Lecture 7: From Policy Shocks to Policy Rules

Christian Wolf

MIT

14.461, Fall 2022

Overview

- Next two lectures will be about our remaining two **substantive questions**:

  2. How should we design business-cycle stabilization policy? How does the choice of policy rule affect the propagation of macro shocks?
  3. What are the sources of business-cycle fluctuations?

- For each we will proceed in **two steps**:

  a) How far can we get with our semi-structural methods alone?
  b) How can we use macro data together with explicit structural models (rather than just the SVMA) to answer 1. & 2.?

- This lecture: predicting the effects of changes in **policy rules**

Outline

1. Policy Shocks vs. Policy Rules
   - Explorations with Sims-Zha
   - A Generalized Sims-Zha Identification Result
   - Practical Implications
2. Application I: Counterfactual Monetary Policy Paths
3. Application II: Optimal Monetary Policy in NK Models
   - Dual Mandate Policymaker
   - Adding distributional objectives
4. Summary

Outline

1. Policy Shocks vs. Policy Rules
    - Explorations with Sims-Zha
    - A Generalized Sims-Zha Identification Result
    - Practical Implications

2. Application I: Counterfactual Monetary Policy Paths

3. Application II: Optimal Monetary Policy in NK Models
    - Dual Mandate Policymaker
    - Adding distributional objectives

4. Summary

A brief history of thought

- How to predict the effects of changes in **policy rules**?
- Two important **methodological approaches**, both heavily use policy shocks:

1. **"Lucas program"**: use fully specified parametric GE model

   See Christiano et al. (1999) for a detailed presentation of this approach. Role of policy shocks: estimation target in IRF matching (e.g. Christiano et al. 2005). Will return to this at the end.

   → Reason for popularity: robustness to Lucas critique

   → Obvious challenge: vulnerability to model mis-specification

2. **Sims-Zha (1995)**: construct policy rule counterfactuals without relying on a model, using *only* identified policy shocks

   → We will focus on this. First review the method & then understand through the lens of models why it's not robust to the Lucas critique.

   → Finally: we'll develop a fix that is consistent with the Lucas critique (under some asn's)

Outline

1. Policy Shocks vs. Policy Rules
   - Explorations with Sims-Zha
   - A Generalized Sims-Zha Identification Result
   - Practical Implications
2. Application I: Counterfactual Monetary Policy Paths
3. Application II: Optimal Monetary Policy in NK Models
   - Dual Mandate Policymaker
   - Adding distributional objectives
4. Summary

The Sims-Zha approach

- The counterfactual policy question
  - Suppose you observe some non-policy shock (e.g., supply) and you estimate its effects on output, inflation, and interest rates $\{\boldsymbol{y}^s, \boldsymbol{\pi}^s, \boldsymbol{i}^s\}$
  - Note that those estimates were generated under the actual policy rule. **Q**: how would this shock have propagated under an alternative rule? Simple example for this slide: $i_t = \tilde{\phi} \times \pi_t$
- **Sims-Zha**: enforce counterfactual rule with a **sequence of policy shocks**
  - Needed input: causal effects of a policy shock to the actual rule $\{\boldsymbol{y}^m, \boldsymbol{\pi}^m, \boldsymbol{i}^m\}$
  - Now choose a policy shock $\boldsymbol{\nu_0^m}$ at date 0 to enforce the **counterfactual rule**:

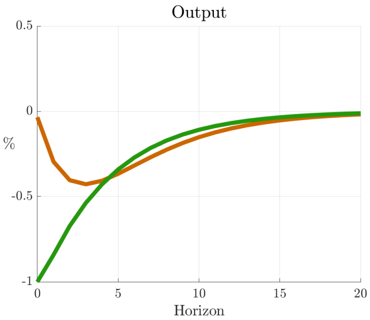
$$i_0^s + \nu_0^m \times i_0^m = \tilde{\phi} \times [\pi_0^s + \nu_0^m \times \pi_0^m]$$

Then iteratively continue for all $t = 1, 2, \ldots$. For $t = 1$:

$$i_1^s + \nu_0^m \times i_1^m + \nu_1^m \times i_0^m = \tilde{\phi} \times [\pi_1^s + \nu_0^m \times \pi_1^m + \nu_1^m \times \pi_0^m]$$

Do you see any problems with this procedure? Why may this not give us the true effects of switching to the alternative rule $i_t = \tilde{\phi} \times \pi_t$? What's the obvious Lucas critique concern?

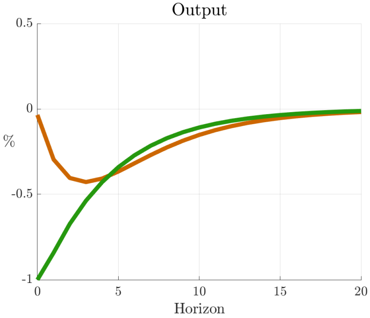
A numerical illustration of Sims-Zha



Experiment: cost-push shock under **base rule** $i_t = \phi_\pi \pi_t$ & **cfnctl rule** $i_t = \tilde{\phi}_\pi \pi_t + \tilde{\phi}_y y_t$

A numerical illustration of Sims-Zha

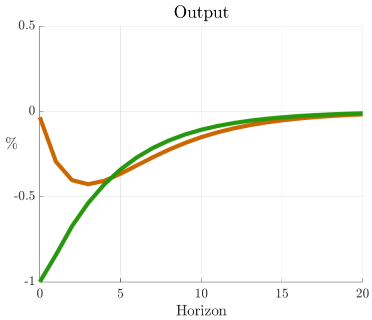
Q: What information would an econometrician need to correctly predict the **counterfactual**?



Experiment: cost-push shock under **base rule** $i_t = \phi_\pi \pi_t$ & **cfnctl rule** $i_t = \tilde{\phi}_\pi \pi_t + \tilde{\phi}_y y_t$

A numerical illustration of Sims-Zha

Q: What information would an econometrician need to correctly predict the **counterfactual**?



- Info: get IRFs to MP shock

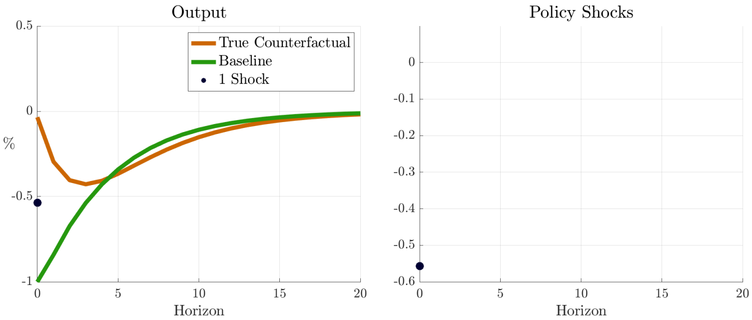
$$i_t = \phi_\pi \pi_t + \nu_t^m$$

Strategy: enforce **counterfactual rule** using sequence of one-time **monetary shocks**

This is Sims-Zha (1995).

A numerical illustration of Sims-Zha

Q: What information would an econometrician need to correctly predict the **counterfactual**?



- Info: get IRFs to MP shock

$$i_t = \phi_\pi \pi_t + \nu_t^m$$

- Set MP shocks to values $\{\nu_0^m, \nu_1^m, \nu_2^m, \ldots\}$ so that $\{i_t, \pi_t, y_t\}$ are related as

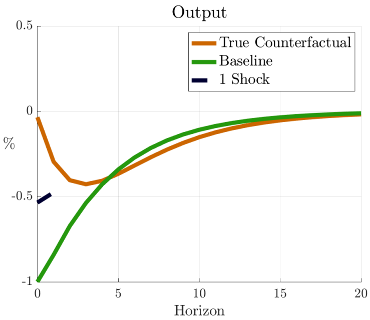
$$i_t = \tilde{\phi}_\pi \pi_t + \tilde{\phi}_y y_t$$

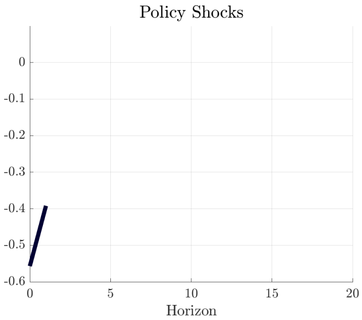
Strategy: enforce **counterfactual rule** using sequence of one-time **monetary shocks**

This is Sims-Zha (1995).

A numerical illustration of Sims-Zha

Q: What information would an econometrician need to correctly predict the **counterfactual**?





- Info: get IRFs to MP shock

$$i_t = \phi_\pi \pi_t + \nu_t^m$$

- Set MP shocks to values $\{\nu_0^m, \nu_1^m, \nu_2^m, \dots\}$ so that $\{i_t, \pi_t, y_t\}$ are related as

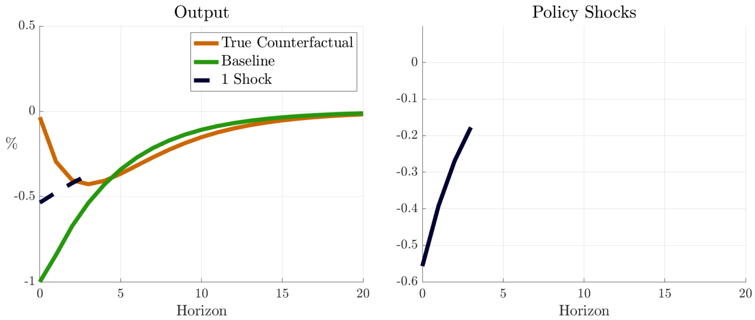
$$i_t = \tilde{\phi}_\pi \pi_t + \tilde{\phi}_y y_t$$

Strategy: enforce **counterfactual rule** using sequence of one-time **monetary shocks**

This is Sims-Zha (1995).

A numerical illustration of Sims-Zha

Q: What information would an econometrician need to correctly predict the **counterfactual**?



- Info: get IRFs to MP shock

$$i_t = \phi_\pi \pi_t + \nu_t^m$$

- Set MP shocks to values $\{\nu_0^m, \nu_1^m, \nu_2^m, \ldots\}$ so that $\{i_t, \pi_t, y_t\}$ are related as

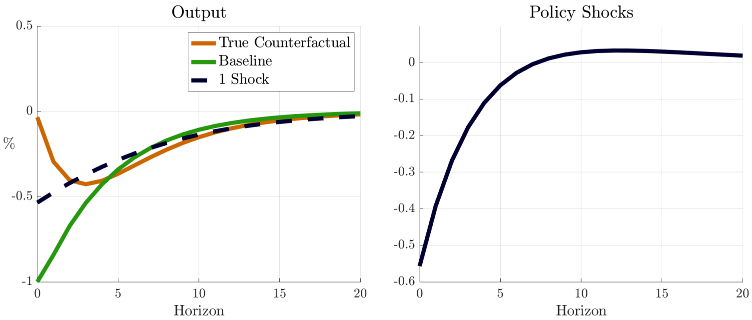
$$i_t = \tilde{\phi}_\pi \pi_t + \tilde{\phi}_y y_t$$

Strategy: enforce **counterfactual rule** using sequence of one-time **monetary shocks**

This is Sims-Zha (1995).

A numerical illustration of Sims-Zha

Q: What information would an econometrician need to correctly predict the **counterfactual**?



- Info: get IRFs to MP shock

$$i_t = \phi_\pi \pi_t + \nu_t^m$$

- Set MP shocks to values $\{\nu_0^m, \nu_1^m, \nu_2^m, \dots\}$ so that $\{i_t, \pi_t, y_t\}$ are related as

$$i_t = \tilde{\phi}_\pi \pi_t + \tilde{\phi}_y y_t$$

Strategy: enforce **counterfactual rule** using sequence of one-time **monetary shocks**

This is Sims-Zha (1995). Problem: at each t, private sector expects return to old rule from $t+1$ onwards.

Outline

Some preliminary intuition

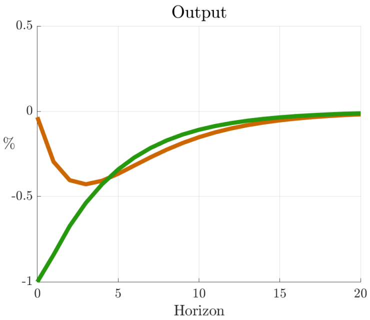
- Obvious **concern** so far: we are missing expectational effects
- But: having access to **multiple distinct** policy shocks may help ...
  - Concrete example: consider the rule + shocks

$$i_t = \phi_\pi \pi_t + \nu_{0,t}^m + \sum_{\ell=1}^{\infty} \nu_{\ell,t-\ell}^m$$

    and suppose we can estimate the effects of the first n policy shocks
  - Now we have more degrees of freedom: we could enforce the counterfactual rule ex post in eq'm, but also in ex ante expectation for $n-1$ periods
- **Q**: does that get us closer to the truth? what happens as n gets large?

Graphical explorations

Q: What information would an econometrician need to correctly predict the **counterfactual**?



- Info: get IRFs to MP shocks

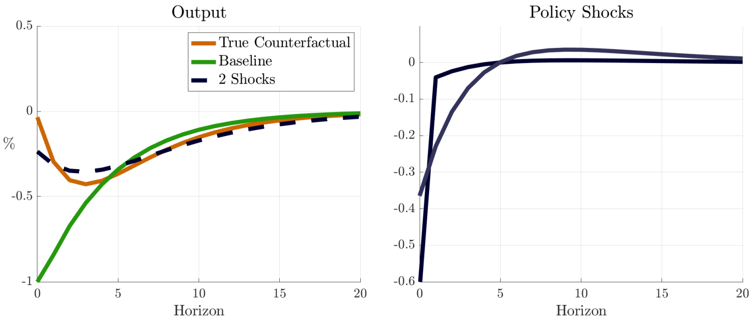
$$\boldsymbol{i_t = \phi_\pi \pi_t} + \nu^m_{0,t} + \nu^m_{1,t-1} + \dots$$

Alternative: with **multiple monetary shocks** we can start to also match *expectations*

With n shocks we can enforce the rule today and in expectation for the next $n-1$ periods.

Graphical explorations

Q: What information would an econometrician need to correctly predict the **counterfactual**?



- Info: get IRFs to MP shocks

  $$i_t = \phi_\pi \pi_t + \nu_{0,t}^m + \nu_{1,t-1}^m + \dots$$

- Set MP shocks so that $\{i_t, \pi_t, y_t\}$ are related as

  $$i_t = \tilde{\phi}_\pi \pi_t + \tilde{\phi}_y y_t$$

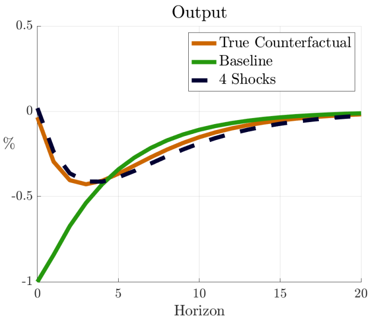
  not just ex post on eq'm path but also *in expectation*

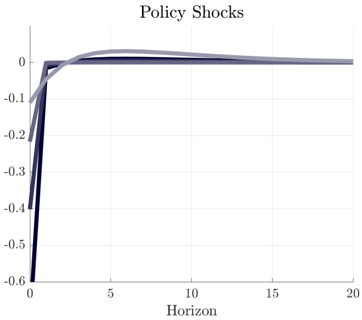
Alternative: with **multiple monetary shocks** we can start to also match *expectations*

With n shocks we can enforce the rule today and in expectation for the next $n-1$ periods.

Graphical explorations

Q: What information would an econometrician need to correctly predict the **counterfactual**?





- Info: get IRFs to MP shocks

  $$i_t = \phi_\pi \pi_t + \nu^m_{0,t} + \nu^m_{1,t-1} + \dots$$

- Set MP shocks so that $\{i_t, \pi_t, y_t\}$ are related as

  $$i_t = \tilde{\phi}_\pi \pi_t + \tilde{\phi}_y y_t$$

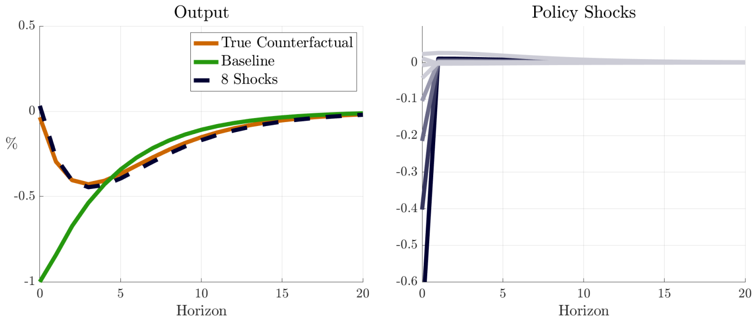
  not just ex post on eq'm path but also *in expectation*

Alternative: with **multiple monetary shocks** we can start to also match *expectations*

With n shocks we can enforce the rule today and in expectation for the next $n-1$ periods.

Graphical explorations

Q: What information would an econometrician need to correctly predict the **counterfactual**?



- Info: get IRFs to MP shocks

$$i_t = \phi_\pi \pi_t + \nu^m_{0,t} + \nu^m_{1,t-1} + \ldots$$

- Set MP shocks so that $\{i_t, \pi_t, y_t\}$ are related as

$$i_t = \tilde{\phi}_\pi \pi_t + \tilde{\phi}_y y_t$$

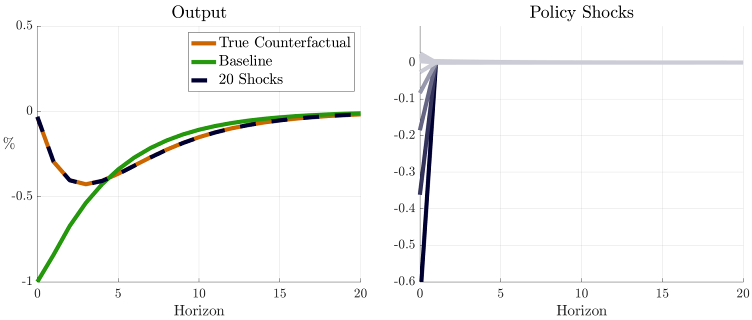
not just ex post on eq'm path but also *in expectation*

Alternative: with **multiple monetary shocks** we can start to also match *expectations*

With n shocks we can enforce the rule today and in expectation for the next $n-1$ periods.

Graphical explorations

Q: What information would an econometrician need to correctly predict the **counterfactual**?



- Info: get IRFs to MP shocks

  $$i_t = \phi_\pi \pi_t + \nu^m_{0,t} + \nu^m_{1,t-1} + \dots$$

- Set MP shocks so that $\{i_t, \pi_t, y_t\}$ are related as

  $$i_t = \tilde{\phi}_\pi \pi_t + \tilde{\phi}_y y_t$$

  not just ex post on eq'm path but also *in expectation*

Limit: with **many monetary shocks** we seem to recover the **true counterfactual**

Note: counterfactual rule is enforced *ex-post* and *ex-ante* using only date-0 shocks. No ex-post surprises.

Towards a general identification result

Q: Across what space of structural models does this logic work out?

- **Model** [perfect foresight = 1st-order perturbation w/ aggregate risk]

$$\mathcal{H}_x \boldsymbol{x} + \mathcal{H}_z \boldsymbol{z} + \mathcal{H}_\varepsilon \boldsymbol{\varepsilon} = \mathbf{0} \qquad n_x \times T \text{ eqn's} \tag{1}$$

$$\mathcal{A}_x \boldsymbol{x} + \mathcal{A}_z \boldsymbol{z} + \boldsymbol{\nu} = \mathbf{0} \qquad n_z \times T \text{ eqn's} \tag{2}$$

Boldface denotes time paths, e.g. $\boldsymbol{x} = (x_0, x_1, x_2, \dots)'$. Solution of (1) - (2) = impulse responses.

Towards a general identification result

Q: Across what space of structural models does this logic work out?

- **Model** [perfect foresight = 1st-order perturbation w/ aggregate risk]

$$
\begin{aligned}
\mathcal{H}_x \boldsymbol{x} + \mathcal{H}_z \boldsymbol{z} + \mathcal{H}_\varepsilon \boldsymbol{\varepsilon} &= \boldsymbol{0} \qquad n_x \times T \text{ eqn's} \qquad (1) \\
\boldsymbol{\mathcal{A}}_{\boldsymbol{x}} \boldsymbol{x} + \boldsymbol{\mathcal{A}}_{\boldsymbol{z}} \boldsymbol{z} + \boldsymbol{\nu} &= \boldsymbol{0} \qquad n_z \times T \text{ eqn's} \qquad (2)
\end{aligned}
$$

Boldface denotes time paths, e.g. $\boldsymbol{x} = (x_0, x_1, x_2, \ldots)'$. Solution of (1) - (2) = impulse responses.

  - Illustrative example: three-equation NK model

Towards a general identification result

Q: Across what space of structural models does this logic work out?

- **Model** [perfect foresight = 1st-order perturbation w/ aggregate risk]

$$\mathcal{H}_x \boldsymbol{x} + \mathcal{H}_z \boldsymbol{z} + \mathcal{H}_\varepsilon \boldsymbol{\varepsilon} = \boldsymbol{0} \qquad n_x \times T \text{ eqn's} \tag{1}$$

$$\mathcal{A}_x \boldsymbol{x} + \mathcal{A}_z \boldsymbol{z} + \boldsymbol{\nu} = \boldsymbol{0} \qquad n_z \times T \text{ eqn's} \tag{2}$$

Boldface denotes time paths, e.g. $\boldsymbol{x} = (x_0, x_1, x_2, \dots)'$. Solution of (1) - (2) = impulse responses.

  - Illustrative example: three-equation NK model

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

Towards a general identification result

Q: Across what space of structural models does this logic work out?

- **Model** [perfect foresight = 1st-order perturbation w/ aggregate risk]

$$\mathcal{H}_x \boldsymbol{x} + \mathcal{H}_z \boldsymbol{z} + \mathcal{H}_\varepsilon \boldsymbol{\varepsilon} = \boldsymbol{0} \qquad n_x \times T \text{ eqn's} \tag{1}$$

$$\mathcal{A}_x \boldsymbol{x} + \mathcal{A}_z \boldsymbol{z} + \boldsymbol{\nu} = \boldsymbol{0} \qquad n_z \times T \text{ eqn's} \tag{2}$$

Boldface denotes time paths, e.g. $\boldsymbol{x} = (x_0, x_1, x_2, \dots)'$. Solution of (1) - (2) = impulse responses.

  - Illustrative example: three-equation NK model

$$\begin{pmatrix} 1 & -\beta & 0 & \dots \\ 0 & 1 & -\beta & \dots \\ 0 & 0 & 1 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \boldsymbol{\pi} - \kappa \boldsymbol{y} - \begin{pmatrix} 1 \\ \rho_\varepsilon \\ \rho_\varepsilon^2 \\ \vdots \end{pmatrix} \varepsilon_0 = \boldsymbol{0}$$

Towards a general identification result

Q: Across what space of structural models does this logic work out?

- **Model** [perfect foresight = 1st-order perturbation w/ aggregate risk]

$$\mathcal{H}_x \boldsymbol{x} + \mathcal{H}_z \boldsymbol{z} + \mathcal{H}_\varepsilon \boldsymbol{\varepsilon} = \mathbf{0} \qquad n_x \times T \text{ eqn's} \tag{1}$$

$$\boldsymbol{\mathcal{A}}_{\boldsymbol{x}} \boldsymbol{x} + \boldsymbol{\mathcal{A}}_{\boldsymbol{z}} \boldsymbol{z} + \boldsymbol{\nu} = \mathbf{0} \qquad n_z \times T \text{ eqn's} \tag{2}$$

Boldface denotes time paths, e.g. $\boldsymbol{x} = (x_0, x_1, x_2, \dots)'$. Solution of (1) - (2) = impulse responses.

  - Illustrative example: three-equation NK model

$$\boldsymbol{i} - \boldsymbol{\phi} \boldsymbol{\pi} - \begin{pmatrix} \nu_0 & \nu_1 & \nu_2 & \dots \end{pmatrix}' = \mathbf{0}$$

Towards a general identification result

Q: Across what space of structural models does this logic work out?

- **Model** [perfect foresight = 1st-order perturbation w/ aggregate risk]

$$\begin{aligned} \mathcal{H}_x \boldsymbol{x} + \mathcal{H}_z \boldsymbol{z} + \mathcal{H}_\varepsilon \boldsymbol{\varepsilon} &= \boldsymbol{0} \qquad n_x \times T \text{ eqn's} \qquad (1) \\ \boldsymbol{\mathcal{A}}_{\boldsymbol{x}} \boldsymbol{x} + \boldsymbol{\mathcal{A}}_{\boldsymbol{z}} \boldsymbol{z} + \boldsymbol{\nu} &= \boldsymbol{0} \qquad n_z \times T \text{ eqn's} \qquad (2) \end{aligned}$$

Boldface denotes time paths, e.g. $\boldsymbol{x} = (x_0, x_1, x_2, \dots)'$. Solution of (1) - (2) = impulse responses.

  - Illustrative example: three-equation NK model
  - More generally: many linearized models fit (1) - (2) [RBC, NK-DSGE, HANK, ...]

Towards a general identification result

Q: Across what space of structural models does this logic work out?

- **Model** [perfect foresight = 1st-order perturbation w/ aggregate risk]

$$\mathcal{H}_x \boldsymbol{x} + \mathcal{H}_z \boldsymbol{z} + \mathcal{H}_\varepsilon \boldsymbol{\varepsilon} = \mathbf{0} \qquad n_x \times T \text{ eqn's} \tag{1}$$

$$\boldsymbol{\mathcal{A}_x x} + \boldsymbol{\mathcal{A}_z z} + \boldsymbol{\nu} = \mathbf{0} \qquad n_z \times T \text{ eqn's} \tag{2}$$

Boldface denotes time paths, e.g. $\boldsymbol{x} = (x_0, x_1, x_2, \dots)'$. Solution of (1) - (2) = impulse responses.

  - Illustrative example: three-equation NK model
  - More generally: many linearized models fit (1) - (2) [RBC, NK-DSGE, HANK, ...]
  - This environment has **two key features**:

    (i) Private sector responds only to current & future values of policy instrument, not rule *per se*

    (ii) Linearity in aggregates: can restrict attention to expected values [not sign, size, state, ...]

Towards a general identification result

Q: Across what space of structural models does this logic work out?

- **Model** [perfect foresight = 1st-order perturbation w/ aggregate risk]

$$\mathcal{H}_x \boldsymbol{x} + \mathcal{H}_z \boldsymbol{z} + \mathcal{H}_\varepsilon \boldsymbol{\varepsilon} = \mathbf{0} \qquad n_x \times T \text{ eqn's} \tag{1}$$

$$\boldsymbol{\mathcal{A}}_x \boldsymbol{x} + \boldsymbol{\mathcal{A}}_z \boldsymbol{z} + \boldsymbol{\nu} = \mathbf{0} \qquad n_z \times T \text{ eqn's} \tag{2}$$

Boldface denotes time paths, e.g. $\boldsymbol{x} = (x_0, x_1, x_2, \dots)'$. Solution of (1) - (2) = impulse responses.

- **Equilibrium**: bounded $\{\boldsymbol{x}, \boldsymbol{z}\}$ that solves (1) - (2) given bounded $\{\boldsymbol{\varepsilon}, \boldsymbol{\nu}\}$
  - Assume that, under baseline rule, eq'm exists & is unique
  - Write IRFs to path $\boldsymbol{\varepsilon}$ under baseline rule as $\{\boldsymbol{x}(\boldsymbol{\varepsilon}), \boldsymbol{z}(\boldsymbol{\varepsilon})\}$

Towards a general identification result

Q: Across what space of structural models does this logic work out?

- **Model** [perfect foresight = 1st-order perturbation w/ aggregate risk]

$$\begin{aligned} \mathcal{H}_x \boldsymbol{x} + \mathcal{H}_z \boldsymbol{z} + \mathcal{H}_\varepsilon \boldsymbol{\varepsilon} &= \mathbf{0} \qquad n_x \times T \text{ eqn's} && (1) \\ \boldsymbol{\mathcal{A}_x} \boldsymbol{x} + \boldsymbol{\mathcal{A}_z} \boldsymbol{z} + \boldsymbol{\nu} &= \mathbf{0} \qquad n_z \times T \text{ eqn's} && (2) \end{aligned}$$

Boldface denotes time paths, e.g. $\boldsymbol{x} = (x_0, x_1, x_2, \dots)'$. Solution of (1) - (2) = impulse responses.

- **Equilibrium**: bounded $\{\boldsymbol{x}, \boldsymbol{z}\}$ that solves (1) - (2) given bounded $\{\boldsymbol{\varepsilon}, \boldsymbol{\nu}\}$
  - Assume that, under baseline rule, eq'm exists & is unique
  - Write IRFs to path $\boldsymbol{\varepsilon}$ under baseline rule as $\{\boldsymbol{x}(\boldsymbol{\varepsilon}), \boldsymbol{z}(\boldsymbol{\varepsilon})\}$
- **Object of interest**: IRFs $\{\tilde{\boldsymbol{x}}(\boldsymbol{\varepsilon}), \tilde{\boldsymbol{z}}(\boldsymbol{\varepsilon})\}$ under alternative rule $\{\tilde{\boldsymbol{\mathcal{A}}}_x, \tilde{\boldsymbol{\mathcal{A}}}_z\}$

Dynamic causal effects – the VAR/LP estimands

- Solving the system under the baseline rule gives

$$\begin{pmatrix} x \\ z \end{pmatrix} = \underbrace{-\begin{pmatrix} \mathcal{H}_x & \mathcal{H}_z \\ \mathcal{A}_x & \mathcal{A}_z \end{pmatrix}^{-1} \times \begin{pmatrix} \mathcal{H}_\varepsilon & 0 \\ 0 & I \end{pmatrix}}_{\equiv \Theta} \times \begin{pmatrix} \boldsymbol{\varepsilon} \\ \boldsymbol{\nu} \end{pmatrix}, \quad \Theta \equiv \begin{pmatrix} \Theta_{x,\varepsilon} & \Theta_{x,\nu} \\ \Theta_{z,\varepsilon} & \Theta_{z,\nu} \end{pmatrix}$$

- **Informational requirements** to construct $\{\tilde{x}(\boldsymbol{\varepsilon}), \tilde{z}(\boldsymbol{\varepsilon})\}$

  1. **Non-policy**: causal effects of particular non-policy shock $\boldsymbol{\varepsilon}$ under base rule $\{x(\boldsymbol{\varepsilon}), z(\boldsymbol{\varepsilon})\}$
     Note: this is one-dimensional. One particular shock path.
  2. **Policy**: causal effects of all current *and news* shocks $\boldsymbol{\nu}$ to the base rule, $\{\Theta_{x,\nu}, \Theta_{z,\nu}\}$
     - This is multi-dimensional—each column gives the IRF to a particular policy shock
     - First column = contemp. shock, later columns = news shocks

$$\underbrace{\hphantom{xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx}}_{\text{individual } \textbf{VARs/LPs} \text{ give } \{x(\boldsymbol{\varepsilon}), z(\boldsymbol{\varepsilon})\} \text{ \& } \textit{(avg's of) columns} \text{ of } \{\Theta_{x,\nu}, \Theta_{z,\nu}\}}$$

Counterfactual policy rules

Proposition

For any $\{\tilde{\mathcal{A}}_x, \tilde{\mathcal{A}}_z\}$ that induces a unique eq'm, we can recover the counterfactuals $\tilde{x}(\varepsilon)$ and $\tilde{z}(\varepsilon)$ as the impulse responses under the baseline rule to $\{\varepsilon, \tilde{\nu}\}$, where $\tilde{\nu}$ solves

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

In words: select date-0 policy shocks $\tilde{\nu}$ so that cnfct'l rule holds following $\{\varepsilon, \tilde{\nu}\}$.

- **Key intuition**: private sector only cares about (expected) instrument path
  $\Rightarrow$ use many date-0 shocks to enforce new instrument path in date-0 expectation, not just ex-post
- Let's provide a sketch of the proof on the board ...

Optimal policy rules

- Same argument applies for **optimal policy**. Consider a policymaker with loss function

$$\mathcal{L} = \frac{1}{2}\sum_{i=1}^{n_x} \lambda_i \boldsymbol{x}_i' W \boldsymbol{x}_i, \quad W = \text{diag}(1, \beta, \beta^2, \dots) \tag{3}$$

- Let's begin by computing FOCs in the usual way:
  - Policy problem is to minimize loss subject to private-sector block. This gives

$$\begin{aligned} (\Lambda \otimes W)\boldsymbol{x} + \mathcal{H}_x'(I \otimes W)\boldsymbol{\varphi} &= \mathbf{0} \quad (4) \\ \mathcal{H}_z' W \boldsymbol{\varphi} &= \mathbf{0} \quad (5) \end{aligned}$$

    where $\Lambda = \text{diag}(\lambda_1, \lambda_2, \dots)$ and φ is the multiplier on (1)

  - Denote solutions of FOCs + (1) by $\{\boldsymbol{x}^*(\boldsymbol{\varepsilon}), \boldsymbol{z}^*(\boldsymbol{\varepsilon}), \boldsymbol{\varphi}^*(\boldsymbol{\varepsilon})\}$.

Optimal policy rules

- Now let's consider the artificial problem of picking the best shocks $\boldsymbol{\nu}^*$ to the rule (2)
  - This gives the FOCs

$$\begin{aligned} (\Lambda \otimes W)\boldsymbol{x} + \mathcal{H}_x'(I \otimes W)\boldsymbol{\varphi} + \mathcal{A}_x' W \boldsymbol{\varphi}_z &= \mathbf{0} & (6) \\ \mathcal{H}_z'(I \otimes W)\boldsymbol{\varphi} + \mathcal{A}_z' W \boldsymbol{\varphi}_z &= \mathbf{0} & (7) \\ W \boldsymbol{\varphi}_z &= \mathbf{0} & (8) \end{aligned}$$

    The policy rule multiplier $\boldsymbol{\varphi}_z$ is equal to $\mathbf{0}$, so they are the same problem. Interpretation?
  - Let's re-write the constraint set of this alternative artificial problem as

$$\begin{pmatrix} \boldsymbol{x} \\ \boldsymbol{z} \end{pmatrix} = \Theta \times \begin{pmatrix} \boldsymbol{\varepsilon} \\ \boldsymbol{\nu}^* \end{pmatrix}$$

Optimal policy rules

- Now let's consider the artificial problem of picking the best shocks $\boldsymbol{\nu}^*$ to the rule (2)
  - Solving the problem with this re-written constraint set, we then get an FOC in the form of an optimal policy rule
$$\sum_{i=1}^{n_x} \lambda_i \Theta'_{x_i,\nu} W \boldsymbol{x}_i = 0 \tag{9}$$
  - In words: pick the best combination of your targets $\boldsymbol{x}$ that's attainable via policy shocks
  - vs. policy practice: (9) is in the form of a so-called "forecast target criterion"—a restriction on current and future values of the policymaker targets
    You can read up on the classical optimal policy literature on such rules. Standard references by Svensson and Woodford. Will return to this in our later applications.
- Key: private sector doesn't care whether you achieve this best combination through some policy rule or through shocks to a different policy rule

Discussion

How general is the identification result? When does it break down?

(i) **Sufficiency of policy instrument**

- Model restrictions: most notably fails in signal extraction problems
  E.g.: in Lucas island economy rule matters above and beyond nominal demand growth

(ii) **Linearity**

- Convenient, but not essential: allows to reduce measurement problem to T dimensions
  See McKay-Wolf (2022) for a general global identification result. Simple idea: before just needed to match means (= expectations), now need to match entire distribution of z through policy shocks.
- Rule restrictions: cnfctl rule should not affect the system's steady state
  E.g.: response to a given rate path may be different in high- vs. low-inflation economies

general **"sufficient statistics"** result for business-cycle models

Outline

Practical implications

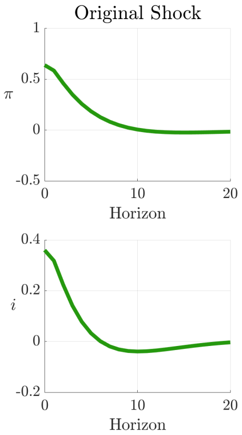
How do these identification results relate to the two dominant **methodological approaches** that we reviewed at the beginning?

1. **Lucas program**
    - Provides a justification of IRF matching for model estimation. The full set of policy shock IRFs Θ are "sufficient statistics", so targeting a single one seems like a good idea.
    - Later: application to optimal monetary policy in HANK
    - Important but open **Q**: can we give a tighter robustness interpretation to this?
2. **Sims-Zha**
    - Violates the Lucas critique only because of expectational effects. The natural solution is simply to get evidence on more distinct policy shocks.
    - Later: application to counterfactual federal funds rate paths

Outline

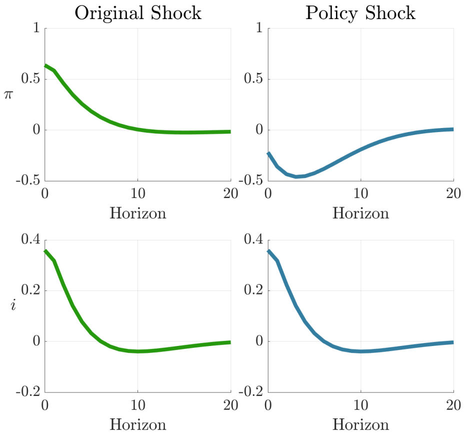
Connecting with empirical evidence: example

Q: How would this **cost-push shock** have propagated in the absence of a monetary reaction?



Connecting with empirical evidence: example

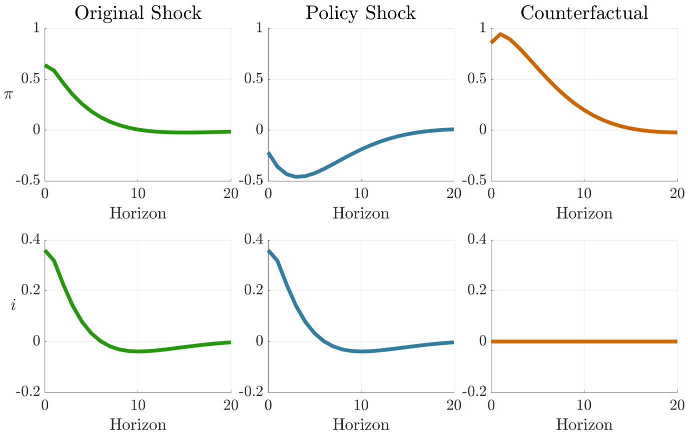
Q: How would this **cost-push shock** have propagated in the absence of a monetary reaction?



- ID result: find a **monetary shock** inducing the same rate response
  $\Rightarrow$ move $\mathbb{E}_0(i_t)$ just like **cost-push shock**

Connecting with empirical evidence: example

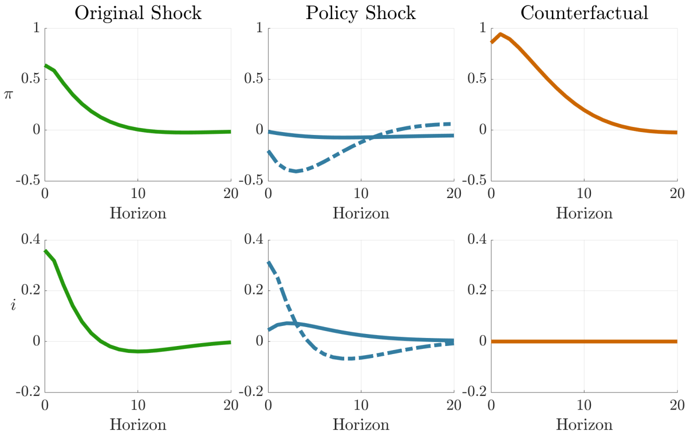
Q: How would this **cost-push shock** have propagated in the absence of a monetary reaction?



- ID result: find a **monetary shock** inducing the same rate response
  $\Rightarrow$ move $\mathbb{E}_0(i_t)$ just like **cost-push shock**
- $\Longrightarrow$ if a model matches **(1)** & **(2)**, it will agree with **cnfct'l (3)**

Connecting with empirical evidence: example

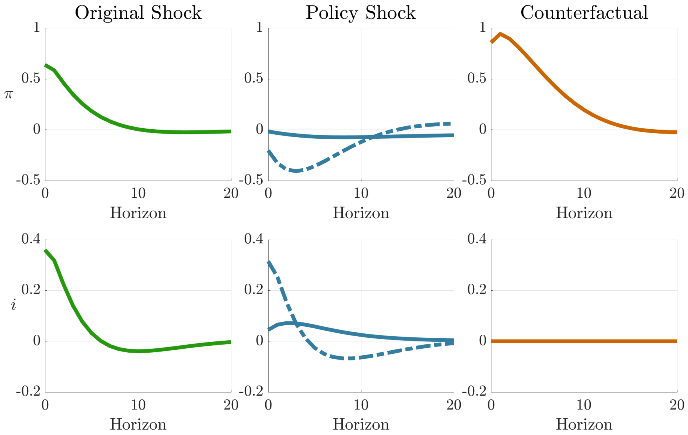
Q: How would this **cost-push shock** have propagated in the absence of a monetary reaction?



- ID result: find a **monetary shock** inducing the same rate response
  $\Rightarrow$ move $\mathbb{E}_0(i_t)$ just like **cost-push shock**
- $\Longrightarrow$ if a model matches **(1)** & **(2)**, it will agree with **cnfct'l (3)**
  [Same result for combo of MP shocks.]

Connecting with empirical evidence: example

Q: How would this **cost-push shock** have propagated in the absence of a monetary reaction?



- ID result: find a **monetary shock** inducing the same rate response
  $\Rightarrow$ move $\mathbb{E}_0(i_t)$ just like **cost-push shock**
- $\Longrightarrow$ if a model matches **(1)** & **(2)**, it will agree with **cnfct'l (3)**
  [Same result for combo of MP shocks.]
- Emp. method: enforce **cnfct'l rule** *as well as possible* using linear combo of ***date-0* policy shocks**

Counterfactuals with "a few" shocks

The practically relevant case is where you observe $1 < n \ll \infty$ **policy shocks**, giving the columns of the (small) IRF matrices $\{\Omega_{x,\nu}, \Omega_{z,\nu}\}$. What can you do with those?

- **Method**: enforce cnfctl rule *as well as possible* **using only a few $t = 0$ shocks**
  Note: no ex-post shocks, so fully Lucas critique robust, but imperfect rule match.
  - Solve problem:
    $$\min_{\boldsymbol{v}} \underbrace{\left|\left|\tilde{A}_x \left(x(\varepsilon) + \Omega_{x,\nu} \times v\right) + \tilde{A}_z \left(z(\varepsilon) + \Omega_{z,\nu} \times v\right)\right|\right|}_{\text{rule inaccuracy}}$$
    This selects linear combo of time-0 shocks to implement rule *as well as possible*
- **Discussion**
  - Clearly not enough to estimate all possible counterfactuals
  - But perhaps we can recover *some* interesting counterfactuals?

A review of monetary policy shocks

- What kind of shocks can we get from the **empirical monetary policy literature**?
- Key: monetary policy is **multi-dimensional**, and thus so are IVs for policy shocks

$$i_t = f(\Omega_t) + \underbrace{\nu^m_{0,t} + \nu^m_{1,t-1} + \nu^m_{2,t-2} + \dots}_{\text{a policy shock IV correlates with those } \nu^m\text{'s}}$$

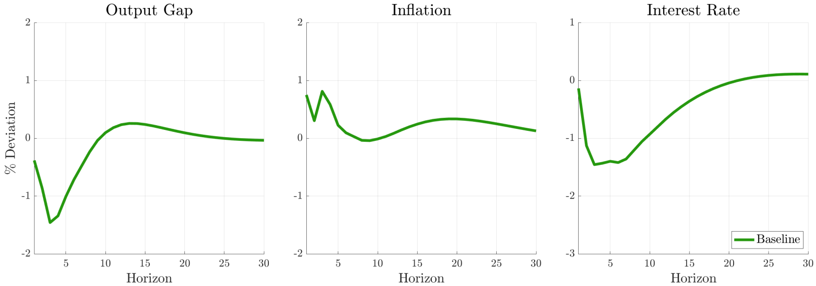
[Notation: $f(\bullet)$ is the systematic policy rule and Ω_t is the date-t information set.]

  - In application on next slide will use two canonical monetary shock series:
    1. Romer-Romer: transitory innovation to short-term rates
    2. Gertler-Karadi: persistent innovation/greater forward guidance component
  - Some work actually explicitly splits monetary innovations by their effects on the yield curve
    Gurkaynak-Sack-Swanson, Antolin-Diaz-Petrella-Rubio-Ramirez, Inoue-Rossi

Application: investment technology shocks

Q: How would **investment demand shocks** propagate under different **monetary rules**?

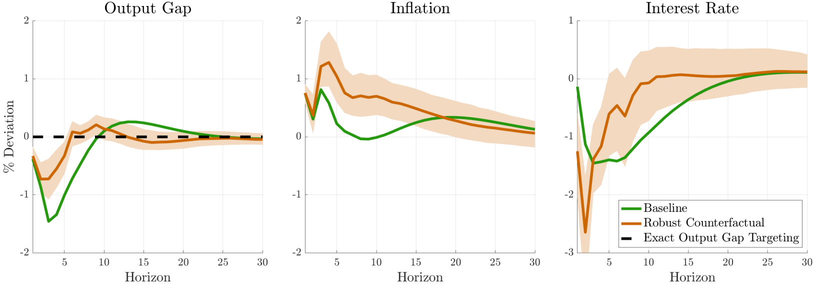
- **Inputs**
  - **Original shock**: contractionary innovation to inv. technology [Ben Zeev-Khan]
  - **Policy shocks**: two different interest rate response paths [Romer-Romer & Gertler-Karadi]
- **Results**:



Application: investment technology shocks

Q: How would **investment demand shocks** propagate under different **monetary rules**?

- **Inputs**
  - **Original shock**: contractionary innovation to inv. technology [Ben Zeev-Khan]
  - **Policy shocks**: two different interest rate response paths [Romer-Romer & Gertler-Karadi]
- **Results**: **strict output gap targeting**

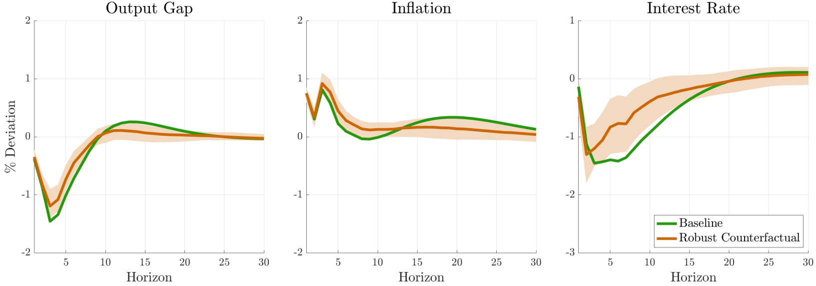


Wolf

Application: investment technology shocks

Q: How would **investment demand shocks** propagate under different **monetary rules**?

- **Inputs**
  - **Original shock**: contractionary innovation to inv. technology [Ben Zeev-Khan]
  - **Policy shocks**: two different interest rate response paths [Romer-Romer & Gertler-Karadi]
- **Results**: **optimal dual mandate policy**



Other applications

- Lots of other interesting questions one could tackle with this approach ...
- **Some ideas**:
  1. **Monetary policy**
     - Consider the current inflationary episode. Should the Fed have tightened earlier? And what effects would such earlier tightening have had on the rest of the world?
     - How far would the Fed have needed to cut rates in 2008/2009 to stabilize the economy? In other words, how much of a constraint was the ZLB?
     - Are there any past historical episodes in which, with the benefit of hindsight, the Fed tightened too much/too little?
  2. **Fiscal policy**
     - How much did the Biden stimulus contribute to the recent inflation?
     - How big of a fiscal expansion would have been needed in 2008/2009 to stabilize the economy at the ZLB?

Outline

Optimal monetary policy in NK models

- Now let's return to the **Lucas program**—sometimes empirical evidence is not enough, so we'll need to rely on structural models
- Natural strategy: model estimation via **impulse response matching**
  - Basic idea: can learn about parts of the shock causal effects Θ from the data, then extrapolate to get all of Θ using a structural model
  - Implementation details ▸ Details

  As said before: a great open question is to more formally justify the appeals of this—intuitively, there should be some kind of robustness argument.
- Our **application**: optimal monetary policy rules. Will consider two loss functions:
  1. A conventional dual mandate policymaker
     Arguably practically relevant. This is the mandate of the Fed.
  2. A policymaker that also cares about inequality
     Note: can be microfounded as a Ramsey problem in a structural HANK model.

Outline

The dual mandate policy problem

- A dual-mandate policymaker has loss function

$$\mathcal{L}^{DM} \equiv \frac{1}{2}\sum_{t=0}^{\infty}\beta^t\left[\lambda_\pi \widehat{\pi}_t^2 + \lambda_y \widehat{y}_t^2\right]$$

- We know from our general derivations above that the optimal rule takes the form

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

- Now suppose that $\Theta_{\pi,\nu}$ is invertible. Then we can write this as

  In words: the policymaker can implement any possible sequence of inflation.

$$\lambda_\pi \boldsymbol{\pi} + \lambda_y \underbrace{\left(\Theta'_{\pi,\nu} W\right)^{-1}\left(\Theta'_{y,\nu} W\right)}_{\text{what can we say about this?}} \boldsymbol{y} = \mathbf{0}$$

Leveraging Phillips curve structure

- How much **model structure** do we need to say something about $(\Theta'_{\pi,\nu})^{-1} \times (\Theta'_{y,\nu})$?
  - Note that, if the monetary authority moves interest rates to move inflation by $d\boldsymbol{\pi}$, then the effect on output is
  $$d\boldsymbol{y} = \Theta_{y,\nu}\Theta_{\pi,\nu}^{-1}d\boldsymbol{\pi}$$
  - Key insight: in models with an NKPC, this mapping is purely governed by the NKPC—all other parts of the model are irrelevant Simple logic: MP moves us along the NKPC ...
  - We can use thus our insights from Lecture Note 6 to estimate optimal policy rules ...
- Simple but instructive example: canonical NKPC $[\pi_t = \kappa y_t + \beta\pi_{t+1}]$
  - Can show that with this particular NKPC we have [exercise!]
  $$(\Theta'_{\pi,\nu}W)^{-1} \times (\Theta'_{y,\nu}W) = \frac{1}{\kappa} \times \begin{pmatrix} 1 & 0 & 0 & \dots \\ -1 & 1 & 0 & \dots \\ 0 & -1 & 1 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$
  - But that's the optimal monetary policy rule in Woodford/Gali! All roads lead to Rome ...

Outline

The objective function

- What happens if the policymaker also cares about inequality?
    - Nothing happens with complete markets. Under weak assumptions, everyone just scales up and down with the aggregate economy
    - Recent HANK literature: incomplete markets. Policymakers use their instruments to "endogenously" try to complete markets
      Bhandari et al., Acharya et al., McKay-Wolf, ... But this theory is not our focus.

- Takeaway for our purposes: under some conditions can write objective as

$$\mathcal{L}^{HA} \approx \frac{1}{2}\sum_{t=0}^{\infty}\beta^t\left[\lambda_\pi\widehat{\pi}_t^2 + \lambda_y\widehat{y}_t^2 + \int \lambda_i\widehat{\omega}_{it}^2 di\right],$$

    - Notation: the λ's depend on primitives & the ω_i's are consumption *shares* of individual i
    - In words: the planner seeks to stabilize y, π, and the consumption shares of all i's

The optimal policy rule

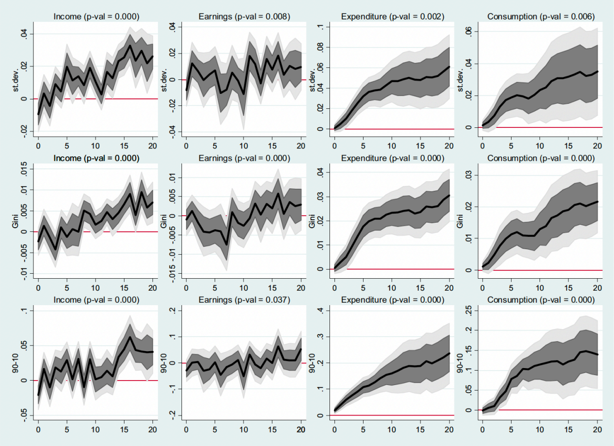
- Applying our logic from before:
  - Let $\{\Theta_{\pi,i}, \Theta_{y,i}, \Theta_{\omega_i,i}\}$ denote the effects of interest rate policy i on the three policy targets in the Ramsey loss function. Then the optimal rule is

$$\lambda_\pi \Theta'_{\pi,i}\widehat{\boldsymbol{\pi}} + \lambda_y \Theta'_{y,i}\widehat{\boldsymbol{y}} + \int \lambda_i \Theta'_{\omega_i,i}\widehat{\boldsymbol{\omega}}_i = \mathbf{0}$$

  - Same intuition: set instruments to align all of the policy targets as well as possible
- **Discussion**
  - Suppose monetary policy doesn't affect inequality, i.e. $\Theta_{\omega_i,i} = \mathbf{0}$. Then same rule as in for dual mandate! Intuition: MP is useless to complete markets. Special case: Werning (2015)
  - Thus: deviate from dual mandate if and only if MP has meaningful distributional effects
- This is ultimately an empirical question. So what does the evidence say?

Monetary policy & inequality

Results from Coibion et al. (2017)



Results from Cloyne et al. (2020)

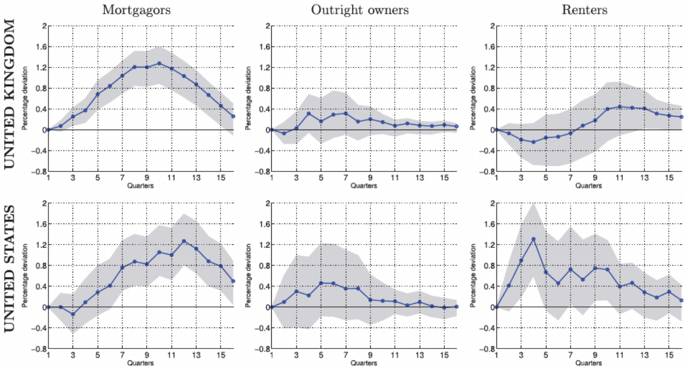
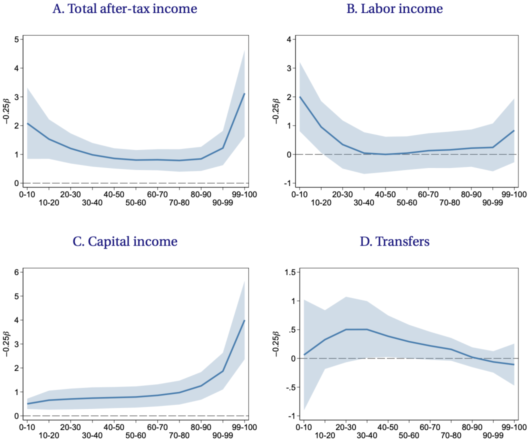


FIGURE 4

Dynamic effects of a 25 bp unanticipated interest rate cut on the expenditure of durable goods by housing tenure group. Gray areas are bootstrapped 90% confidence bands. Top row: U.K. (FES/LCFS data). Bottom row: U.S. (CEX data).

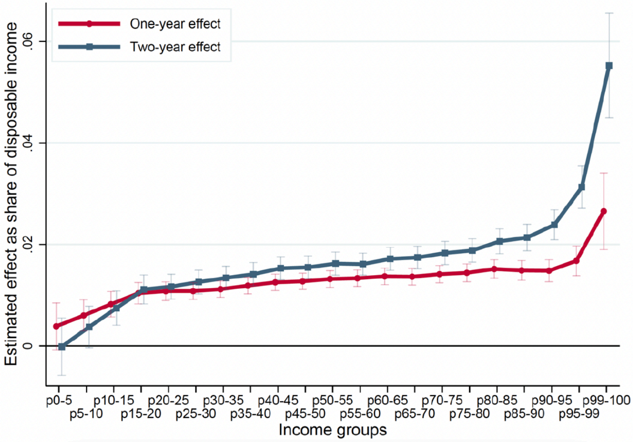
Monetary policy & inequality

Results from Amberg et al. (2020)



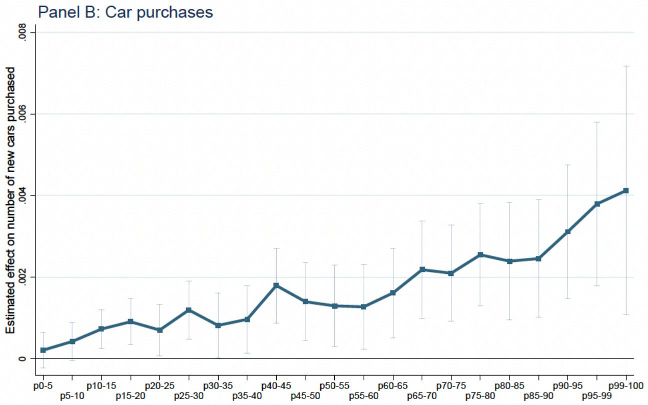
Monetary policy & inequality

Results from Andersen et al. (2020)



Monetary policy & inequality

Results from Andersen et al. (2020)



Outline

Summary

- Main takeaway: IRFs to **policy shocks**—what we've been studying using our time series methods—can identify the effects of switching to different **policy rules**

  Key condition: only (expected) policy instrument path matters to private sector.

- **Practical implications**

  1. Pay particular attention to policy instrument IRFs corresponding to identified time series shocks. More shocks = can construct more counterfactuals.

  2. Role of structural modeling in policy counterfactual analysis is to complete the causal effect matrices Θ. IRF matching is a very natural approach. Christiano-Eichenbaum-Evans

Appendix

IRF matching as minimum-distance estimation

- Model estimation via IRF matching is simply **minimum-distance estimation**
- Formally:
  - The discrepancy is
    $$G_T(\psi \mid Y) = \widehat{m}_T(Y) - \mathbb{E}\left[\widehat{m}_T(Y) \mid \psi\right]$$
    where $\psi \in \Psi$ is the model parameter vector, Y denotes the data, and $\widehat{m}_T(Y)$ is a sample moment of the data
  - The minimum distance estimator is defined as
    $$\widehat{\psi}_{md} \equiv \underset{\psi \in \Psi}{\operatorname{argmin}} \, ||\widehat{m}_T(Y) - \mathbb{E}\left[\widehat{m}_T(Y) \mid \psi\right]||_{W_T}$$
  - As usual, we can characterize the sampling distribution of $\widehat{\psi}_{md}$ using a second-order approximation of the loss function

IRF matching in practice

- From **identification results** to **IRF matching strategy**
  - Our results imply that $\widehat{m}_T(Y) =$ IRFs to identified policy shocks is a particularly promising estimation target
  - Note: best-practice is to make sure that $\mathbb{E}\left[\widehat{m}_T(Y) \mid \psi\right]$ is constructed as the model analogue of the empirical shock identification approach
    See Chari-Kehoe-McGrattan (2008) for a discussion of this point.
    - $\rightarrow$ Thus: to use actual model-implied policy shock IRFs as $\mathbb{E}\left[\widehat{m}_T(Y) \mid \psi\right]$, you need to make sure that your empirical ID asn's hold in the model
- Important **prior examples** of IRF matching for model estimation
  Rotemberg-Woodford (1997), Christiano-Eichenbaum-Evans (2005), Altig et al. (2011)

▸ back