of the general form (13) - (14) and consistent with our empirical estimates of monetary transmission will agree with those policy counterfactuals.

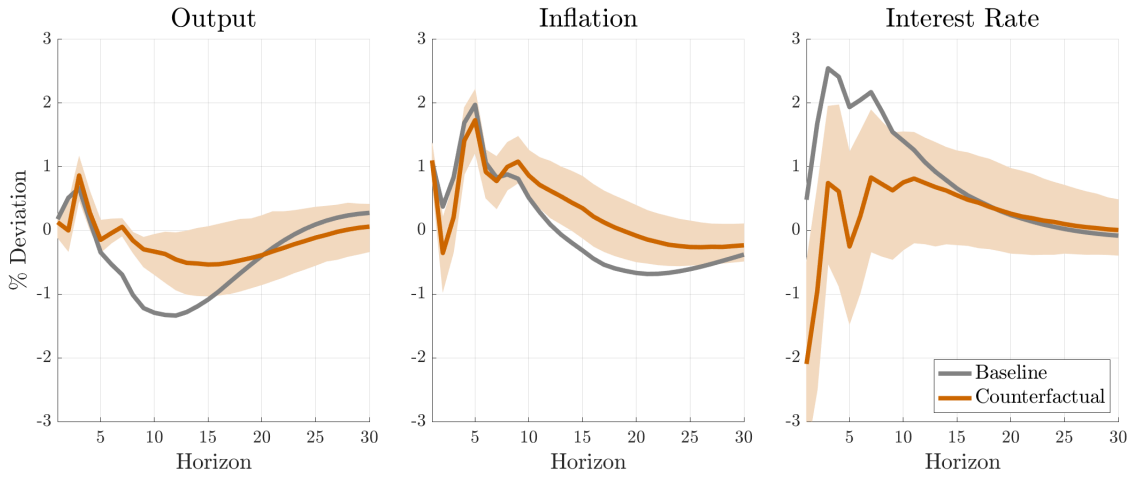


Figure 3: Output, inflation and interest rate impulse responses to a contractionary investment-specific technology shock under the prevailing baseline rule (grey) and the best feasible approximation to a rule that stabilizes output (orange). The shaded areas correspond to 16th and 84th percentile confidence bands.

3.2 Optimal policy in identified subspaces

For optimal policy, we follow the same steps as in the proof of Proposition 2 to now consider the problem of minimizing the policymaker loss function (20) within the identified subspace of policy changes. This problem gives the optimality condition

$$\sum_{i=1}^{n_x} \lambda_i \Theta'_{x_i,\tilde{\nu},\mathcal{A}} W \boldsymbol{x}_i = 0 \tag{41}$$

(41) can be interpreted in two ways. First, it gives $n_{\tilde{\nu}}$ restrictions that *any* solution to the full optimal policy problem must satisfy.[13] Second, it fully characterizes the optimal rule in the $n_{\tilde{\nu}}$-dimensional identified subspace of dynamic causal effects. The larger that space is, the more meaningful is the derived constrained optimal policy rule. In particular we by Proposition 2 know that, for $n_{\tilde{\nu}} \to \infty$, (41) fully characterizes the optimal policy rule.

[13]Equation (41) is related to Barnichon & Mesters (2020b), who propose to use a condition of this form to test the optimality of a given policy. Since their analysis relies on fixed private sector expectations, they do not draw any implications for optimal policy *rules*, unlike our approach.

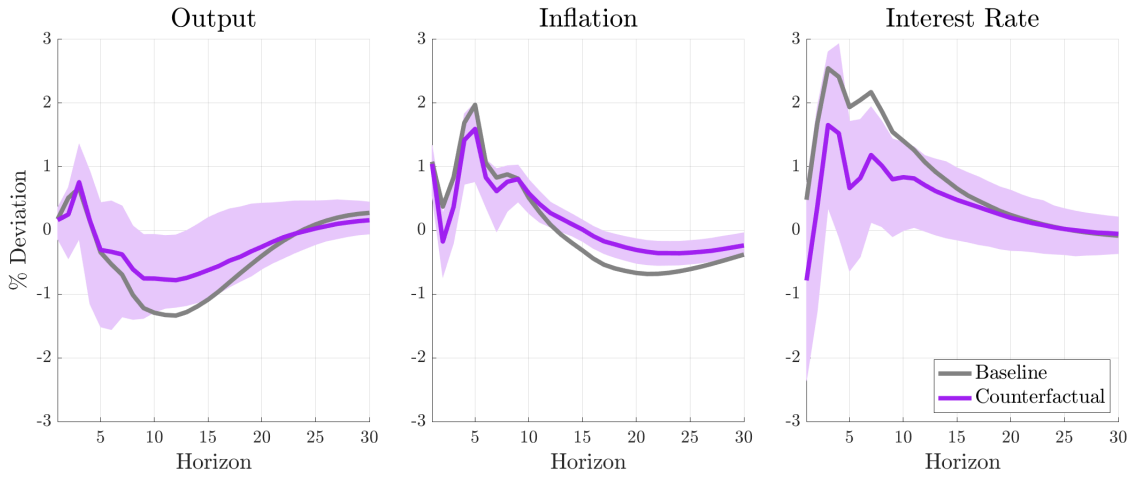


Figure 4: Output, inflation and interest rate impulse responses to a contractionary investment-specific technology shock under the prevailing baseline rule (grey) and the optimal policy rule for a dual-mandate policymaker ($\lambda_\pi = \lambda_y = 1$) within the identified subspace (purple). The shaded areas correspond to 16th and 84th percentile confidence bands.

Application. We illustrate our conclusions for optimal policy rules with another application to the investment-specific technology shocks discussed in Section 3.1. As before, our analysis leverages estimates of the dynamic causal effects of monetary policy based on Christiano et al. (1999) and Gertler & Karadi (2015). We consider a dual-mandate policymaker with equal weight on output and inflation ($\lambda_\pi = \lambda_y = 1$). We then use (41) to recover the constrained optimal policy rule in the identified subspace as well as the corresponding counterfactual paths of both the policy instrument as well as the two targets.

Figure 4 presents our results. Our analysis reveals that the (constrained) optimal policy rule increases interest rates somewhat less than the observed baseline policy response. As a result, the path for output is somewhat closer to trend than in the baseline and the deviation in inflation is somewhat smaller at longer horizons. On the whole, however, the differences between the implied optimal policy and the baseline policy are fairly small, suggesting that the observed policy response was close to optimal for an equal-weight dual-mandate policy.

4 Additional identifying information

Our second approach to implementing our method with limited evidence makes use of additional prior information to translate the dynamic causal effects of the policy shocks that

we do observe—the finite-shock maps $\{\Theta_{x,\tilde{\nu},\mathcal{A}}, \Theta_{z,\tilde{\nu},\mathcal{A}}\}$—into those that we did not observe—the rest of $\{\Theta_{x,\nu,\mathcal{A}}, \Theta_{z,\nu,\mathcal{A}}\}$. We are faced with a matrix completion problem: we know (or can estimate) certain linear combinations of the columns of the dynamic causal effect maps, and would like to learn about the rest. Our general approach to solving this problem is to impose (dogmatic) priors on the causal effect matrices, motivated by economic theory. Together with this prior structure, estimates of $\{\Theta_{x,\tilde{\nu},\mathcal{A}}, \Theta_{z,\tilde{\nu},\mathcal{A}}\}$ are enough to pin down the entirety of $\{\Theta_{x,\nu,\mathcal{A}}, \Theta_{z,\nu,\mathcal{A}}\}$.

We begin in Section 4.1 by showing that the conventional VAR-based approach of Sims & Zha (2006) can be interpreted as one example of such a dogmatic prior. This prior, however, turns out to be at odds with typical (rational-expectations) macroeconomic models. We thus in Section 4.2 present an alternative approach to eliciting a prior structure on the Θ's, and use it to revisit our applications of Sections 3.1 and 3.2.[14]

4.1 Counterfactuals as repeated surprises

Consider a researcher that, by estimating a conventional VAR, has estimated the dynamic causal effects of a contemporaneous policy shock ν_0 to the prevailing policy rule; that is, she knows the first column of the maps in $\{\Theta_{x,\nu,\mathcal{A}}, \Theta_{z,\nu,\mathcal{A}}\}$.

To construct counterfactuals for an alternative rule (19), Sims & Zha propose to subject the economy to a *sequence* of contemporaneous policy shocks such that (19) holds for all $t = 0, 1, 2, \ldots$. Translated to our notation, this approach is equivalent to imposing the dogmatic prior that the causal effect maps $\{\Theta_{x,\nu,\mathcal{A}}, \Theta_{z,\nu,\mathcal{A}}\}$ are lower-triangular, with the particular structure

$$\Theta_{q,\nu,\mathcal{A}} = \begin{pmatrix} \Theta_{q,\nu,\mathcal{A}}(1,1) & 0 & 0 & \ldots \\ \Theta_{q,\nu,\mathcal{A}}(2,1) & \Theta_{q,\nu,\mathcal{A}}(1,1) & 0 & \ldots \\ \Theta_{q,\nu,\mathcal{A}}(3,1) & \Theta_{q,\nu,\mathcal{A}}(2,1) & \Theta_{q,\nu,\mathcal{A}}(1,1) & \ldots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \quad q \in \{x, z\} \tag{42}$$

where $\Theta(i,j)$ denotes the (i,j)th entry of a map Θ. In words, the dynamic causal effects of news shocks are simply time-shifted versions of the impulse responses for a contemporaneous shock, with no effect prior to the shock's realization. Intuitively, if this is so, then surprising the economy with a suitable new shock each period—the approach of Sims & Zha (2006)—is

[14] For a third example of a possible prior on the Θ's see Appendix A.5.

the same as announcing a sequence of contemporaneous and news shocks at $t = 0$—our identification result.

Equation (42) is an approach to cross-column impulse response extrapolation that is valid in any model with fully myopic agents. For example, in a variant of the behavioral New Keynesian model of Gabaix (2020) with full discounting in both the consumer Euler equation and the firm-side Phillips curve, news shocks have no effect prior to their realization, so (42) holds. Typical (rational-expectations) macroeconomic models with forward-looking agents, however, are inconsistent with this prior.

4.2 A prior on the output-inflation space

Our objective is to find a prior on the Θ's that is informative enough to pin down *all* causal effects from empirical evidence on just a few shocks (just like the approach of Sims & Zha), yet consistent with a large family of structural models. In this section we present one example of such a prior: the restrictions on output and inflation co-movements implied by a standard Phillips curve relationship. We first discuss the general structure that such a relationship imposes on the causal effect maps $\Theta_{\pi,\nu,\mathcal{A}}$ and $\Theta_{y,\nu,\mathcal{A}}$, and then leverage this prior to revisit the applications of Sections 3.1 and 3.2.

Dynamic Phillips curve restrictions. Many structural macroeconomic models feature a Phillips curve—a link between inflation and leads and lags of output (or the output gap). Using our perfect-foresight notation of Section 2, we can write a general Phillips curve relationship as

$$\boldsymbol{\pi} \;=\; \Pi_y \;\times\; \boldsymbol{y} \;+\; \Pi_\varepsilon \;\times\; \boldsymbol{\varepsilon} \tag{43}$$

Here Π_y is the linear map summarizing the structural relationship between inflation and leads and lags of output, up to (non-policy) shocks $\Pi_\varepsilon \times \boldsymbol{\varepsilon}$. For example, in the New Keynesian model of Section 2.1, Π_y would take the particular form

$$\Pi_y = \begin{pmatrix} \kappa & \kappa\beta & \kappa\beta^2 & \dots \\ 0 & \kappa & \kappa\beta & \dots \\ 0 & 0 & \kappa & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

The crucial implication of (43) is that, conditional on policy shocks $\boldsymbol{\nu}$, the co-movements of output and inflation are fully characterized by the map Π_y:

$$\Theta_{\pi,\nu,\mathcal{A}} = \Pi_y \times \Theta_{y,\nu,\mathcal{A}}$$

In words, we can map output into inflation impulse responses (and vice-versa) using only Π_y. It then follows from our discussion of relative impulse responses in Section 2.3 that knowledge of Π_y is sufficient to[15] (a) construct counterfactuals for alternative policy rules of the general form

$$\tilde{\mathcal{A}}_\pi \boldsymbol{\pi} + \tilde{\mathcal{A}}_y \boldsymbol{y} = 0, \tag{44}$$

and (b) find the optimal policy rule for a dual mandate policymaker with preferences

$$\mathcal{L} = \lambda_\pi \boldsymbol{\pi}'\boldsymbol{\pi} + \lambda_y \boldsymbol{y}'\boldsymbol{y} \tag{45}$$

Importantly, this conclusion holds independently of any further model details, including preferences, technology, the nature of expectation formation, and so on—as long as two structural models agree on the Phillips curve map Π_y they also invariably agree on the desired policy counterfactuals. We have thus reduced the problem of extrapolating across columns of the two maps $\Theta_{\pi,\nu,\mathcal{A}}$ and $\Theta_{y,\nu,\mathcal{A}}$ to one of learning only about the single map Π_y.

Our preferred approach to estimation of Π_y proceeds as follows. First, we use economic theory to motivate a particular (dogmatic) prior for Π_y. Second, we use that prior in conjunction with empirical evidence on individual shocks—the individual columns $\{\Theta_{\pi,\tilde{\nu},\mathcal{A}}, \Theta_{y,\tilde{\nu},\mathcal{A}}\}$—to estimate Π_y.[16] In the next subsection we operationalize this approach to revisit our application to monetary policy and investment-specific technology shocks.

Application. We use our insights regarding the centrality of the Phillips curve relationship to construct the same two investment-specific technology shock policy counterfactuals as in Sections 3.1 and 3.2, but now using the extrapolated full map Π_y rather than the limited identified subspaces of causal effects.

[15] Strictly speaking, we additionally require the assumption of invertibility of $\Theta_{\pi,\nu,\mathcal{A}}$—that is, the policymaker can implement any possible path of inflation. This assumption is generally satisfied in standard business-cycle models. For example, in the simple model of Section 2.1, it is straightforward to verify that $\Theta_{\pi,\nu,\mathcal{A}}$ is an upper-triangular, invertible matrix. We provide further details in Appendix A.4.

[16] Equivalently, this approach may be interpreted as impulse response matching *for one model block*. Our key insight is that, for some interesting policy counterfactuals, it is actually not necessary to specify the entire model—only the implied prior on the output-inflation relationship matters.

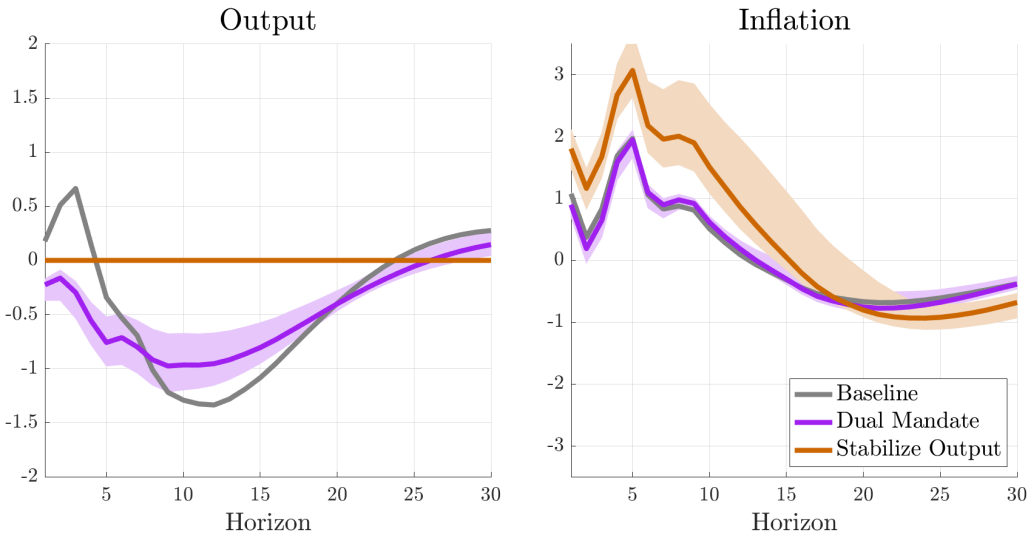


Figure 5: Output and inflation impulse responses to a contractionary investment-specific technology shock under the prevailing baseline rule (grey), a counterfactual rule that perfectly stabilizes output (grey), and the optimal policy rule for a dual-mandate policymaker ($\lambda_\pi = \lambda_y = 1$) (purple). The shaded areas correspond to 16th and 84th percentile confidence bands.

Our dogmatic prior on Π_y takes the form of a hybrid Phillips curve relationship:

$$\pi_t = \gamma_b \pi^4_{t-1} + \gamma_f \mathbb{E}_t \left[\pi^4_{t+4}\right] + \kappa y_t + \varepsilon_t \tag{46}$$

where $\pi^4_{t-1} = \frac{1}{4} \times (\pi_{t-1} + \pi_{t-2} + \pi_{t-3} + \pi_{t-4})$. Appendix A.4 shows the linear map Π_y corresponding to this Phillips curve specification. We then estimate the parameters $\{\gamma_b, \gamma_f, \kappa\}$ (and so all of Π_y) using evidence on identified monetary policy shocks. The econometric challenge is that estimates of $\{\Theta_{\pi,\tilde{\nu},\mathcal{A}}, \Theta_{y,\tilde{\nu},\mathcal{A}}\}$ will not perfectly align with the parametric structure imposed by (46); thus, following Barnichon & Mesters (2020a), we simply find the best possible fit. Our estimation uses the identified monetary policy shocks of Gertler & Karadi (2015), already discussed in Section 3.

Given an estimate of Π_y, we can construct the two desired counterfactuals: output and inflation responses to investment-specific technology shocks under counterfactual policy rules that (a) perfectly stabilize output and (b) are optimal for a dual mandate policymaker with equal weights. The results, reported in Figure 5, echo those of Section 3.[17] First,

[17]Note that Π_y is sufficient to characterize the (relative) implementable space of output and inflation, but does not allow us to solve for the nominal interest rate path that is required to engineer those allocations.

perfect output stabilization implies persistently elevated inflation relative to the baseline rule outcome. Second, the output and inflation impulse response paths under the optimal dual-mandate policy are relatively close to observed outcomes, but with somewhat smoother output dynamics. With a Phillips curve of the form (46), we can by Proposition 2 in fact explicitly characterize this optimal dual mandate policy rule as

$$\lambda_\pi \pi_t + \lambda_y \frac{1}{\kappa} \left[y_t - \frac{1}{4}\gamma_b \sum_{h=1}^{4} y_{t+h} - \frac{1}{4}\gamma_f \sum_{h=1}^{4} y_{t-h} \right] = 0, \quad \forall t = 0, 1, 2, \ldots \tag{47}$$

It follows from our analysis that any structural model that (i) fits into the general form (13) - (14), (ii) features a Phillips curve relationship of the form (46) and (iii) is consistent with the empirical monetary policy shock estimates of Gertler & Karadi (2015) will produce the same counterfactuals as in Figure 5, and yield the optimal dual-mandate policy rule (47).

5 Conclusions

The standard approach to counterfactual policy analysis relies on fully-specified, dynamic general equilibrium models. Our identification results instead point in a different direction: for valid policy counterfactuals, researchers can estimate dynamic causal effects of policy shocks and combine them to form policy counterfactuals and optimal policy responses. These counterfactuals are valid in a large class of models that encompasses the majority of structural business-cycle models that are currently used for policy analysis.

The main challenge in implementing this strategy is that evidence on an infinitely large set of policy perturbations is unattainable. One natural reaction is simply to try to get as close as possible; viewed in this light, further empirical evidence on policy news shocks would be extremely valuable. Every additional piece of empirical evidence will allow researchers to (a) expand the space of alternative, counterfactual policy rules that we can analyze and (b) find further restrictions on optimal policy rules. The obvious alternative path forward is to use outside (probabilistic or dogmatic) prior information to map the available limited empirical evidence into estimates of the entire space of policy-relevant causal effects. We have presented an example of this approach, and constructed counterfactuals under a dogmatic Phillips curve prior. Refining and extending this approach remains an important avenue for further research.

References

Auclert, A., Bardóczy, B., Rognlie, M., & Straub, L. (2021). Using the sequence-space jacobian to solve and estimate heterogeneous-agent models. Working Paper.

Auclert, A., Rognlie, M., & Straub, L. (2020). Micro jumps, macro humps: Monetary policy and business cycles in an estimated hank model. Technical report, National Bureau of Economic Research.

Barnichon, R. & Mesters, G. (2020a). Identifying modern macro equations with old shocks. *The Quarterly Journal of Economics*, *135*(4), 2255–2298.

Barnichon, R. & Mesters, G. (2020b). Testing Macroeconomic Policies with Sufficient Statistics. Working Paper.

Barsky, R. B., Basu, S., & Lee, K. (2015). Whither news shocks? *NBER Macroeconomics Annual*, *29*(1), 225–264.

Baxter, M. & King, R. G. (1993). Fiscal policy in general equilibrium. *The American Economic Review*, *83*(3), 315–334.

Ben Zeev, N. & Khan, H. (2015). Investment-specific news shocks and us business cycles. *Journal of Money, Credit and Banking*, *47*(7), 1443–1464.

Beraja, M. (2020). Counterfactual equivalence in macroeconomics. Working Paper.

Bernanke, B. S., Gertler, M., Watson, M., Sims, C. A., & Friedman, B. M. (1997). Systematic monetary policy and the effects of oil price shocks. *Brookings papers on economic activity*, *1997*(1), 91–157.

Boppart, T., Krusell, P., & Mitman, K. (2018). Exploiting MIT shocks in heterogeneous-agent economies: the impulse response as a numerical derivative. *Journal of Economic Dynamics and Control*, *89*, 68–92.

Christiano, L. J., Eichenbaum, M., & Evans, C. L. (1999). Monetary policy shocks: What have we learned and to what end? *Handbook of macroeconomics*, *1*, 65–148.

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

Eberly, J. C., Stock, J. H., & Wright, J. H. (2019). The federal reserve's current framework for monetary policy: A review and assessment. NBER Working Paper.

Fernández-Villaverde, J., Rubio-Ramírez, J. F., & Schorfheide, F. (2016). Solution and estimation methods for DSGE models. In *Handbook of Macroeconomics*, volume 2 (pp. 527–724). Elsevier.

Gabaix, X. (2020). A behavioral new keynesian model. *American Economic Review*, *110*(8), 2271–2327.

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

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

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

Guren, A., McKay, A., Nakamura, E., & Steinsson, J. (2021). What do we learn from cross-regional empirical estimates in macroeconomics? *NBER Macroeconomics Annual*, *35*(1), 175–223.

Hebden, J. & Winker, F. (2021). Impulse-Based Computation of Policy Counterfactuals. Working Paper.

Kaplan, G., Moll, B., & Violante, G. (2018). Monetary Policy according to HANK. *American Economic Review*, *108*(3), 697–743.

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

McKay, A. & Wieland, J. F. (2021). Lumpy durable consumption demand and the limited ammunition of monetary policy. Forthcoming in Econometrica.

Ottonello, P. & Winberry, T. (2020). Financial heterogeneity and the investment channel of monetary policy. *Econometrica*, *88*(6), 2473–2502.

Plagborg-Møller, M. & Wolf, C. K. (2021). Local projections and vars estimate the same impulse responses. *Econometrica*, *89*(2), 955–980.

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

Romer, C. D. & Romer, D. H. (2004). A new measure of monetary shocks: Derivation and implications. *American Economic Review*, *94*(4), 1055–1084.

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

Sims, C. A. (1987). A rational expectations framework for short-run policy analysis. In Barnett, W. A. & Singleton, K. J. (Eds.), *New approaches to monetary economics*, (pp. 293–308)., Cambridge, UK. Cambridge University Press.

Sims, C. A. & Zha, T. (2006). Does monetary policy generate recessions? *Macroeconomic Dynamics*, *10*(2), 231–272.

Smets, F. & Wouters, R. (2007). Shocks and frictions in us business cycles: A bayesian dsge approach. *American economic review*, *97*(3), 586–606.

Svensson, L. E. (1997). Inflation forecast targeting: Implementing and monitoring inflation targets. *European economic review*, *41*(6), 1111–1146.

Wolf, C. K. (2020). The missing intercept: A demand equivalence approach. Working Paper.

Wolf, C. K. (2021). Interest rate cuts vs. stimulus payments: An equivalence result. Working Paper.

Woodford, M. (2011). *Interest and prices: Foundations of a theory of monetary policy*. princeton university press.

A Further results

This appendix presents several supplementary theoretical results. Appendix A.1 extends Proposition 2 to a more general loss function. Appendix A.2 then expresses the simple three-equation model in our general linearized perfect foresight notation, Appendix A.3 gives further details for the HANK model of Sections 2.2 and 2.4, and Appendix A.4 provides additional information for our dogmatic Phillips curve prior. Finally in Appendix A.5 we present a third approach to the matrix completion problem.

A.1 More general loss functions

Proposition 2 can be generalized to allow for a non-separable quadratic loss function. Suppose the policymaker's loss function takes the form

$$\mathcal{L} = \boldsymbol{x}' Q \boldsymbol{x} \tag{A.1}$$

where Q is a weighting matrix. Following the same steps as the proof of Proposition 2, we can formulate the policy problem as minimizing the loss function (A.1) subject to (34). The first-order conditions of this problem are

$$\Theta'_{\nu,x,A}(Q + Q')\boldsymbol{x} = 0$$

so we can recover the optimal policy rule as

$$\begin{aligned} \mathcal{A}^*_x &= \Theta'_{\nu,x,A}(Q + Q') \\ \mathcal{A}^*_z &= \mathbf{0} \end{aligned}$$

Outside of the quadratic case, the causal effects of policy shocks on $\boldsymbol{x}$ are still enough to formulate a set of necessary conditions for optimal policy, but in this general case the resulting optimal policy rule will not fit into the linear form (14), so we do not consider this case here.

A.2 Linear maps for the canonical New Keynesian model

We begin with the non-policy block. The Phillips curve can be written as

$$\begin{pmatrix} 1 & -\beta & 0 & \dots \\ 0 & 1 & -\beta & \dots \\ 0 & 0 & 1 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \boldsymbol{\pi} - \kappa \boldsymbol{y} - \boldsymbol{\varepsilon}^s = 0,$$

while the Euler equation can be written as

$$-\sigma \begin{pmatrix} 0 & 1 & 0 & \dots \\ 0 & 0 & 1 & \dots \\ 0 & 0 & 0 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \boldsymbol{\pi} + \begin{pmatrix} 1 & -1 & 0 & \dots \\ 0 & 1 & -1 & \dots \\ 0 & 0 & 1 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \boldsymbol{y} + \sigma \boldsymbol{i} = 0.$$

These linear maps can be stacked into the form (13). Finally the policy rule can be written as

$$\phi_\pi \boldsymbol{\pi} - \boldsymbol{i} + \boldsymbol{\nu} = 0,$$

which fits into the form of (14).

A.3 Model details & parameterization

The HANK model sketched in Section 2.2 and used for our quantitative illustration in Section 2.4 is exactly the same as in Wolf (2021) (including the parameterization), with only one change: rather than imposing uniform hours worked $\ell_{it} = \ell_t$ for all households i, we consider a labor rationing rule that ensures that

$$w_t \ell_{it} e_{it} + d_{it} = e_{it} y_t$$

for all households i. This rationing rule is feasible since $\int_0^1 w_t \ell_{it} e_{it} + \int_0^1 d_{it} = w_t \ell_t + d_t = y_t$, and it allows us to write the perfect foresight consumer demand block as

$$\boldsymbol{c} = \mathcal{C}\left(\boldsymbol{y}, \boldsymbol{\pi}, \boldsymbol{i}, \boldsymbol{\tau}, \boldsymbol{\varepsilon}^d\right)$$

Linearizing, we recover (16).

A.4 Phillips curve & policy counterfactuals

Consider the augmented Phillips curve (46). Along a perfect foresight transition path, we can write this relationship as

$$\underbrace{\begin{pmatrix} 1 & -\frac{1}{4}\gamma_f & -\frac{1}{4}\gamma_f & -\frac{1}{4}\gamma_f & -\frac{1}{4}\gamma_f & 0 & \dots \\ -\frac{1}{4}\gamma_b & 1 & -\frac{1}{4}\gamma_f & -\frac{1}{4}\gamma_f & -\frac{1}{4}\gamma_f & -\frac{1}{4}\gamma_f & \dots \\ -\frac{1}{4}\gamma_b & -\frac{1}{4}\gamma_b & 1 & -\frac{1}{4}\gamma_f & -\frac{1}{4}\gamma_f & -\frac{1}{4}\gamma_f & \dots \\ -\frac{1}{4}\gamma_b & -\frac{1}{4}\gamma_b & -\frac{1}{4}\gamma_b & 1 & -\frac{1}{4}\gamma_f & -\frac{1}{4}\gamma_f & \dots \\ -\frac{1}{4}\gamma_b & -\frac{1}{4}\gamma_b & -\frac{1}{4}\gamma_b & -\frac{1}{4}\gamma_b & 1 & -\frac{1}{4}\gamma_f & \dots \\ 0 & -\frac{1}{4}\gamma_b & -\frac{1}{4}\gamma_b & -\frac{1}{4}\gamma_b & -\frac{1}{4}\gamma_b & 1 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}}_{\equiv \Pi_\pi} \times \boldsymbol{\pi} \quad = \quad \kappa \times \boldsymbol{y}$$

We thus have

$$\Pi_y \quad \equiv \quad \Pi_\pi^{-1} \quad \times \quad \kappa$$

Policy counterfactuals. Knowledge of Π_y—together with the assumption that $\Theta_{\pi,\nu,\mathcal{A}}$ is invertible, i.e., any path of inflation is in principle implementable—is sufficient to construct several interesting counterfactuals. First, for a counterfactual policy rule of the form (44), we can recover counterfactual outcomes by solving the system

$$\begin{aligned} \tilde{\mathcal{A}}_\pi \boldsymbol{\pi} + \tilde{\mathcal{A}}_y \boldsymbol{y} &= 0 \\ \boldsymbol{\pi} &= \boldsymbol{\pi}_{\mathcal{A}}(\boldsymbol{\varepsilon}) + \boldsymbol{\nu} \\ \boldsymbol{y} &= \boldsymbol{y}_{\mathcal{A}}(\boldsymbol{\varepsilon}) + \Pi_y^{-1} \boldsymbol{\nu} \end{aligned}$$

for the three unknowns $\{\boldsymbol{\pi}, \boldsymbol{y}, \boldsymbol{\nu}\}$. Second, for a policymaker loss function of the form (45), the optimal policy rule is given as

$$\lambda_\pi \boldsymbol{\pi} + \lambda_y (\Pi_y')^{-1} \boldsymbol{y} = 0$$

Solving out this equation we find that, for all $t \geq 0$, (47) must hold, as claimed.

Invertibility of $\Theta_{\pi,\nu,\mathcal{A}}$. Strictly speaking, our results leveraging Π_y impose the additional assumption that the monetary policymaker can in principle implement any desired path of inflation. This assumption is routinely satisfied in standard business-cycle models.

For example, in our simple model of Section 2.1, it is straightforward to verify that $\Theta_{\pi,\nu,\mathcal{A}}$ is an upper-triangular matrix with

$$\Theta_{\pi,\nu,\mathcal{A}}(i,i) = -\frac{\kappa\sigma}{1+\kappa\sigma\phi_\pi}$$

and $\Theta_{\pi,\nu,\mathcal{A}}(i,j)$ for $i<j$ defined recursively via the system

$$\begin{aligned}\Theta_{y,\nu,\mathcal{A}}(i,j) &= -\sigma(\phi_\pi\Theta_{\pi,\nu,\mathcal{A}}(i,j) - \Theta_{\pi,\nu,\mathcal{A}}(i+1,j)) + \Theta_{y,\nu,\mathcal{A}}(i+1,j)\\ \Theta_{\pi,\nu,\mathcal{A}}(i,j) &= \kappa\Theta_{y,\nu,\mathcal{A}}(i,j) + \beta\Theta_{\pi,\nu,\mathcal{A}}(i+1,j)\end{aligned}$$

A.5 Leveraging time invariance

This section presents a third possible approach to the matrix completion problem of Section 4, based upon a property of the causal effect maps $\{\Theta_{x,\nu,\mathcal{A}}, \Theta_{z,\nu,\mathcal{A}}\}$ that we refer to as asymptotic time invariance.

Definition. Asymptotic time invariance formalizes the idea that the different columns of the causal effect maps are not completely independent objects – for example, impulse responses to a forward guidance shock eight quarters out should be very similar to forward guidance shocks nine quarters out, just shifted by one period. The precise definition is that, for all $s \in \mathbb{N}$,

$$\lim_{t\to\infty} \Theta_{x,\nu,\mathcal{A}}(t+s,t) = \bar{\Theta}_{x,\nu,\mathcal{A}}(s), \quad \lim_{t\to\infty} \Theta_{z,\nu,\mathcal{A}}(t+s,t) = \bar{\Theta}_{z,\nu,\mathcal{A}}(s) \tag{A.2}$$

where $\bar{\Theta}_{x,\nu,\mathcal{A}}$ and $\bar{\Theta}_{z,\nu,\mathcal{A}}$ are two sequences. Figure A.1 provides an illustration of this property in the quantitative HANK model of Section 2.4, showing output impulse responses to various different contemporaneous and forward guidance monetary shocks. We see that, for forward guidance shocks far into the future (large shock horizon h), the output impulse responses are left- and right-translations of each other, exactly as expected.

Time invariance as a dogmatic prior. Imposing (A.2) after some finite horizon H reduces the problem of dynamic causal effect identification from an infinite-dimensional one to an $H+1$-dimensional one. For example, imposing time invariance from horizon $H=0$

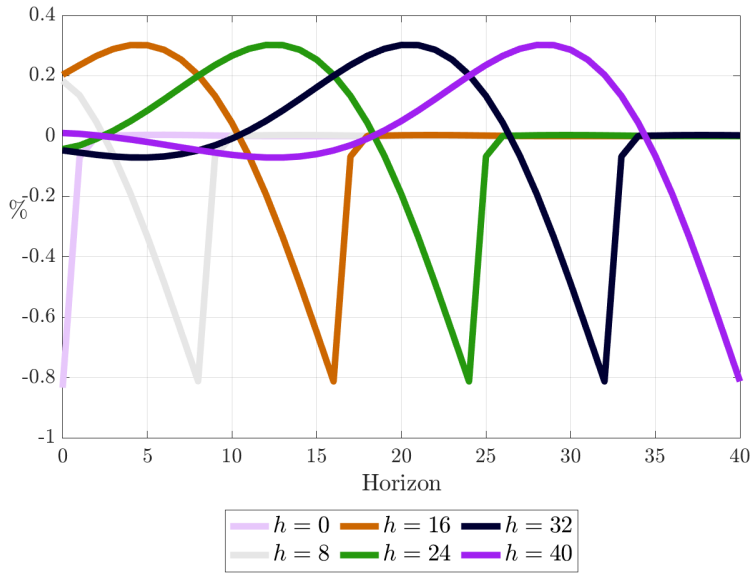


Figure A.1: Output impulse responses to contemporaneous and forward guidance monetary policy shocks in the HANK model of Appendix A.3.

onwards implies that the causal effect matrices have the following particular structure:

$$\Theta_{q,\nu,\mathcal{A}} = \begin{pmatrix} \bar{\Theta}_{q,\nu,\mathcal{A}}(0) & \bar{\Theta}_{q,\nu,\mathcal{A}}(-1) & \bar{\Theta}_{q,\nu,\mathcal{A}}(-2) & \dots \\ \bar{\Theta}_{q,\nu,\mathcal{A}}(1) & \bar{\Theta}_{q,\nu,\mathcal{A}}(0) & \bar{\Theta}_{q,\nu,\mathcal{A}}(-1) & \dots \\ \bar{\Theta}_{q,\nu,\mathcal{A}}(2) & \bar{\Theta}_{q,\nu,\mathcal{A}}(1) & \bar{\Theta}_{q,\nu,\mathcal{A}}(0) & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \quad q \in \{x, z\} \tag{A.3}$$

The sequence $\bar{\Theta}_{q,\nu,\mathcal{A}}$ can be estimated using empirical evidence on a forward guidance shock sufficiently far into the future. (A.3) then provides the mapping from the sequence $\bar{\Theta}_{q,\nu,\mathcal{A}}$ into the entire causal effect map.

B Empirical appendix

This appendix elaborates on our empirical estimation. First, in Appendix B.1, we discuss the causal effects estimated for the counterfactuals in Section 3. Second, in Appendix B.2, we briefly review the approach of Barnichon & Mesters (2020a) to Phillips curve estimation, used for our counterfactuals in Section 4.

B.1 Shock & policy dynamic causal effects

Our analysis of investment-specific technology shocks follows Ben Zeev & Khan (2015), while our monetary policy shock identification mimics that of (i) Christiano et al. (1999) and (ii) Gertler & Karadi (2015).

Outcomes. We are interested in impulse responses of three outcome variables: output, inflation, and the policy rate. For output, we follow Ramey (2016) and deflate per-capita nominal GDP by the GDP deflator. For inflation, we consider the annualized growth rate of the GDP deflator. All series are obtained from the replication files for Ramey (2016). Finally, we consider the federal funds rate as our measure of the policy rate, obtained from the St. Louis Federal Reserve FRED database. All series are quarterly.

Shocks & identification. We take the investment-specific technology shock series from Ben Zeev & Khan (2015) and the high-frequency monetary policy surprise series from Gertler & Karadi (2015), aggregated to quarterly frequency through simple averaging. Recursive shocks are identified through the estimated VAR itself.

Estimation details. For maximal consistency, we try to estimate all impulse responses within a common empirical specification. For the investment-specific technology shocks, we order the shock measure first in a recursive VAR containing our outcomes of interest (following Plagborg-Møller & Wolf, 2021), estimated on a sample from 1969:Q1–2007:Q4. Our estimation of the Gertler-Karadi shock is identical, except for the fact that—because of constraints on when the shock is actually available—the sample runs from 1990:Q1 – 2012:Q4. Finally, for a recursive shock, we return to our baseline long sample period, and now identify a monetary shock as the *last* recursively ordered shock in a system containing output, inflation, and the nominal rate.

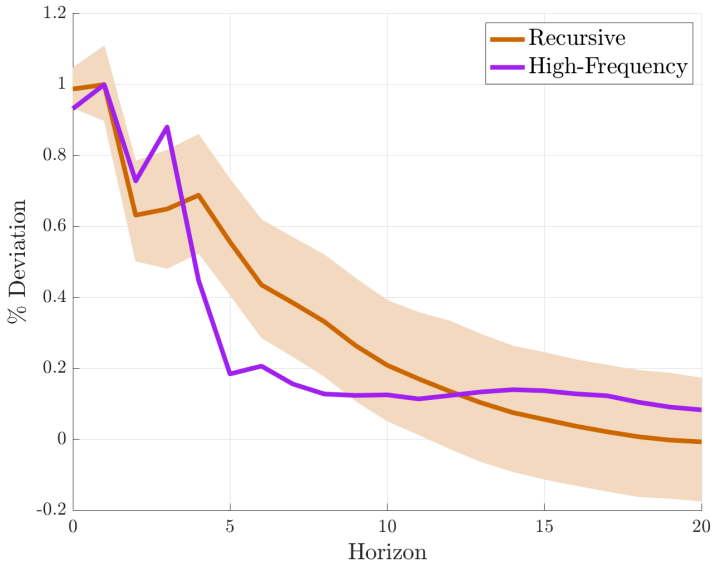


Figure B.1: Federal funds rate impulses to the recursive (orange) and high-frequency (purple) monetary policy shocks of Christiano et al. (1999) and Gertler & Karadi (2015), with the peak impulse normalized to 1 per cent.

We estimate all VARs using four lags, a constant, and deterministic linear and quadratic trends. For the baseline investment-specific technology shock we fix the OLS point estimates. We then construct policy counterfactuals using our identified monetary policy shocks, taking into account their estimation uncertainty. To do so we separately draw from the different monetary policy model posteriors and then compute the counterfactuals for each draw, thus effectively imposing independence across the estimated VARs.

Interest rate paths. Figure B.1 shows impulse responses of the federal funds rate to the two estimated monetary policy shocks. Consistent with previous work, we find that the recursively identified shock induces much more persistent interest rate movements than a shock identified via high-frequency surprises.[18]

[18]The third well-known example of an identified monetary policy shock—that of Romer & Romer (2004)—induces interest rate movements very similar to our recursively identified shock, so it adds little to our construction of policy counterfactuals.

B.2 NKPC estimation

Barnichon & Mesters (2020a) show how to use estimates of monetary policy impulse responses to identify a Phillips curve relationship of the form (46). For our empirical analysis in Section 4.2 we rely on the point estimates and the confidence region corresponding to their Gertler & Karadi analysis (which imposes the additional constraint that $\gamma_f + \gamma_b = 1$), reported in Table IV and Figure V of their paper.