DISCUSSION PAPER SERIES

DP14338
(v. 2)

THE LONG-RUN EFFECTS OF MONETARY POLICY

Òscar Jordà, Sanjay R. Singh and Alan M. Taylor

ECONOMIC HISTORY

INTERNATIONAL MACROECONOMICS AND FINANCE

MACROECONOMICS AND GROWTH

MONETARY ECONOMICS AND FLUCTUATIONS

ISSN 0265-8003

THE LONG-RUN EFFECTS OF MONETARY POLICY

Òscar Jordà, Sanjay R. Singh and Alan M. Taylor

Discussion Paper DP14338
First Published 21 January 2020
This Revision 10 September 2020

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

This Discussion Paper is issued under the auspices of the Centre's research programmes:

- Economic History
- International Macroeconomics and Finance
- Macroeconomics and Growth
- Monetary Economics and Fluctuations

THE LONG-RUN EFFECTS OF MONETARY POLICY

Abstract

Does monetary policy have persistent effects on the productive capacity of the economy? Yes, we find that such effects are economically and statistically significant and last for over a decade based on: (1) identification of exogenous monetary policy fluctuations using the trilemma of international finance; (2) merged data from two new international historical cross-country databases reaching back to the nineteenth century; and (3) econometric methods robust to long-horizon inconsistent estimates. Notably, the capital stock and total factor productivity (TFP) exhibit strong hysteresis, whereas labor does not; and money is non-neutral for a much longer period of time than is customarily assumed. We show that a New Keynesian model with endogenous TFP growth can reconcile these empirical findings.

JEL Classification: E01, E30, E32, E44, E47, E51, F33, F42, F44

Keywords: monetary policy, money neutrality, hysteresis, trilemma, instrumental vari- ables, local projections

Òscar Jordà - ojorda@ucdavis.edu
Federal Reserve Bank of San Francisco

Sanjay R. Singh - sjrsingh@ucdavis.edu
University of California, Davis

Alan M. Taylor - amtaylor@ucdavis.edu
University of California, Davis and CEPR

Acknowledgements
We are thankful to Gadi Barlevy, Susanto Basu, James Cloyne, Stéphane Dupraz, John Fernald, Jordi Galí, Yuriy Gorodnichenko, Pierre-Olivier Gourinchas, Valerie Ramey, Ina Simonovska, Andrea Tambalotti, and many seminar and conference participants at the Barcelona Summer Forum, the NBER Summer Institute IFM Program Meeting, the SED Annual Meeting, the Federal Reserve Banks of Richmond, San Francisco, St. Louis, and Board of Governors, the Midwest Macro Conference, Claremont McKenna College, the Swiss National Bank, the Norges Bank, UC Davis, University of Wisconsin-Madison, Universität Zürich, and Vanderbilt University, who provided very helpful comments and suggestions. Antonin Bergeaud graciously shared detailed data from the long-term productivity database created with Gilbert Cette and Rémy Lecat at the Banque de France. All errors are ours. The views expressed herein are solely those of the authors and do not necessarily represent the views of the Federal Reserve Bank of San Francisco or the Federal Reserve System.

The long-run effects of monetary policy*

Òscar Jordà† Sanjay R. Singh‡ Alan M. Taylor§

September 2020

Abstract

Does monetary policy have persistent effects on the productive capacity of the economy? Yes, we find that such effects are economically and statistically significant and last for over a decade based on: (1) identification of exogenous monetary policy fluctuations using the trilemma of international finance; (2) merged data from two new international historical cross-country databases reaching back to the nineteenth century; and (3) econometric methods robust to long-horizon inconsistent estimates. Notably, the capital stock and total factor productivity (TFP) exhibit strong hysteresis, whereas labor does not; and money is non-neutral for a much longer period of time than is customarily assumed. We show that a New Keynesian model with endogenous TFP growth can reconcile these empirical findings.

JEL classification codes: E01, E30, E32, E44, E47, E51, F33, F42, F44.

Keywords: monetary policy, money neutrality, hysteresis, trilemma, instrumental variables, local projections.

*We are thankful to Gadi Barlevy, Susanto Basu, James Cloyne, Stéphane Dupraz, Gauti Eggertsson, John Fernald, Jordi Galí, Yuriy Gorodnichenko, Pierre-Olivier Gourinchas, Juan Pablo Nicolini, Valerie Ramey, Ina Simonovska, Andrea Tambalotti, and many seminar and conference participants at the Barcelona Summer Forum, the NBER Summer Institute IFM Program Meeting, the Econometric Society World Congress, the SED Annual Meeting, the Federal Reserve Banks of Richmond, San Francisco, St. Louis, and Board of Governors, the Midwest Macro Conference, Claremont McKenna College, the Swiss National Bank, the Norges Bank, UC Davis, University of Wisconsin-Madison, Universität Zürich, and Vanderbilt University, who provided very helpful comments and suggestions. Antonin Bergeaud graciously shared detailed data from the long-term productivity database created with Gilbert Cette and Rémy Lecat at the Banque de France. All errors are ours. The views expressed herein are solely those of the authors and do not necessarily represent the views of the Federal Reserve Bank of San Francisco or the Federal Reserve System.

†Federal Reserve Bank of San Francisco and Department of Economics, University of California, Davis (oscar.jorda@sf.frb.org; ojorda@ucdavis.edu).

‡Department of Economics, University of California, Davis (sjrsingh@ucdavis.edu).

§Department of Economics and Graduate School of Management, University of California, Davis; NBER; and CEPR (amtaylor@ucdavis.edu).

> Are there circumstances in which changes in aggregate demand can have an appreciable, persistent effect on aggregate supply?
>
> — Yellen (2016)

1. Introduction

What is the effect of monetary policy on the long-run productive capacity of the economy? Since at least Hume (1752), macroeconomics has largely operated under the assumption that money is neutral in the long-run, and a vast literature spanning centuries has gradually built the case (see, e.g., King and Watson, 1997, for a review). Contrary to this monetary canon, we find evidence rejecting long-run neutrality.

Our investigation of monetary neutrality rests on three pillars. First, it is essential to identify exogenous movements in interest rates to obtain a reliable measure of monetary effects and avoid confounding. Second, we focus on long-run outcomes, so we need a large sample based on large-sample time series data and, if possible, a wide panel of countries to obtain statistical power. Third, as we show below, the empirical method used can make a big difference: common approaches are designed to maximize short-horizon fit, but we need methods that are consistent over longer spans of time. We discuss how we build on each of these three pillars next.

On identification, the first pillar, we exploit the trilemma of international finance (see, e.g., Obstfeld, Shambaugh, and Taylor, 2004, 2005; Shambaugh, 2004). The key idea is that when a country pegs its currency to some base currency, but allows free movement of capital across borders, it effectively loses control over its own domestic interest rate: a correlation in home and base interest rates is induced, which is exact when the peg is hard and arbitrage frictionless, but is generally less than one otherwise. Insofar as base rates are determined by base country conditions alone, they provide a potential source of exogenous variation in home rates. We theoretically ground this identification strategy in a canonical New Keynesian small open economy model (Schmitt-Grohé and Uribe, 2016; Fornaro and Romei, 2019). Specifically, we derive analytical results to show formally, for the first time, how a trilemma-based identification approach recovers the exact monetary policy impulse response function of interest, with or without the presence of spillover effects.

Second, moving on to the data pillar, we rely on two new macro-history databases spanning 125 years and 17 advanced economies. First, we use the data in Jordà, Schularick, and Taylor (2017), available at `www.macrohistory.net/data`. This "JST Database" contains key macroeconomic series, such as output, interest rates, as well as inflation, credit, and many other potentially useful control variables for our analysis. Second, to allow a Solow decomposition of output into its components, we incorporate data from Bergeaud, Cette, and Lecat (2016), available at `http://www.longtermproductivity.com`.[1] Their data series include observations on investment in machines and buildings, number of employees, and hours worked. With these variables, we can construct measures of total factor productivity (TFP), as we show later, and decompose impulse responses for output into TFP, capital input, and labor input, to pinpoint the important channels of the hysteresis mechanism that we have uncovered.

The third and final pillar of our analysis has to do with the econometric approach. We use local projections (Jordà, 2005) in order to get more accurate estimates of the impulse response function (IRF) at longer horizons. As we show formally, as long as the truncation lag in local projections is chosen to grow with the sample size (at a particular rate that we make specific below), local projections estimate the impulse response consistently at any horizon. Other procedures commonly used to estimate impulse responses do not have this property (see, e.g., Lewis and Reinsel, 1985; Kuersteiner, 2005), and this—among several other reasons—may explain the failure of the prior literature to discern the highly persistent effects we document here.

Supported by these three pillars we show that, surprisingly, monetary policy affects TFP, capital accumulation, and the productive capacity of the economy for a very long time. In response to an exogenous monetary shock, output declines and does not return to its pre-shock trend even twelve years thereafter. Next, we investigate the source of this hysteresis and find that capital and TFP experience similar trajectories to output. In contrast, total hours worked (both hours per worker and number of workers) return more quickly to the original trend. Hence, our new findings are distinct from the usual labor hysteresis mechanism previously emphasized in the literature (see, e.g., Blanchard and Summers, 1986; Galí, 2015a; Blanchard, 2018). These results have important implications for how we think about standard models of monetary economies.

How do our findings stack up against the state of knowledge? A voluminous literature based on post-WW2 U.S. data has examined the causal effects of monetary policy (see, e.g., Ramey, 2016; Nakamura and Steinsson, 2018, for a detailed review), but the evidence on long-run neutrality is, at best, mixed (King and Watson, 1997). An important exception is the work of Bernanke and Mihov (1998), which fails to reject long-run neutrality, but finds that the point estimates of GDP response to monetary innovations do not revert to zero even after ten years. Mankiw (2001) interprets this non-reversal as potential evidence of long-run non-neutrality.[2] Recently, Moran and Queralto (2018) use three-equation VAR

[1] We are particularly thankful to Antonin Bergeaud for sharing some of the disaggregated series from their database that we use to construct our own series of adjusted TFP.

[2] Mankiw notes (emphasis added): "Bernanke and Mihov estimate a structural vector autoregression and present the impulse response functions for real GDP in response to a monetary policy shock. (See their Figure

models that find a similar a link between monetary policy and TFP growth, which we also see as a key factor in understanding non-neutrality.

To reconcile our empirical results with traditional economic models, we develop a theoretical model using the formulation in Stadler (1990) to characterize endogenous TFP growth in a parametrically convenient way in an extension of our small open economy New Keynesian model. In particular, we estimate the *hysteresis elasticity* directly from our estimated responses of TFP and output using a two-step, classical minimum distance procedure. Our estimated elasticity is similar, for example, to that assumed in DeLong and Summers (2012). In our model, a contractionary monetary policy shock lowers output, which temporarily slows down TFP growth. In turn, this slowdown in TFP growth results in permanently lower levels of output and capital, even though labor returns to the stationary equilibrium quickly, as we find in the data.

Our results and our model also connect to a more recent literature on productivity research. Barro (2013), for example, provides evidence that high levels of inflation result in a loss in the *rate* of economic growth. Work by Gopinath, Kalemli-Özcan, Karabarbounis, and Villegas-Sanchez (2017) links the level of interest rates to the *level* of productivity. Baqaee and Farhi (2019) construct a general framework where monetary shocks may affect allocative efficiency. Work by Benigno and Fornaro (2018) links low interest rates to the rate of *growth* of productivity. Moreover, although in this paper we do not need to take a stand on the precise details of the hysteresis mechanism, many are consistent with our setup and several recent papers have examined different potential micro-foundations of endogenous hysteresis effects via TFP growth (Fatás, 2000; Barlevy, 2004; Anzoategui, Comin, Gertler, and Martinez, 2019; Bianchi, Kung, and Morales, 2019).[3] Going beyond our paper, hysteresis matters for how we build models of monetary economies and what optimal monetary policy is in those models: the welfare implications could be substantial.

2. Identification

The trilemma of international finance gives a theoretically justified source of exogenous variation in interest rates (Jordà, Schularick, and Taylor, 2020). The logic is straightforward

III.) Their estimated impulse response function does not die out toward zero, as is required by long-run neutrality. Instead, the point estimates imply a large impact of monetary policy on GDP even after ten years. Bernanke and Mihov don't emphasize this fact because the standard errors rise with the time horizon. Thus, if we look out far enough, the estimated impact becomes statistically insignificant. *But if one does not approach the data with a prior view favoring long- run neutrality, one would not leave the data with that posterior. The data's best guess is that monetary shocks leave permanent scars on the economy.*"

[3] See Cerra, Fatás, and Saxena (2020) for a detailed review of literature on hysteresis and business cycles. In a recent work, Meier and Reinelt (2020) provide evidence of increased misallocation following contractionary monetary policy shocks.

under a hard peg with free capital mobility short-term rates in two countries, home and base, will be arbitraged. With *strict* interest parity, rates are exactly correlated.

But even under soft pegs or dirty floats, with frictions or imperfect arbitrage, a non-zero interest rate correlation between home and base is enough for identification using instrumental variables, as is well-known. In this section we present an open economy model to make formal the conditions for identification, even in the presence of spillovers via non–interest rate channels. This level of detail allows us to construct the econometric estimation procedures around the propositions derived from the model.

2.1. Baseline Model

We build on a standard open economy setup widely used today as in Schmitt-Grohé and Uribe (2016), and Fornaro and Romei (2019).[4] Various elements of this framework appear also in Benigno, Fornaro, and Wolf (2020), Farhi and Werning (2017), Fornaro (2015), Gourinchas (2018) among others. Our aim is not a new model, but how theory maps rigorously into our trilemma identification scheme and guides our econometric approach. As the model is standard, many details are relegated to the Appendix. In Appendix H we also show how the same results hold in a Mundell-Fleming-Dornbusch model with additional financial channels, as in Gourinchas (2018).

We assume there is perfect foresight. The environment features incomplete international markets with nominal rigidities. We focus on two countries: a large economy that we label the *base* and a small open economy, the home country. If home pegs its currency to the base, then we label it as a *peg*; if it floats with respect to the base, then we label it as a *float*. We begin by describing the small economies first.

Consumers. The household consists of a continuum of consumers, normalized to measure one, each of which supplies labor while taking as given the market wage. There is perfect consumption insurance within the household. The household has CRRA preferences over a composite good and derives disutility from supplying labor. The household problem is

$$\max_{\{C_t, h_t, B_{t+1}\}} \sum_{t=0}^{\infty} \beta^t \left[\log(C_t) - \varphi \frac{h_t^{1+\nu}}{1+\nu} \right] .$$

The composite C_t is a Cobb-Douglas aggregate $C_t = \left(\frac{C_{Tt}}{\omega}\right)^{\omega} \left(\frac{C_{Nt}}{1-\omega}\right)^{1-\omega}$ of a tradable good C_{Tt} and a non-tradable good C_{Nt}, where $\omega \in (0,1)$ is the tradable share, ν is (inverse)

[4]Relative to Fornaro and Romei (2019), we do not model a continuum of small open economies and instead focus on individual small open economies.

Frisch elasticity of labor supply, and φ is a scaling parameter to normalize $h = 1$ in the steady state.

Households can trade in one-period riskless real and nominal bonds. Real bonds are denominated in units of the tradable consumption good and pay gross interest rate R_t, taken as given (i.e., a world real interest rate). Nominal bonds issued by the domestic central bank are denominated in units of domestic currency, and pay gross nominal interest rate R_t^n. That is, R_t^n is the policy rate of interest.

The households' budget constraint in units of domestic currency is as follows

$$P_{Tt}C_{Tt} + P_{Nt}C_{Nt} + P_{Tt}d_t + B_t \;=\; W_t h_t + P_{Tt}Y_{Tt} + P_{Tt}\frac{d_{t+1}}{R_t} + \frac{B_{t+1}}{R_t^n} + T_t + Z_t\,,$$

where P_{Tt} and P_{Nt} are the prices of tradable and non-tradable goods in local currency; d_t is the level of real debt in units of tradable good assumed in period $t-1$ and due in period t; B_t is the level of nominal debt in units of local currency assumed in period $t-1$ and due in period t; W_t is the nominal wage; T_t are nominal lump-sum transfers from the government; Z_t are nominal profits from domestic firms owned by households; and $Y_{Tt} > 0$ is the endowment of tradable goods received by the households.

The household chooses a sequence of $\{C_{Tt}, C_{Nt}, h_t, d_{t+1}, B_{t+1}\}$ to maximize lifetime utility subject to the budget constraint, taking initial bond holdings as given. Labor is immobile across countries, so the wage level is local to each small open economy. The world real interest rate is taken as given, so there can be dependence on initial conditions.

The first-order conditions for the household's optimization problem are

$$\frac{1}{C_{Tt}} = \frac{\beta}{C_{Tt+1}}\; R_t\,, \tag{1}$$

$$\frac{1}{C_{Tt}} = \frac{\beta}{C_{Tt+1}}\; \frac{R_t^n P_{Tt}}{P_{Tt+1}}\,, \tag{2}$$

$$\frac{P_{Nt}}{P_{Tt}} = \frac{(1-\omega)C_{Tt}}{\omega C_{Nt}}\,, \tag{3}$$

$$\varphi\, h_t^{\nu}\, C_{Tt} = \frac{W_t}{P_{Tt}}\,. \tag{4}$$

We assume that law of one price on the tradable good holds. Let $\mathcal{E}_t$ be the nominal exchange rate for home relative to the base, and let P_t^* be the base price of the tradable good denominated in base currency.[5] Then, we have that $P_{Tt} = \mathcal{E}_t\, P_t^*$. From Equation 1 and

[5] It is common in the small open economy literature to treat price level in the base economy P_t^* as synonymous for price level of tradable goods in the base economy P_{Tt}^*.

Equation 2 we can then derive the interest rate parity condition,

$$R_t^n = R_t \frac{P_{Tt+1}}{P_{Tt}} = R_t \frac{\mathcal{E}_{t+1}}{\mathcal{E}_t} \frac{P_{t+1}^*}{P_t^*} . \tag{5}$$

To ensure stationarity under incomplete markets, we follow Schmitt-Grohé and Uribe (2003); Uribe and Schmitt-Grohé (2017) and assume that the home real interest rate is related to foreign real interest rate through a debt-elastic interest rate premium,

$$R_t = R_t^* + \psi(e^{d_{t+1}-\bar{d}} - 1) . \tag{6}$$

We emphasizes that this is a standard technical requirement for solving these types of models, but we will henceforth work in the limit case $\psi \to 0$, so the financial constraint is vanishingly small.[6]

Production and nominal rigidities. The non-tradable consumption good is a Dixit-Stiglitz aggregate over a continuum of products $C_{Nt}(i)$ produced by monopolistically competitive producers indexed by i, with $C_{Nt} \equiv (\int_0^1 C_{Nt}(i)^{(\epsilon_p-1)/\epsilon_p} di)^{\epsilon_p/(\epsilon_p-1)}$. Each firm i in home produces a homogenous good with technology given by $Y_{Nt}(i) = L_{Nt}(i)$, taking the demand for its product as given by $C_{Nt}(i) = (P_{Nt}(i)/P_{Nt})^{-\epsilon_p} C_{Nt}$, where we use the price index of the non-tradable good composite, $P_{Nt} = (\int_0^1 P_{Nt}(i)^{1-\epsilon_p} di)^{1/(1-\epsilon_p)}$. Individual firms reset prices at random intervals with Calvo (1983) price setting. Full details are found in the appendix.

Fiscal policy and the bond markets. The portfolio allocation between the real and nominal bond is not determinate in this type of model. To ensure determinacy, and since all agents at home are identical, we now assume that home domestic nominal bonds are in net zero supply, i.e., $B_{t+1}^n = 0$. We also assume that the home fiscal authority follows a balanced budget every period.[7]

Market clearing. We impose that the non-tradable goods market has to clear at home, implying that production of non-tradable goods (net of misallocation costs due to price dispersion) must equal the consumption demand for non-tradable goods. However, since we will focus on a first-order approximation around a deterministic steady state, this price

[6] Equivalently, we could introduce a borrowing constraint on the household's end-of-period debt position. The vanishingly small ψ is not necessary for our results, and is assumed for transparent analysis.

[7] We assume appropriate government subsidies financed by lumpsum taxes to eliminate monopoly rents in the intermediate goods sector.

dispersion term will be treated as zero, so

$$L_{Nt} = Y_{Nt} = C_{Nt}.^{8} \tag{7}$$

Finally, the external budget constraint of the economy must be satisfied every period:

$$C_{Tt} + d_t \;=\; Y_{Tt} + \frac{d_{t+1}}{R_t}\,. \tag{8}$$

Construction of small open economy GDP. Our key outcome variable of interest is the real GDP in the small economy. To make the connection with our empirical counterparts, and to keep our baseline discussion focused, for now we construct this real GDP variable using constant aggregation weights implied by the Cobb-Douglas aggregator.

Clearly, variation in aggregation weights can cause changes in real GDP in a multiple sector economy, and this definition abstracts from such potential index number problems. That said, we present analytical results in an environment with time-varying aggregation weights in Appendix D.

Monetary policy. The policy rate is the home nominal interest rate on one-period domestic currency bonds. For our small open economies hit with a base economy interest rate shock, we first consider the following two possible home policy configurations, a hard peg and a pure float (we will relax this and look at a dirty float or soft peg below):

- A *hard peg* fixes the nominal exchange rate at a given level. Without loss of generality, we assume the rule

  $$\mathcal{E}_t = 1\,. \tag{9}$$

  By Equation 62, and in the limit when $\psi \to 0$, there is perfect passthrough from base economy interest rate changes into home nominal interest rates, hence $R^n_t - \bar{R}^n = R^*_t - \bar{R}^*$, where $\bar{R}^n$ and $\bar{R}^*$ denote the steady state levels of nominal interest rates in the home and the base economy, respectively.

- Instead, a *pure float* economy sterilizes base interest rate movements so that $R^n_t - \bar{R}^n = 0$. In our model, this policy scenario can be implemented with a strict producer inflation targeting rule. We assume that the net (gross) producer price inflation target is zero (one).[9]

  $$\Pi_{Nt} = 1. \tag{10}$$

[8] In models with industry-specific labor, the price dispersion term does not appear in the aggregate resource constraints.

[9] This is a standard assumption in New Keynesian models. We could equivalently introduce a different value for the producer price inflation target while assuming perfect indexation to this target by all producers.

As shown by Uribe and Schmitt-Grohé (2017), strict producer price inflation targeting is the optimal monetary policy rule in such an environment.

Based on these two alternatives, the question now is to determine the conditions under which, using *base* interest rate shocks as instruments, one can recover exactly the same impulse response as that generated by a standard domestic monetary policy shock. We label this latter reference scenario as the *benchmark*, where policy is characterized as follows:

- In the *benchmark* economy, the home nominal interest rate follows an exogenous path subject to policy shocks ε_t,

$$R_t^n = \bar{R}^n e^{\varepsilon_t}\,. \tag{11}$$

  Since we are simulating responses to one-time shocks, we interpret this policy rule assumption as equivalent to that of *temporary interest rate peg* made in the zero lower bound literature (Eggertsson and Woodford, 2003; Werning, 2011). Once the shock abates, a policy rule that maintains local determinacy (Blanchard and Kahn, 1980) is expected to hold in those environments with temporary interest rates at the zero lower bound. We assume a similar equilibrium selection device that the economy returns back to the same deterministic steady state.[10,11]

The base economy. The small country takes the path of prices P_t^* and real interest rates R_t^* in the base as given. Without loss of generality, we therefore assume rigid prices in the base economy, with $P_t^* = 1$. Our focus is then the impulse response of small open economy output—under peg or float—following a shock in R_t^*, and how it compares to the impulse response of a benchmark economy following a domestic policy shock.

Equilibrium and solution method. We present the equilibrium conditions in Appendix B. We log-linearize the model around a deterministic steady state, and solve the model backwards from the deterministic steady state, assuming perfect foresight. Variables in hats will denote deviations from steady state. In the long-run PPP is assumed to hold, and the economy returns to $d = \bar{d}$, where $\bar{d}$ is the level of debt in a deterministic steady state. Policy shock sequences are i.i.d. changes to $\{\hat{R}_t^*\}$ in the peg and float home economy configurations, and to $\{\varepsilon_t\}$ in the benchmark economy configuration.

[10]Similar solution methods to do counterfactual policy simulations have been developed for economies away from the zero lower bound (Laséen and Svensson, 2011; Guerrieri and Iacoviello, 2015; Christiano, 2015). Embedding an endogenous policy transmission through inflation targeting, while the shock is on, does not change our theoretical results since we are identifying responses to non-systematic components of monetary policy.

[11]We solve the economy for the perfect foresight solution, and assuming that the economy returns back to the initial steady state. While interest rate pegs are known to cause indeterminacy issues, we maintain this assumption here to keep our results comparable to the hard peg and pure float economy.

2.2. Identification with trilemma: analytical results

We now present the core theoretical results of our paper as a series of propositions. We focus on closed form analytical results. We begin by noting that tradable good consumption, as well as real debt choice, is independent of the monetary policy regime.[12] This greatly simplifies the analysis.

Proposition 1. *The responses to a base interest rate shock of tradable consumption and the domestic real interest rate (on bonds denominated in tradable goods) do not depend on whether the home economy pegs or floats.*

Proof. The proof follows directly from Equation 1, Equation 62, and Equation 8, which define the competitive equilibrium for $\{C_{Tt}, r_t, d_{t+1}\}$ a given sequence of $\{r_t^*\}$.[13] □

The upshot of this well-known result is that we can now separate the determination of all remaining variables from $\{C_{Tt}, r_t, d_{t+1}\}$. Crucially, we will take the path of these variables as given across various policy regimes for the same foreign shock.

Definition 1. Expressions 12, 13, 14 below summarize the log-linear equilibrium conditions for $\{Y_{Nt}, \hat{\Pi}_{Nt}, \hat{R}_t^n, \hat{\mathcal{E}}_t\}$ under perfect foresight in a small open economy. In addition to these expressions, the policy regime is summarized by either Equation (peg), (float), or (benchmark). These conditions are for a given sequence of $\{\hat{Y}_{Tt}, \hat{C}_{Tt}, \hat{d}_{t+1}, \hat{R}_t, \hat{R}_t^*\}$ and assuming $\psi \to 0$, specifically:

$$\hat{Y}_{Nt} = \hat{Y}_{Nt+1} - (\hat{R}_t^n - \hat{\Pi}_{Nt+1})\,, \tag{12}$$

$$\hat{\Pi}_{Nt} = \beta\hat{\Pi}_{Nt+1} + \kappa\hat{Y}_{Nt}\,, \tag{13}$$

$$\hat{\mathcal{E}}_{t+1} - \hat{\mathcal{E}}_t = \hat{R}_t^n - \hat{R}_t^*\,, \tag{14}$$

$$\hat{\mathcal{E}}_{t+1} - \hat{\mathcal{E}}_t = 0\,, \tag{peg}$$

$$\hat{R}_t^n = 0\,, \tag{float}$$

$$\hat{R}_t^n = \varepsilon_t\,. \tag{benchmark}$$

[12] This result is well noted in the literature at least since Obstfeld and Rogoff (1995a, Appendix) in the case with fixed base economy interest rates. Uribe and Schmitt-Grohé (2017, Section 9.5) generalized the result to settings where the inter-temporal elasticity of substitution is equal to the intra-temporal elasticity of substitution between tradable and non-tradable goods.

[13] The key difference between a peg and a float comes from whether the nominal exchange rate is used to counter the passthrough of foreign rates into domestic policy rates. There is an extant literature in open economy macroeconomics that has emphasized this insight, most recently articulated by Farhi and Werning (2012), Fornaro (2015), as well as Schmitt-Grohé and Uribe (2016) upon which we build.

Based on these expressions, hence consider the log-linear equilibrium of a small open economy under a (hard) peg and the log-linear equilibrium of the benchmark economy with a domestic policy shock. Assume real GDP is constructed with constant and identical aggregation weights in the two economies. Then the following proposition holds:

Proposition 2 (impulse response equivalence: hard pegs)**.** *The response of real GDP to a base interest rate shock in a peg is identical to the response of real GDP to a domestic policy shock of the same magnitude and persistence in a benchmark economy.*

Proof. The policy rule under a peg prevents any adjustment in nominal exchange rates, i.e., $\hat{\mathcal{E}}_{t+1} - \hat{\mathcal{E}}_t = 0$. Hence the path of nominal interest rates in a peg economy, $\hat{R}^*_t$, is identical to the path in the benchmark economy for $\varepsilon_t = \hat{R}^*_t$. The equilibrium conditions summarized by Equation 12 and Equation 13, with the same terminal condition, then solve for an identical sequence of $\{\hat{Y}_{Nt}, \hat{\Pi}_{Nt}\}$ in the two economies. We provide an exact solution in Appendix C. Since the tradable output is an exogenous endowment, and we have assumed constant aggregation weights, the response of real GDP is identical across the two economies. □

2.3. Departures from the baseline model

We now extend the baseline model in two ways. First, we allow for imperfect interest rate pass-through from the base rate into the home economy. This can happen either because the home economy is in a soft peg or in dirty float regime. Given this setting, we then show that one can still use base country rates to construct the equivalent response to a monetary shock in the benchmark economy. Second, we consider other channels through which base interest rate shocks can spill over into the home economy. Hence we show how one can adjust the response to base country rates and still obtain the equivalent benchmark economy response to a monetary shock.

2.3.1 Soft pegs and dirty floats

Define the imperfect pass-through (whether for a soft peg, or a dirty float) of base rates to home rates using a coefficient $0 < \lambda \leq 1$ such that: $R^n_t - \bar{R}^n = \lambda(R^*_t - \bar{R}^*)$.

Proposition 3 (impulse response equivalence: imperfect pass-through)**.** *Consider the log-linear equilibrium of a small open economy with imperfect pass-through and the log-linear equilibrium of the benchmark economy with a domestic policy shock. Assume real GDP is constructed with constant and identical aggregation weights in the two economies. The response of real GDP with imperfect pass-through to a base economy interest rate shock is a fraction λ of the response of real GDP in a benchmark economy to a domestic policy shock of same magnitude and persistence.*

Proof. In an imperfect pass-through economy, we can write

$$\hat{R}_t^n = \lambda \hat{R}_t^* \, ; \quad \lambda \in (0,1] \, . \tag{15}$$

From the UIP conditions in Equation 14, the expected exchange rate appreciation is now $(1-\lambda)\hat{R}_t^*$. With the terminal condition of $\hat{Y}_{Nt} = \hat{\Pi}_{Nt} = 0$, it follows that the response of $\hat{Y}_{Nt}$ to $\hat{R}_t^*$ under an imperfect pass-through economy is λ times the response of $\hat{Y}_{Nt}$ to $\varepsilon_t = \hat{R}_t^*$ in the benchmark economy. □

2.3.2 Spillovers

If there are other channels through which base interest rates can affect the model equilibrium, these *spillovers* will affect the previous results derived for pegs and imperfect pass-through economies. The equivalency with the impulse response of output in the benchmark economy will break down.

To see this, consider the following postulated relationship between tradable output and the base real interest rate,

$$\hat{Y}_{Tt} = \alpha \hat{R}_t^* \, . \tag{16}$$

where $\alpha < 0$.[14] Such a relationship is often embedded into open economy models through modeling of the export demand (e.g, see Galí and Monacelli, 2016).

Intuitively, the home economy's ability to sell its export good to the base (or any economy pegged to the base) is now demand constrained. This demand is not perfectly elastic, but depends on the state of consumption demand in the base economy, which in turn depends on the base real rate. Equation 16 is a reduced-form expression of this dependence.[15]

Proposition 4 (spillovers in a peg)**.** *Consider the log-linear equilibrium of a small open economy under a peg with spillovers (i.e., extended with Equation 16), and the log-linear equilibrium of the benchmark economy with a domestic policy shock. Assume real GDP is constructed with constant and identical aggregation weights in the two economies. Denote the response of real GDP in a peg to a unit and i.i.d. base economy interest rate shock with γ_p, and the response of real GDP in the*

[14]With persistent shocks, α may not be time-invariant. Here we ignore the time-subscript for analytical clarity.

[15]It is interesting to note that α can be positive in models with endogenous production of tradable goods. In Appendix E, we show that this is indeed the case with endogenous production of tradable good with labor input derived from an economy-wide labor market. In that environment, our baseline estimates have a downward bias. We think the case of $\alpha < 0$ is more realistic in a world where contractionary policy shocks in the US economy have contractionary spillovers into rest of the world.

benchmark economy to a unit and i.i.d. domestic policy shock with β. Then,

$$\beta = \gamma_p - \frac{P_T Y_T}{PY}\alpha\,.$$

Proof. From Definition 1, which delineates the log-linear equilibrium conditions in a small open economy, the response of non-tradable output Y_{Nt} in the peg economy to a base-economy interest rate shock is identical to that in the benchmark economy to a similar ε_t shock sequence. In the presence of the spillover, tradable output contracts with an increase in base interest rates, while it is unaffected in the benchmark economy.

Using the construction of real GDP described in Section 2.1, we can compute the exact difference in the impulse responses of real GDP as:

$$\hat{Y}_t^{peg} - \hat{Y}_t^{benchmark} \quad = \quad \underbrace{\frac{P_T Y_T}{PY}}_{\substack{\text{tradable output}\\ \text{share in GDP}}} \times \left(\hat{Y}_{Tt}^{peg} - \hat{Y}_{Tt}^{benchmark}\right) \;=\; \frac{P_T Y_T}{PY}\alpha \hat{R}_t^*\,.$$

Now we assume that the base shock equals the benchmark policy shock, $\hat{R}_t^* = \varepsilon_t$, so we have that

$$\underbrace{\frac{\hat{Y}_t^{peg}}{R_t^*}}_{\equiv \gamma_p} - \underbrace{\frac{\hat{Y}_t^{benchmark}}{\epsilon_t}}_{\equiv \beta} = \frac{P_T Y_T}{PY}\alpha\,.$$

Hence,

$$\beta = \gamma_p - \frac{P_T Y_T}{PY}\alpha \tag{17}$$

□

A corollary of Proposition 4 applies to an imperfect pass-through economy.

Corollary 1. *Consider the log-linear equilibrium of a small open imperfect pass-through economy (extended with Equation 16 and the log-linear equilibrium of the benchmark economy with a domestic policy shock. Assume real GDP is constructed with constant and identical aggregation weights in the two economies. Denote the response of real GDP in the imperfect pass-through economy to a unit, i.i.d. base economy interest rate shock with γ_p, and the response of real GDP in the benchmark economy to a unit, i.i.d. domestic policy shock with β. Then,*

$$\beta = \frac{\gamma_p}{\lambda} - \frac{P_T Y_T}{PY}\alpha$$

To sum up, this last result shows that the same logic applies to the continuum of regimes from hard peg ($\lambda = 1$) to pure float ($\lambda = 0$), with appropriate scaling of responses by λ. Thus, for estimation purposes, we may draw on information from any economy within this continuum, not just those with regimes at the extremes.

2.4. Model implications for econometric identification

The final model just introduced, with spillovers, explains how base country monetary policy can affect the output of tradable goods (via export demand shifts) as well as the output of nontradable goods (via interest arbitrage and conventional domestic demand shifts). These spillover effects onto smaller open economies depend on the share of tradable output in their GDP. Using the insights and notation from the model, in this section we explore its implications for the identification of our impulse responses.

Disciplining the spillover coefficient. As in Equation 15, we assume imperfect pass-through of base rates into home rates. In regression form, this can be expressed as

$$\hat{R}^n_t = \lambda \hat{R}^*_t + v_t \,, \tag{18}$$

where, as before, $\hat{R}^n_t$, and $\hat{R}^*_t$ are in deviations from steady state, and $\lambda \in [0,1]$ is the spillover coefficient, and is possibly different for country-time pairs nominally classified as *pegs* versus *floats*. We omit the constant term without loss of generality and we assume that v_t is a well-behaved, white noise error term. For now, it is convenient to leave more complex dynamic specifications aside to convey the intuition simply.

Similarly, Equation 17 in regression form can be expressed as

$$\hat{Y}_t = \hat{R}^n_t \beta + \hat{R}^*_t \theta + u_t \,, \tag{19}$$

where here too $\hat{Y}_t$, $\hat{R}^n_t$, and $\hat{R}^*_t$ are deviations from steady state. For now, we leave unspecified whether $\hat{Y}_t$ belongs to a peg or a float. Note that under Equation 17, we have

$$\theta = \frac{P_T Y_T}{PY} \alpha \,, \tag{20}$$

that is, the share of tradable export output in GDP, which we denote $\Phi = {}^{P_T Y_T}/_{PY}$, scaled by the parameter α, which determined how $\hat{R}^*_t$ affects tradable output.

According to the model, there are two main reasons we might expect $\theta \to 0$. One reason is that home output is dominated by non-tradable output. In the JST database we find indeed that in our advanced economies, over 150 years of history, tradable export shares are 30% at most, and usually lower in the 10%–20% range, so $\Phi \leq 0.3$. The second is the parameter α, which measures the spillover effect of base country rates $\hat{R}_t^*$ on tradable export demand at home. It is fair to assume that this effect will be *at most* as strong as the effect of domestic rates on tradable output, so $\alpha \leq \beta$. These observations will come in handy later as a way to provide a bound on potential biases due to spillover effects.

IV estimator with no spillovers. How then can we estimate β using the trilemma approach? In practice, as is well known, direct OLS estimation of Equation 19 likely delivers a biased estimate of the desired coefficient β since $\hat{R}_t^n$ will be endogenously determined by the monetary authority, depending on the exchange rate regime of the country considered.

However, one can exploit the model (and the trilemma mechanism) as follows. Consider premultiplying Equation 19 by $\hat{R}_t^*$. Taking expectations, and noting that $E(\hat{R}_t^* u_t) = 0$ since base country rates are assumed to be independent of small country factors, we get

$$\frac{E(\hat{R}_t^* Y_t)}{E(R_t^{*2})} = \frac{E(\hat{R}_t^* \hat{R}_t^n)}{E(R_t^{*2})} \beta + \theta \,. \tag{21}$$

Now notice that the LHS of Equation 21 is the reduced-form coefficient γ (introduced in the previous section) in the auxiliary regression

$$\hat{Y}_t = \hat{R}_t^* \gamma + \eta_t \,, \tag{22}$$

where η_t is a well behaved, white noise error term uncorrelated with u_t and v_t. Similarly, the ratio of expectations from the RHS of Equation 21 is the coefficient λ in Equation 18. Of course, Equation 18 and Equation 22 are just another way of computing the familiar traditional IV estimator for Equation 19.

Now in the absence of spillovers, $\theta = 0$, and putting these elements together we obtain

$$\gamma = \lambda \beta \quad \Longrightarrow \quad \beta = \frac{\gamma}{\lambda} \,, \tag{23}$$

where estimates for γ and λ are easily obtained from reduced-form regressions Equation 18 and Equation 22. Hence, β is identified and the estimator based on the two reduced-form regressions is simply the traditional IV estimator of Equation 19 using $\hat{R}_t^*$ as the instrument.

Our data come from two subpopulations, pegs and floats, which principally differ in the degree to which $\lambda \to 1$. In practice we hesitate to impose the same parameters across both subpopulations thus allowing for different γ and λ, so the reduced form regressions are

$$\hat{Y}_t = D_t^P \hat{R}_t^* \gamma_P + D_t^F \hat{R}_t^* \gamma_F + \eta_t \,, \tag{24}$$
$$\hat{R}_t^n = D_t^P \hat{R}_t^* \lambda_P + D_t^F \hat{R}_t^* \lambda_F + v_t \,, \tag{25}$$

where $D_t^P = 1$ for pegs, 0 otherwise, and similarly $D_t^F = 1$ for floats, 0 otherwise.

In other words, if there are no spillovers, the IV estimator of β will be the ratio of the weighted average of the γ over the weighted average of the λ: we will be estimating a "model average" β using information from both of the two subpopulations, pegs and floats.

IV estimator with spillovers. What happens if $\theta \neq 0$? In that case, if we used the above estimator, the exclusion restriction is violated and θ introduces bias in our estimates of β. The approach that we take here is to provide a bound for the possible values that $\theta = \Phi\alpha$ can take based on our model, as we discussed earlier. Two quantities guide our choices: (1) the share of tradables in GDP; and (2) the effect of base country rates on tradable output. Note that (1) is directly measurable and, as we argued above, falls typically in the range $\Phi \in [0.1, 0.3]$ in the JST database. Regarding (2), we assume that effect of $\hat{R}_t^*$ on tradable output is, in any case, no larger than the effect of $\hat{R}_t^n$; that is, we assume that domestic interest rates are more influential on domestic output than base rates are, so we impose as a conservative upper bound that $\alpha = \beta$.

Based on these assumptions, we can write $\theta = \Phi\beta$ and employ the calibrated range of values of Φ. Then it is easy to see that one can transform the original Equation 19 as,

$$\hat{Y}_t = (\hat{R}_t^n + \hat{R}_t^* \Phi)\beta + u_t \,, \tag{26}$$

and one can estimate β with this expression using instrumental variables along the lines just discussed using the subpopulations of pegs and floats, that is, with the first stage given by Equation 25.

To sum up, in the empirical work which follows, we estimate the following IV model,

$$\hat{Y}_t = (\hat{R}_t^n + \hat{R}_t^* \Phi)\beta + u_t \,, \tag{27}$$
$$\hat{R}_t^n = D_t^P \hat{R}_t^* \lambda_P + D_t^F \hat{R}_t^* \lambda_F + v_t \,. \tag{28}$$

which we have shown will recover the true impulse response for the benchmark model based on impulse responses for pegs and floats.

3. Data and series construction

Macroeconomic time series. The empirical features motivating our analysis rest on two major international and historical databases.

Data on macro aggregates and financial variables, including assumptions on exchange rate regimes and capital controls, can be found in `www.macrohistory.net/data`. This database covers 17 advanced economies reaching back to 1870 at annual frequency. Detailed descriptions of the sources of the variables contained therein, their properties, and other ancillary information are discussed in Jordà, Schularick, and Taylor (2017) and Jordà, Schularick, and Taylor (2020), as well as references therein. Importantly, we will rely on the trilemma instrument discussed in Jordà, Schularick, and Taylor (2016), and more recently Jordà, Schularick, and Taylor (2020), as the source of exogenous variation in interest rates. The instrument construction details will become clearer in the next section.

The second important source of data relies on the work by Bergeaud, Cette, and Lecat (2016) and available at `http://www.longtermproductivity.com`. This historical database adds to our main database observations on capital stock (machines and buildings), hours worked, and number of employees, and the Solow residuals (raw TFP). In addition, we construct time-varying capital and labor utilization corrected series using the procedure discussed in Imbs (1999) with the raw data from Bergeaud, Cette, and Lecat (2016) to construct our own series of utilization-adjusted TFP. We went back to the original sources so as to filter out cyclical variation in input utilization rates in the context of a richer production function that allows for factor hoarding. We explain the details of this correction in Appendix G.[16]

Trilemma instruments. Guided by our model and identification strategy as discussed above, we divide our sample into three subpopulations of country-year observations. The *bases* will refer to those economies whose currencies serve as the anchor for the subpopulation of pegging economies, labeled as the *pegs*. Other economies, the *floats*, allow their exchange to be freely determined by the market.

Base and peg country codings can be found in Jordà, Schularick, and Taylor (2020, Table 1 and Appendix A), and are based on older, established definitions (Obstfeld, Shambaugh, and Taylor, 2004, 2005; Shambaugh, 2004; Ilzetzki, Reinhart, and Rogoff, 2019). A country i is defined to be a peg at time t, denoted with the dummy variable $D^P_{i,t} = 1$, if it maintained a peg to its base at dates $t-1$ *and* t. This conservative definition serves to eliminate

[16] Our construction of productivity assumes misallocation related-wedges are absent. We have not yet found the data to take into account markups or sectoral heterogeneity in our productivity estimates. See Basu and Fernald (2002) and Syverson (2011) for extensive discussions on what determines productivity.

opportunistic pegging, and it turns out that transitions from floating to pegging and vice versa represent less than 5% of the sample, the average peg lasting over 20 years. Interestingly, pegs are, on average, more open than floats.[17] Finally, let $D^F_{i,t} = 1 - D^P_{i,t}$ denote a non-peg, i.e., float. The choice of exchange rate regime is treated as exogenous, and indeed we find zero predictability of the regime based on macroeconomic observables in our advanced economy sample. Regimes are also highly persistent in this sample which excludes emerging and developing countries, in contrast to the findings of limited persistence for the full cross-section of countries as in Obstfeld and Rogoff (1995b).

Based on this discussion, we construct an adjusted instrument as follows. Let $\Delta R_{i,t}$ denote changes in country i's short-term nominal interest rate at date t, let $\Delta R_{b(i,t),t}$ denote the change in short term interest rate of country i's base country $b(i,t)$, and let $\Delta \tilde{R}_{b(i,t),t}$ denote its predictable component explained by a variety of base country macroeconomic aggregates. Loosely speaking, think of it as the rate that would be predicted by a policy rule,[18] and hence, using the notation from the previous section, to a first approximation denote $\Delta \widehat{R}_{b(i,t),t} = (\Delta R_{b(i,t),t} - \Delta \tilde{R}_{b(i,t),t})$. However, since countries in a given year may not be perfectly open to capital flows, we then scale the base shock, adjusting for capital mobility using the capital openness index of Quinn, Schindler, and Toyoda (2011), denoted $k_{i,t} \in [0, 1]$.

The resulting trilemma instruments adjusted for capital mobility are thus defined as

$$z^j_{i,t} \equiv D^j_{i,t}\, k_{i,t}\, \Delta \widehat{R}_{b(i,t),t}\,; \quad j = P, F \tag{29}$$

where P refers to *pegs* and F refers to *floats*.

4. Solving inconsistency in long-horizon impulse responses

In thinking about the propagation of a shock, specially to distant horizons, it is generally convenient to allow for generous lag structures—and in the limit, allowing for possibly infinite lags. Infinite dimensional models have a long tradition in econometric theory and form the basis for many standard results. For example, Berk (1974) considers the problem

[17] In the full sample, the capital openness index averages 0.87 for pegs (with a standard deviation of 0.21) and 0.70 for floats (with standard deviation 0.31). After WW2 there is essentially no difference between them. The average is 0.76 for pegs and 0.74 for floats with a standard deviation of 0.24 and 0.30 respectively. See Jordà, Schularick, and Taylor (2020) for further details on the construction of the instrument.

[18] The list of controls used to construct $\widehat{\Delta i}_{b(i,t),t}$ include log real GDP; log real consumption per capita; log real investment per capita; log consumer price index; short-term interest rate (usually a 3-month government bill); long-term interest rate (usually a 5-year government bond); log real house prices; log real stock prices; and the credit to GDP ratio. The variables enter in first differences except interest rates. Contemporaneous terms (except for the left-hand side variable) and two lags are included.

of estimating the spectral density of an infinite order process using finite autoregression. In multivariate settings, Lewis and Reinsel (1985) establish the consistency and asymptotic normality of finite order approximations to an infinite order multivariate system. Kilian (1998) shows that the finite sample biases of the underlying finite order autoregressions can induce severe bias on impulse response bootstrap inference based on vector autoregressions (VARs).

In empirical practice, the well-known biases arising from impulse responses estimated with finite VARs are further aggravated by having to choose relatively short lag lengths due to the parametric loads required in their estimation as Kuersteiner (2005) shows. The solution that we pursue in this paper to avoid these issues, however, is to calculate impulse responses using local projections instead.

Suppose the data are generated by an invertible, reduced-form, infinite moving average process or $VMA(\infty)$—the well-known impulse response representation. Invertibility here means that the space of the vector $\boldsymbol{y}_t$ spans the space of the residual vector, $\boldsymbol{\epsilon}_t$, and that the process can alternatively be expressed as a reduced-form, infinite vector autoregression or $VAR(\infty)$. This assumption allows for very general impulse response trajectories with potentially interesting dynamics at long-horizons. We set aside any discussion on identification since the main issues discussed here do not depend on it. Let

$$\boldsymbol{y}_t = \sum_{h=0}^{\infty} B_h \boldsymbol{\epsilon}_{t-h}; \quad h = 0, 1, \ldots; \quad B_0 = I\,, \tag{30}$$

be the $VMA(\infty)$ representation of the m-dimensional vector $\boldsymbol{y}_t$ (without loss of generality, we omit the constant term). Under the well-known general invertibility assumptions explicitly stated in Appendix A, the $VAR(\infty)$ is:

$$\boldsymbol{y}_t = \sum_{j=1}^{\infty} A_j \boldsymbol{y}_{t-j} + \epsilon_t; \quad j = 1, 2, \ldots\,. \tag{31}$$

The moving average matrices, B_h, and the autoregressive matrices, A_j, follow the well-known recursion due to Durbin (1959):

$$B_h = A_1 B_{h-1} + A_2 B_{h-2} + \ldots + A_k B_{k-h} + \underbrace{A_{k+1} B_{k-h-1} + \ldots + A_{h-1} B_1 + A_h}_{remainder}\,. \tag{32}$$

Lewis and Reinsel (1985) established that, under standard regularity assumptions, a $VAR(p)$ provides consistent estimates of $A_1, \ldots, A_p$ with $p, T \to \infty$ as long as p grows at a rate $p^2/T \to 0$. There are two practical implications of this result. First, a researcher choosing a truncation lag $k < p$ fails to meet the consistency assumption and hence, based on Equation 32, will obtain inconsistent impulse response estimates B_h, even when h is relatively small.

The second and more subtle implication is the following. Suppose that indeed the truncation lag is chosen so that $k = p$ and hence the consistency condition is met. Then, as is clear from Equation 32, estimates of the impulse response for horizons $h = 1, \ldots, k$ will be consistently estimated, but not for horizons $h > k = p$. The reason is that B_h for $h > k = p$ involves the terms $B_1, \ldots, B_{k-h-1}, A_{k+1}, \ldots, A_h$—the remainder term in Equation 32—, which have been truncated and hence their omission introduces inconsistency.

What about local projections? An extension of the proof in Lewis and Reinsel (1985) provided in Appendix A shows that local projections are consistent for any horizon h, even when the lag structure is truncated as long as $p, T \to \infty$ at rate $p^2/T \to 0$. Lusompa (2019) derives a related result in the context of generalized least-squares inference of local projections. Relatedly, Montiel Olea and Plagborg-Møller (2019) use similar asymptotic arguments to show how lag-augmented local projections provide asymptotically valid inference for both stationary and non-stationary data over a wide range of response horizons.

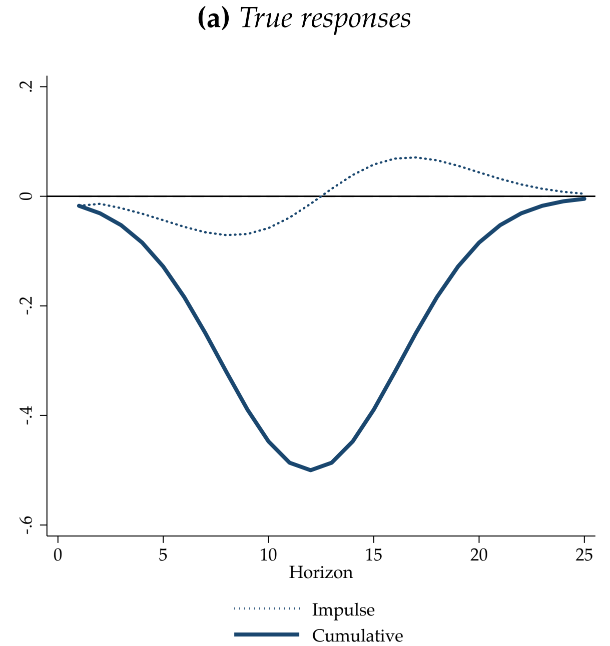
Basically, local projections are direct estimates of the impulse response (moving average) coefficients. Truncating the lag structure, even when $h > k$, has asymptotically vanishing effects on the consistency of the estimator. Truncated VARs on the other hand, have to be inverted to construct the impulse response. Hence the impulse response depends on the entire dynamic specification of the VAR. The cumulation of small sample inconsistencies over increasing horizons can pile up quickly and turn into non-negligible distortions to the impulse response, specially at long horizons.

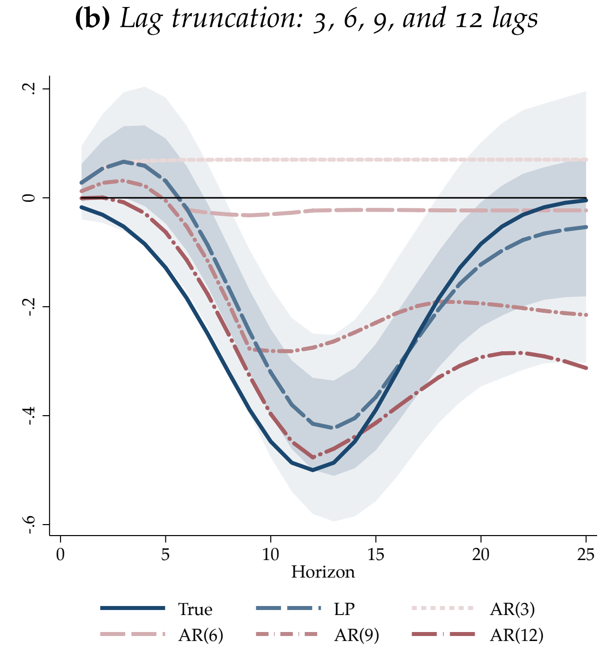
Of course, the solution would be to specify the VAR truncation lag, k, to be large (as long as $k^2/T \to 0$). Setting aside the parametric burden imposed in the estimation, this may not be enough to address the second of the practical issues highlighted earlier, namely the truncation of the remainder term in Equation 32. To illustrate these issues, Figure 1 shows a simple Monte Carlo exercise. We generate an MA process whose coefficients are determined by the impulse response function displayed in panel (a). The implied cumulative response is also shown, as this is the object of interest when we estimate. This impulse response is meant to loosely mimic the shape of the responses we find later in the paper. In cumulative terms, a shock has transitory, but long-lived effects on the variable.[19]

Panel (b) of Figure 1 hence shows Monte Carlo averages from estimates of the cumulative

[19]Further details on the setup of the Monte Carlo exercise along with the specifics of how the two panels of Figure 1 are generated are in Appendix A.

Figure 1: *Estimating cumulative responses: autoregressive versus local projection biases at long horizons.*





Notes: Sample size: 1,000. Monte Carlo replications: 1,000. The shaded error bands are 1 and 2 standard error bands based on the local projection Monte Carlo average. LP refers to cumulative local projections using 2 lags. AR(k) refers to impulse responses cumulated from an autoregressive model with k = 3, 6, 9, and 12 lags. See text.

response from a simple AR model with 3, 6, 9, and 12 lags versus local projections using only 2 lags—a considerable handicap for the local projection. Again, to mimic the empirical analysis, we assume a sample with 1,000 observations (results with 300 observations yield nearly identical results). We repeat the experiment 1,000 times. The error bands displayed are the one and two standard error bands of the local projection Monte Carlo averages.

As is evident from the figure, given the long-lived dynamics of our experiment, truncating the data below 12 lags would generate cumulative effects that are relatively short-lived and far off the true response. The reason is fewer than lags would generally capture the early stages of the impulse response, where not much action has yet taken place, and it would miss entirely the undoing of the dynamics of the first 12 periods that follow over periods 13-24. In contrast, local projections provide quite a close estimate of the response even though the truncation lag is quite severe. As we increase the AR lag length to 12 (the point at which the original negative dynamics die-off as panel (a) illustrates), the AR model with 12 lags picks up the shape of the response very nicely though it gets into trouble once the horizon goes beyond 12 lags, and especially at the tail end, as the theory predicted. In contrast, local projections continue to approximate the response well, even at those long horizons.

Consider our application, which involves 9 variables. A 9-dimensional vector autoregression with 12 lags (as in the Monte Carlo application) involves 108 regressors per equation. The correct lag length, which is 24 in our D.G.P. involves a whopping 216 regressors. Compare that to the 18 regressors for the local projection. Further, note that even truncating the AR at 12 lags is really on the boundary of the order needed to capture the main features of the theoretical impulse response given the D.G.P. Typical information criteria, specially commonly used Bayesian (or Schwartz) information criteria, will tend to select lag lengths that are entirely too small (see Kuersteiner, 2005). Even if long lag lengths are selected, the parametric loads make the task of analyzing the data across subsamples (as we do) even more difficult or often times, impossible.

5. The data show that monetary shocks have long-lived effects

The empirical approach from this point forward relies on local projections, estimated with instrumental variables (LPIV), based on Equation 27 and Equation 34. The instruments, adjusted for capital mobility, are $z_{i,t}^{P}$ and $z_{i,t}^{F}$, as defined earlier, and we estimate the following (cumulative) impulse responses for the baseline, no spillover case ($\Phi = 0$),

$$y_{i,t+h} - y_{i,t-1} = \alpha_{i,h} + \Delta R_{i,t}\, \beta_h + \boldsymbol{x}_{i,t}\, \gamma_h + u_{i,t+h}\,, \tag{33}$$

$$\Delta R_{i,t} = \kappa_i + z_{i,t}^{P}\, \lambda_P + z_{i,t}^{F}\, \lambda_F + \boldsymbol{x}_{i,t}\, \boldsymbol{\zeta} + v_{it}\,. \tag{34}$$

for $h = 0, 1, \ldots, H; i = 1, \ldots, N; t = t_0, \ldots, T$, where $y_{i,t+h}$ is the outcome variable, log real GDP, for country i observed h periods from today, $\alpha_{i,h}$ are country fixed effects at horizon h, $\Delta R_{i,t}$ refers to the instrumented change in the short-term interest rate (usually government bills), our stand-in for the policy rate; and $\boldsymbol{x}_{i,t}$ collects all additional controls including lags of the outcome and interest rates, as well as lagged values of other macro aggregates.[20] Moreover, we control for global business cycle effects through a global world GDP control variable to parsimoniously soak up common global shocks. Estimation is robust with clustering by country.

Table 1 reports the first-stage regression of the pegging country's short term interest

[20] The list of domestic macro-financial controls used include log real GDP; log real consumption per capita; log real investment per capita; log consumer price index; short-term interest rate (usually a 3-month government security); long-term interest rate (usually a 5-year government security); log real house prices; log real stock prices; and the credit to GDP ratio. The variables enter in first differences except for interest rates. Contemporaneous terms (except for the left-hand side variable) and two lags are included. We control for contemporaneous values of other macro-financial variables for two purposes a) base rate movements might be predictable by current home macro-conditions, and b) we wanted to impose restrictions in the spirit of Cholesky ordering whereby real GDP is ordered at the top. Results are robust to excluding contemporaneous home-country controls.

Table 1: *Trilemma instruments: First stage evidence.*

(a) Pegs ($D^P_{i,t} = 1$)	All years	PostWW2
λ_P	0.58***	0.61***
t-statistic	[7.56]	[8.30]
Observations	766	585

(b) Floats ($D^F_{i,t} = 1$)	All years	PostWW2
λ_F	0.26***	0.24***
t-statistic	[3.32]	[3.14]
Observations	379	289

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Full sample: 1870–2015 excluding WW1: 1914–1919 and WW2: 1939–1947. Pre-WW2 sample: 1870–1938 (excluding 1914–1919). Post WW2 sample: 1948–2015. These regressions include country fixed effects as well as up to two lags of the first difference in log real GDP, log real consumption, investment to GDP ratio, credit to GDP, short and long-term government rates, log real house prices, log real stock prices, and CPI inflation. In addition we include world GDP growth to capture global cycles. Estimation is robust with clustering by country. See text.

rate $\Delta R_{i,t}$ on the instruments $z^P_{i,t}, z^F_{i,t}$ and controls $x_{i,t}$, country fixed effects and (robust) clustered standard errors. The interest-rate passthrough is roughly 0.6 for pegs and 0.25 for floats. Thus, neither represents a hard peg or a pure float corner case, further bolstering the case for studying the more general imperfect pass-through case discussed earlier. Both instruments are statistically significant. The peg instrument, $z^P_{i,t}$, has a t-statistic close to 8 in the full and post-WW2 samples and is therefore not a weak instrument. The float instrument, $z^F_{i,t}$, has a t-statistic close to 3 in the full and post-WW2 samples, a weaker instrument, as one would expect though still relevant. Nevertheless, we show that our results are robust to excluding the weaker instrument.

5.1. Main results

The main findings in our paper are shown by the response of real GDP to a shock to domestic interest rates. Before we show the main results, we highlight the value of our instrumental variable by comparing the response calculated using selection-on-observables identification versus identification with our trilemma instrument. This is shown in Table 2.

The table reports coefficient estimates of the (cumulative) impulse response calculated with each identification approach for the full and post-WW2 samples, *LP-OLS* and *LP-IV* using the trilemma instruments. We provide the coefficient estimates by row, with a test of the null hypothesis that *LP-OLS* and *LP-IV* estimates are equal. The differences between identification schemes could not be starker: *LP-IV* estimates are economically and statistically significant, and the *LP-IV* response is considerably larger at all horizons.

We display these results graphically in Figure 2. This figure is organized into two

Table 2: *LP-OLS vs. LP-IV. Attenuation bias of real GDP responses to 100 bps shock. Trilemma instruments.*

Responses of real GDP at years 0 to 12 (100 $\times$ log change from year 0 baseline).

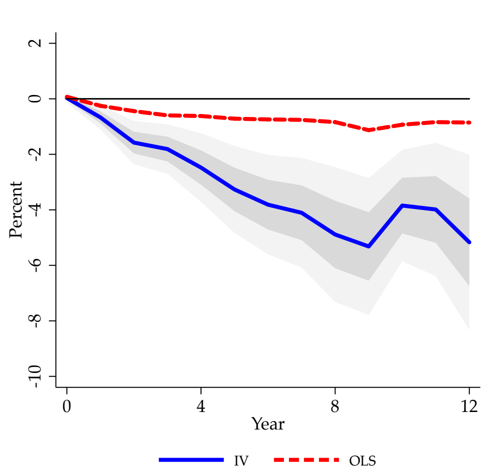
	(a) Full Sample		OLS=IV	(b) Post-WW2		OLS=IV
Year	LP-OLS	LP-IV	p-value	LP-OLS	LP-IV	p-value
	(1)	(2)	(3)	(4)	(5)	(6)
$h = 0$	0.09***	-0.01	0.29	0.05**	0.07	0.81
	(0.03)	(0.09)		(0.03)	(0.07)	
$h = 2$	-0.36***	-1.56***	0.00	-0.29**	-1.09***	0.01
	(0.12)	(0.40)		(0.11)	(0.31)	
$h = 4$	-0.45**	-2.36***	0.00	-0.33*	-1.32***	0.00
	(0.18)	(0.63)		(0.17)	(0.33)	
$h = 6$	-0.54**	-3.54***	0.00	-0.39*	-2.64***	0.00
	(0.27)	(0.90)		(0.24)	(0.58)	
$h = 8$	-0.62**	-4.55***	0.00	-0.42	-2.99***	0.00
	(0.30)	(1.13)		(0.33)	(0.81)	
$h = 10$	-0.72*	-3.55***	0.00	-0.26	-2.78***	0.00
	(0.38)	(0.94)		(0.42)	(0.83)	
$h = 12$	-0.64	-4.68***	0.00	-0.22	-2.98***	0.00
	(0.44)	(1.41)		(0.51)	(0.95)	
KP weak IV		34.77			33.81	
H_0: $LATE = 0$		0.00			0.00	
Observations	1145	1145		874	874	

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Cluster robust standard errors in parentheses. LP-IV (OLS): Response to a 100 bps shock in domestic short-term interest rate instrumented with the trilemma IVs (or using OLS, no instruments). Full sample: 1890–2015 excluding WW1: 1914–1919 and WW2: 1939–1947. Post WW2 sample: 1948–2015. *KP weak IV* refers to the Kleibergen-Paap test for weak instruments. H_0: $LATE = 0$ refers to the p-value of the test of the null hypothesis that the coefficients for $h = 0, ..., 10$ are jointly zero for a given subpopulation. $OLS = IV$ shows the p-value for the test of the null that OLS estimates equal IV estimates. Estimation is robust with clustering by country. See text.

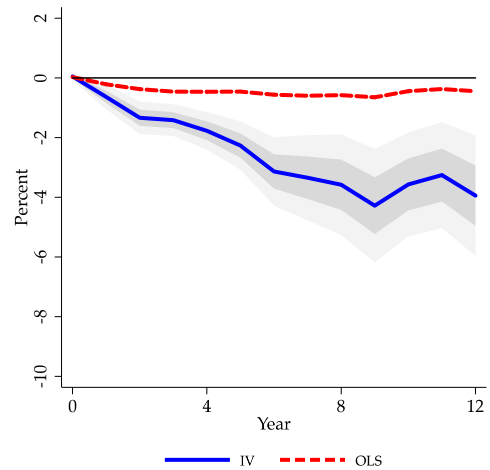
columns, charts (a) and (c) refer to full sample results, and columns (b) and (d) to the post-WW2 sample. In addition, the top row—charts (a) and (b)—is based on using the *peg* and *float* instruments, whereas the second row—charts (c) and (d)—only use the *peg* instrument as a robustness check. Regardless of the sample used, a 1 percentage point increase in domestic short-term interest rates has sizable and long-lasting effects on GDP. In the full sample, GDP declines by 4.68 percent over 12 years. A similar, but moderated, effect is found when we restrict the sample to post-WW2. The drop 12 years after impact is 2.98 percent. Both estimates are significantly different from zero ($p < 0.01$).

Figure 2: *Baseline response to 100 bps shock: Real GDP.*

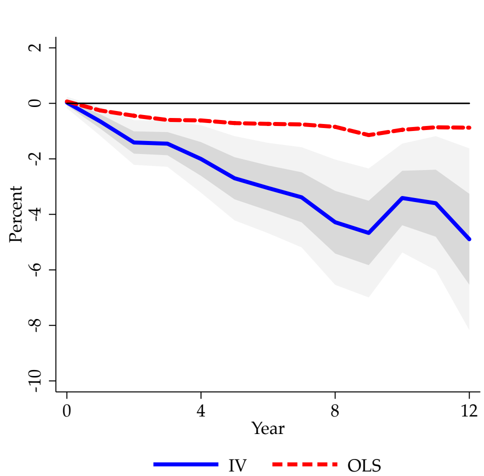
(a) *Full sample: 1890–2015.*



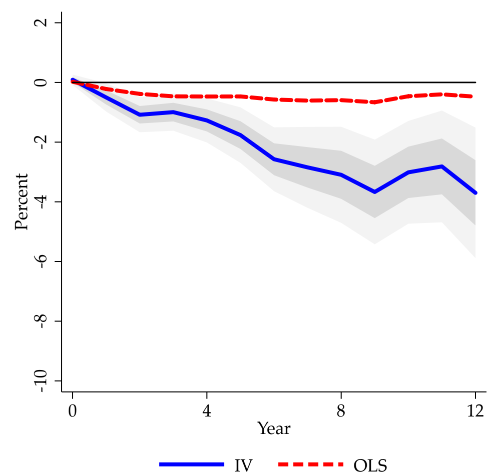
(b) *Post-WW2 sample: 1948–2015.*



(c) *Full sample: 1890–2015, using only the peg IV.*



(d) *Post-WW2 sample: 1948–2015, using only the peg IV.*



Notes: Response to a 100 bps shock in domestic short-term interest rate instrumented with the trilemma IVs. Full sample: 1890–2015 (World Wars excluded). LP-OLS estimates displayed as a dashed red line, LP-IV estimates displayed as a solid blue line with 1 S.E. and 2 S.E. confidence bands. Top row uses both peg and float instruments; bottom row uses only peg instrument. Estimation is robust with clustering by country. See text.

5.2. Inspecting the mechanism

The results in Figure 2 are a far cry from traditional notions of long-run neutrality found in the literature. What is the source of this persistent decline? We employ a Solow decomposition of GDP (Y) into its components, using a Cobb-Douglas production function, to construct hours worked (L, employees times number of hours per employee); capital stock (K, measured capital in machines and buildings); and the Solow residual, labeled as total factor productivity (TFP).

Figure 3 displays the (cumulative) responses of each of these components to the same shock to the domestic short-term interest rate using the trilemma instruments, both for the full and the post-WW2 samples.[21] The chart displays each of the components with one and two standard error confidence bands. In the appendix Section I.2, we provide corresponding figure for post-WW2 sample with similar findings.

Several features deserve mention. Figure 3a shows that there are similar declines in capital and raw TFP. In terms of growth accounting and the negative real GDP response, the capital response component accounts for two-thirds and the TFP response component for about one-third. However, total hours worked exhibits a much flatter response, with no sign of labor hysteresis. Because capital enters the production function with a smaller weight, it should be clear from the figure that most of the decline in GDP is explained by the TFP variable, followed by capital , with total hours worked mostly flat.

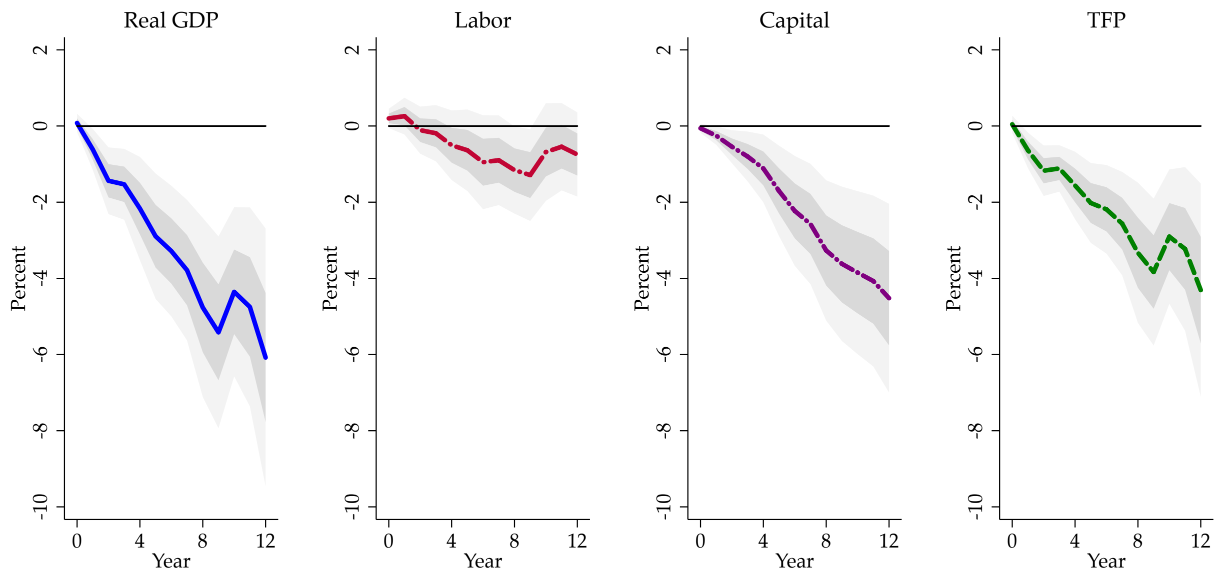
We may note that capital accumulation follows textbook dynamics in the short-run. The capital response is initially muted but builds up over time. But unlike a textbook New Keynesian model (Galí, 2015b), the capital stock does not recover even after 12 years. Similarly, TFP falls gradually rather than suddenly, and also does not recover.

But are these estimates based on the raw data accurate? One serious concern with Solow decompositions, well known at least since the work of Basu and Kimball (1997), is the issue of capacity utilization biases (See also Basu, Fernald, and Kimball 2006). When K measures the capital stock, as here, that is not capital input: input is only the capital being used, possibly much lower in periods of slack when plant and equipment may be idling. Likewise if L measures labor stock, but even if it measures total hours this may be biased upwards in periods of slack if labor is hoarded, and not fully utilized. In such cases, naïve use of the Solow approach will result in mismeasured factor inputs with too weak cyclicality, leaving residual TFP with too strong cyclicality, a pervasive problem that exaggerates the role of TFP shocks as a source of business cycles. Therefore, following the literature, we revise the capital and labor raw data to account for cyclicality in utilization, following the well-established method proposed by Imbs (1999). The results are shown in Figure 3b, and reveal some subtle differences. Overall the responses are similar in terms of shape and statistical significance, so the qualitative story is the same. But quantitatively, the TFP response is now muted in amplitude, as expected, and the factor responses are accordingly larger, suggesting that the Imbs correction captures some utilization-driven

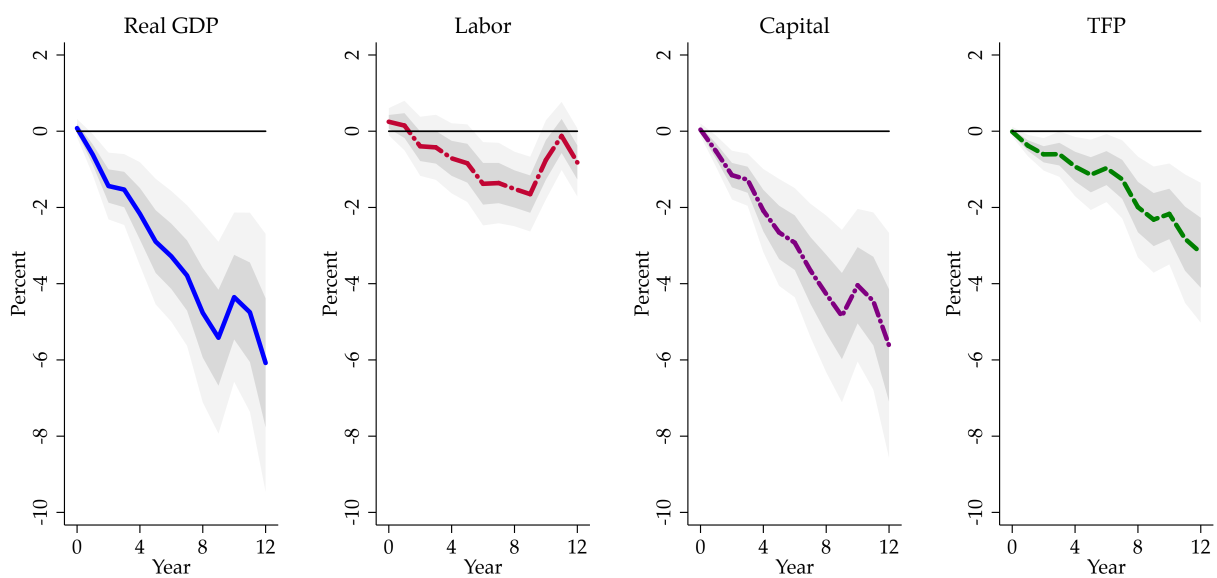
[21]For comparability, we use the same controls in estimating responses of various components of the production function as those used for the GDP response.

Figure 3: *Baseline response to 100 bps shock: Real GDP and Solow decomposition. Full sample, 1890–2015.*

(a) *Estimates using raw data*

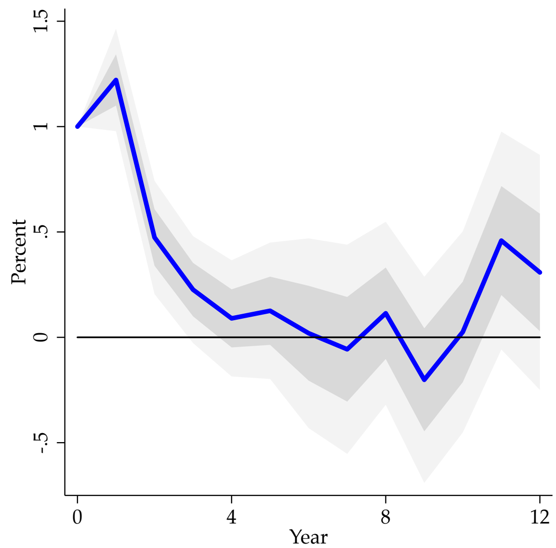


(b) *Estimates using Imbs correction for factor utilization*



Notes: Response to a 100 bps shock in domestic short-term interest rate instrumented with the trilemma IVs. Full sample: 1890–2015 (World Wars excluded). LP-IV estimates displayed as a thick lines and and 1 S.E. and 2 S.E. confidence bands. The upper panel uses raw data on capital stocks and total hours to construct TFP as a residual. The lower panel adjusts the raw data on capital stock and total hours to obtain estimates of actual factor inputs by using the Imbs (1999) correction. Estimation is robust with clustering by country. See text.

Figure 4: *Baseline short term nominal interest rate own response to 100 bps shock. Full sample, 1890–2015.*



Notes: Response to a 100 bps shock in domestic short-term interest rate instrumented with the trilemma IVs. Full sample: 1890–2015 (World Wars excluded). LP-IV estimates displayed as a thick line and and 1 S.E. and 2 S.E. confidence bands. Estimation is robust with clustering by country. See text.

factor slack. We proceed using these Imbs-corrected responses as our baseline henceforth, since they give a smaller (i.e., conservative) response of TFP and we wish to guard against exaggerating our proposed hysteresis channel which runs via TFP.

Finally, we verify that the long-run responses we have found and not a simple mechanical result of unusually persistent shock impulses. One possible explanation for the long-lasting effects of the monetary shock could be that domestic interest rates remain elevated for a long period of time as well. In other words, persistence is generated by a delayed response in interest rates. A simple check shows that this is not the case. Figure 4 shows that the short-term nominal interest rate returns to zero deviation after about 4 years. The response of the nominal interest rate is typical of what has been reported often in the literature (see, e.g., Christiano, Eichenbaum, and Evans, 1999; Ramey, 2016).

5.3. Robustness and discussion

Our baseline LP was quite saturated, and included lags and current values of global GDP growth. This rich specification served multiple purposes. Global shocks that caused bases to change interest rates are controlled for during instrument construction, as well through use of these controls during estimation of the local projections. Comparison with OLS estimates, which control for contemporaneous home economy macro-variables (as in a

Cholesky ordering), further allay some concerns on systematic structural breaks in GDP or TFP growth picked up as regime shifts over decades.

We now discuss further robustness checks to ensure that the persistent effects we identified are not misattributed to monetary policy shocks.

Spillover correction for the trilemma instrument A violation of the exclusion restriction could occur if base rates affect home outcomes through channels other than movements in home rates, as underscored in the theoretical model. Additional influences via such channels are sometimes referred to as *spillover effects*. These could occur if base rates proxy for factors common to all countries. That said, these factors would have to persist despite having included global GDP to soak up such variation. Or they could occur for other reasons, such as spillovers via trade. In addition to the control strategy used in our baseline specification, we now assess such spillover effects more formally by estimating a spillover-corrected IV specification developed in Section 2.4.

Equations 27 and 34 generalize our baseline IV estimator to accommodate spillovers that vary with size of export share in the peg economies. With a range of values for $\Phi \in [0.1, 0.3]$, we estimate the cumulative impulse responses to GDP. Figure 5 shows our spillover-corrected estimates of response of output to a 100 bps monetary policy shock. A light-green shaded area with dashed border shows the spillover corrections. While the impulse response coefficients at year 12 are somewhat smaller than the baseline estimates (solid blue line), monetary policy shocks still exert a sizable persistent effect on output.

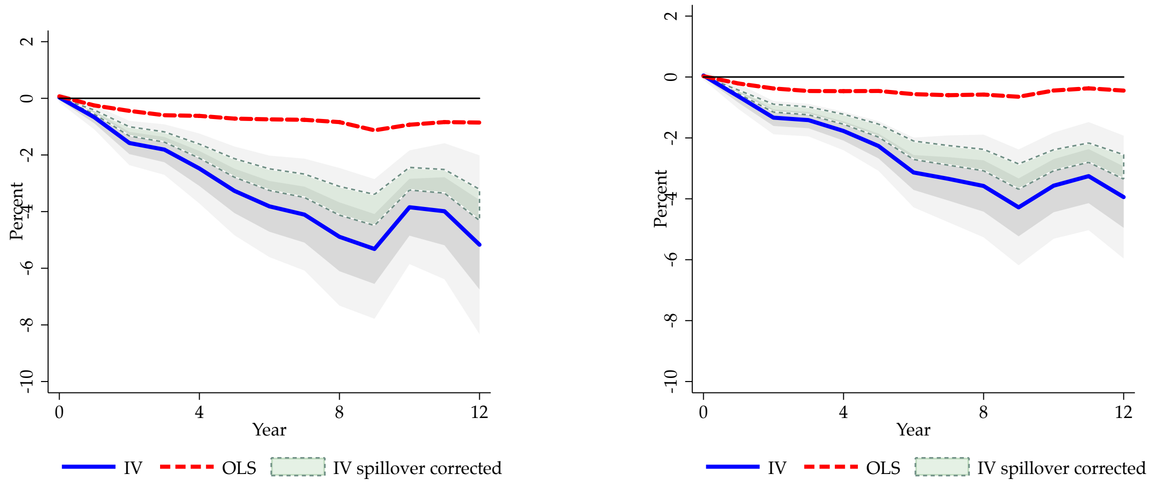
External factors and structural breaks A cruder approach to validate the exclusion restriction is by directly controlling for a primary channel through which the spillover effects may originate. A monetary tightening in the base country may reduce the demand for goods from the pegging economy. This effect would amplify the effect of the trilemma shock on home output. With soft peg regimes, there may be further effects through changes in nominal exchange rates (Gourinchas, 2018). To account for these effects, we control for global GDP growth rate, base country's GDP growth rate, exchange rate of the pegging economy with respect to the USD and the current account of the peg. Since we do not have exchange rate data with respect to other countries, we indirectly control for those spillovers using the current account of the peg country.

Figure 6a plots the IRFs to the trilemma identified shock. Directly controlling for open-economy variables, motivated by export demand channels, does not affect our main result: monetary shocks still have a large and very persistent effect on real GDP.

Fernald (2014) and Gordon (2016) have convincingly argued that there are structural breaks in U.S. TFP growth. One may suspect that there are structural breaks in other

Figure 5: *Response to 100 bps trilemma shock with spillover corrections: Real GDP.*

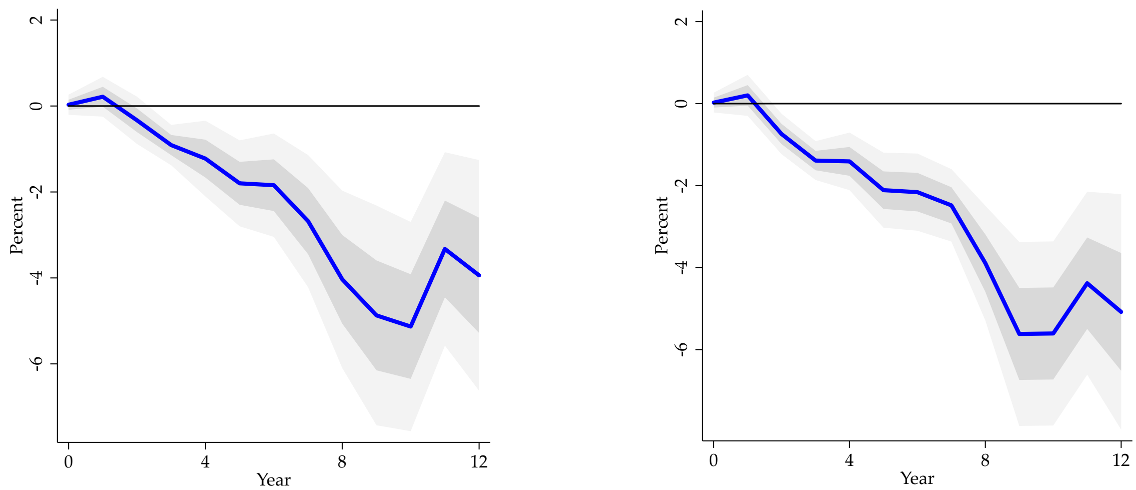
(a) *Full sample: 1890–2015.* **(b)** *Post-WW2 sample: 1948–2015.*



Notes: Response to a 100 bps shock in domestic short-term interest rate instrumented with the trilemma IVs. Full sample: 1890–2015 (World Wars excluded). Post-WW2 sample: 1948–2015. LP-OLS estimates displayed as a dashed red line, LP-IV estimates displayed as a solid blue line with 1 S.E. and 2 S.E. confidence bands, LP-IV spillover corrected estimates displayed as a light green shaded area with dashed border, using $\Phi \in [0.1, 0.3]$. Estimation is robust with clustering by country. See text.

Figure 6: *Response to 100 bps trilemma shock with additional controls: Real GDP. Full sample, 1890–2015.*

(a) *Open economy model based controls* **(b)** *Structural breaks in TFP growth*



Notes: Response to a 100 bps shock in domestic short-term interest rate instrumented with the trilemma IVs. Full sample: 1890–2015 (World Wars excluded). LP-IV estimates displayed as a solid blue line with 1 S.E. and 2 S.E. confidence bands. Estimation is robust with clustering by country. See text.

economies' TFP growth rates as well. If such structural breaks coincide in time with monetary shocks of the same sign, they could bias our results. To address this concern, we first estimate five structural breaks in TFP growth and GDP growth for each country in our sample using the UD-max statistic of Bai and Perron (1998). We report these estimated structural break dates in Appendix I.6. Then in our baseline specification, we allow output growth to lie in either of the five regimes at horizon zero.

Figure 6b plots the estimated impulse response when including structural breaks in TFP growth. As evident, our results are robust to accounting for structural breaks.[22] We conducted a variety of additional robustness checks reported in the Appendix, all of which made no meaningful difference to the main results reported in this section.

6. A small-open economy model with long-run effects

Impulse responses calculated with standard methods that internally favor reversion to the mean will tend to underestimate the value of the response at longer horizons. By relying on local projections, we allow the data to more directly speak as to its long-run properties. The evidence presented in the previous sections strongly indicate that these long-run effects are important and require further investigation. In order to think through a possible mechanism that explains our empirical findings, we augment our baseline model with capital and endogenous productivity growth in a stylized manner.

We assume that physical capital is also used for production in the non-traded sector: $Y_{Nt}(i) = K_{Nt}^{\alpha}(i)\left(A_t L_{Nt}(i)\right)^{1-\alpha}$. As before, traded output is an endowment that grows at rate of growth of A_t for stationarity along a balanced growth path. Non-traded output is used for consumption and investment. As is standard in medium scale DSGE models (Christiano, Eichenbaum, and Evans, 2005; Smets and Wouters, 2007; Justiniano, Primiceri, and Tambalotti, 2013), we assume investment adjustment costs. We leave the formal model to the appendix, and focus on the key departure that allows us to introduce long-run effects on output and capital with a parametrically convenient process.

6.1. Long-run effects

In order to be able to capture the empirical features describe in the previous sections, we examine a richer specification of the low of motion for total factor productivity A_t than is conventional. In particular, we assume that the law of motion for A_t is:

$$\log A_t = \log A_{t-1} + \mu_t + \eta \log\left(Y_{Nt-1}/Y_{Nt-1}^{f,t-1}\right), \tag{35}$$

[22]In the appendix I.5, we report the IRFs allowing for structural breaks in GDP growth.

where μ_t is the exogenous component of the TFP growth rate, that may be subject to trend shocks. Y_{Nt} is non-traded output at time t. $Y_{Nt-1}^{f,t-1}$ is the flexible price level of output in period $t-1$ conditional on A_{t-1}, and will be referred to as the *potential* output at time $t-1$. The second component denotes the endogenous component of TFP growth, where η is the elasticity of TFP growth rate with respect to fluctuations in output due to nominal rigidities. We refer to this parameter as the *hysteresis elasticity* (so as to be consistent with DeLong and Summers 2012).

The above law of motion allows business cycles to affect TFP growth rate only in the presence of nominal rigidities or inadequate stabilization. For clarity, we employ this parametric-convenient functional form for hysteresis. A similar setup was used by Stadler (1990) in his seminal work. A micro-founded model of innovation and productivity growth that yields this exact representation under monetary policy shocks can be found in the recent literature embedding endogenous growth into DSGE models (Bianchi, Kung, and Morales, 2019; Garga and Singh, Forthcoming). The effects of business cycles on TFP growth rate that are unrelated to nominal rigidities can be denoted by time varying values of μ_t, which may depend on other shocks (markup shocks, stationary TFP shocks, discount factor shocks, capital quality shocks etc.). For ease of exposition, we only focus on the hysteresis effects induced by the presence of nominal rigidities and treat μ_t as an exogenous process.

6.2. Government

The central bank follows a Taylor rule in setting the nominal interest rate i_t. It responds to deviations in inflation, output and output growth rate from time-t natural allocations.

$$\frac{1+i_t}{1+i_{ss}} = \left(\frac{1+i_{t-1}}{1+i_{ss}}\right)^{\rho_R} \left[\left(\frac{\pi_t}{\pi_{ss}}\right)^{\phi_\pi} \left(\frac{Y_t}{Y_t^{f,t}}\right)^{\phi_y}\right]^{1-\rho_R} \epsilon_t^{mp}, \tag{36}$$

where i_{ss} is the steady state nominal interest rate, π_t is gross inflation rate, π_{ss} is the steady state inflation target, $Y_t^{f,t}$ is the time-t natural output, ρ_R determines interest-rate smoothing and $\epsilon_t^{mp} \sim N(0, \sigma_r)$ is the monetary policy shock.[23]

We assume government balances budget every period, where total lump-sum taxes go as a production subsidy to intermediate good producers, and a wage subsidy to workers.

[23]In the presence of endogenous state variables, the timing assumptions on flexible prices matter in defining appropriate natural rate. We follow Woodford (2003, Sec 5.4) in defining natural rate as the level of variables at time t such that prices are set flexibly beginning at time t taking as given the evolution of state-variables up to time t. See also Garga and Singh (Forthcoming) for a detailed discussion.

Table 3: *Point estimates for hysteresis elasticity η*

	1890–2015	1948–2015
η	0.25	0.67
[95% confidence interval]	[0.21,0.30]	[0.34,0.99]

Notes: The point estimates are estimated with a two-step classical minimum distance approach using the IRF of utilization-adjusted TFP and the IRF of cyclical gaps to the monetary policy shock in the second step. See text.

6.3. Calibration and simulation

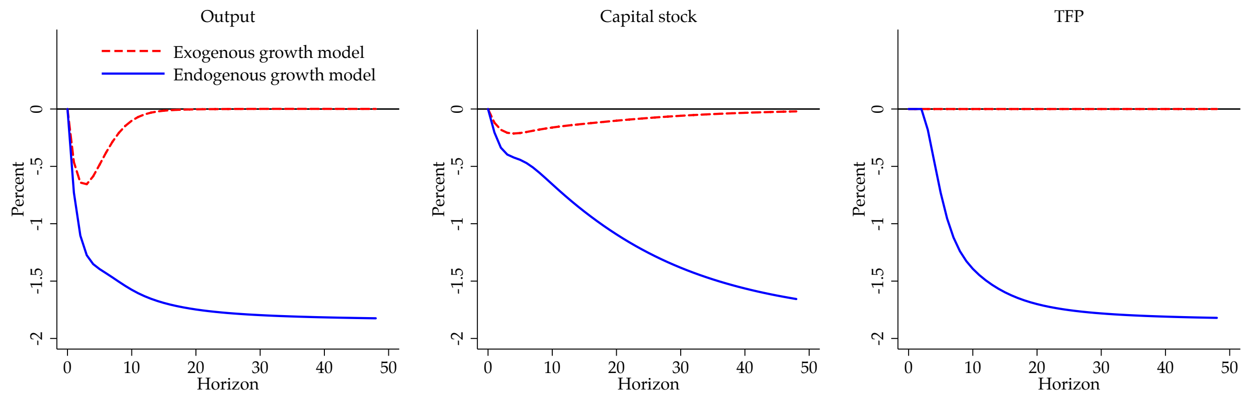
As the extended model is intentionally stylized, we take parameters from papers in the literature (Justiniano et al., 2013; Schmitt-Grohé and Uribe, 2016). We report these in the appendix in Table A1. The new parameter in our model, relative to the business cycles literature, is η: the hysteresis elasticity. Table 3 reports the point estimates for η implies by the estimated impulse responses. We use a two-step classical minimum distance approach to recover η. In the first step, we estimate the IRFs of utilization-adjusted TFP and a measure of cyclical gap to monetary policy shock. Using the estimated coefficients (see Figure 3), we then estimate η using Equation 35.[24] Following the persistent drop in output after the Great Recession in the US, DeLong and Summers (2012) infer that this parameter could be as high as 0.24. While our estimate is on the higher side, there is considerably large confidence interval. In our calibration henceforth, we use the value of 0.25.

Figure 7 plots the model-implied impulse responses for output, capital stock, and utilization-adjusted TFP after a monetary policy shock. Solid blue line reports the IRFs for endogenous growth model with $\eta = 0.25$, and dashed blue line reports IRFs for the comparable exogenous growth benchmark i.e. $\eta = 0$. The IRFs for output, capital stock and TFP are plotted in percent deviations from an exogenous trend. Time is in quarters.

The model replicates the estimated empirical patterns. There is a persistent decline in capital stock, output and TFP. A monetary shock reduces aggregate output and investment in exogenous growth model, and they return back to their long-run levels. The effects of monetary shock in an exogenous growth model do last for nearly six years, as can be seen by slow recovery of capital stock to the pre-shock level. On the other hand, capital stock and

[24] Formally, using the concept of time-t natural output (taking evolution of state variables up to time t as given), the cyclical output gap in the model with utilization adjustments is $\log Y_t - \log Y_t^{f,t} = \alpha u_t + (1-\alpha)e_t + (1-\alpha)\log L_t$, where u_t is capital utilization rate, e_t is labor-utilization rate, and L_t is labor input. Now, let λ_h be the estimated IRF of utilization-adjusted TFP (A_t) at horizon h, and let β_h be the estimated IRF of the cyclical gap. We empirically estimate λ_h and β_h at $h = 0, ..., 12$ as in Figure 3b. Using Equation 35, we can recover the estimating equation that delivers estimates for η as $\lambda_h = \eta \sum_{k=0}^{h} \beta_k$.

Figure 7: *Model response of output, capital stock, and TFP to a 100 bps increase in nominal interest rate*



Notes: The figure plots the model-implied IRFs for output, capital stock and utilization-adjusted TFP to a transitory shock to the assumed Taylor rule. Solid blue line reports the IRFs for endogenous growth model with $\eta = 0.25$, and dashed red line reports IRFs for the comparable exogenous growth benchmark i.e. $\eta = 0$. Time is in quarters. IRFs are traced following a one-time exogenous shock in the federal funds rate of 100 basis points. The IRFs are plotted in percent deviations from an exogenous trend.

output do not return to pre-shock trend even after twelve years when TFP is endogenous. Furthermore, the endogenous growth model exhibits considerable amplification to the transitory shock because of the large hysteresis elasticity.

7. Conclusion

This paper challenges the widely accepted view that money is neutral in the long-run. We find that monetary policy has real effects that last for a decade or more. We spent considerable time and energy with the three pillars of our empirical strategy—identification, data, and methods—to ensure the reader that our results are solid. We believe they are.

The source of the main hysteresis result—that monetary policy shocks have long-lasting effects on output—surprised us. We find that capital and TFP growth are the main drivers of this result, but not hours worked, in contrast to standard models of labor hysteresis. Our findings do not negate the influence of labor frictions in shaping the business cycle. Instead, we do not find a strong role for such labor scarring in explaining why monetary policy has such long-lived effects.

Our next task was to provide an economic framework to understand where our results come from. We do this using the same open economy framework that we used to justify our identification of monetary policy shocks. Simple extensions to the existing paradigm replicate the empirical findings that we document here.

There is much that is left unexplored in this paper. Determining the micro-foundations that explain TFP growth hysteresis would require a different paper devoted to the topic with a completely different data set. Exploring the optimality of the monetary policy rule in more general settings, and the welfare consequences of the hysteresis results documented here are of first order importance for policymakers. Perhaps more importantly, our paper challenges long-held views that require a reexamination of standard business cycle models.

References

Anzoategui, Diego, Diego Comin, Mark Gertler, and Joseba Martinez. 2019. Endogenous Technology Adoption and R&D as Sources of Business Cycle Persistence. *American Economic Journal: Macroeconomics* 11(3): 67–110.

Bai, Jushan, and Pierre Perron. 1998. Estimating and Testing Linear Models with Multiple Structural Changes. *Econometrica* 66(1): 47–78.

Baqaee, David Rezza, and Emmanuel Farhi. 2019. Productivity and Misallocation in General Equilibrium. *Quarterly Journal of Economics* 135(1): 105–163.

Barlevy, Gadi. 2004. The cost of business cycles under endogenous growth. *American Economic Review* 94(4): 964–990.

Barro, Robert J. 2013. Inflation and Economic Growth. *Annals of Economics and Finance* 14(1): 85–109.

Basu, Susanto, and John G Fernald. 2002. Aggregate productivity and aggregate technology. *European Economic Review* 46(6): 963–991.

Basu, Susanto, John G. Fernald, and Miles S. Kimball. 2006. Are Technology Improvements Contractionary? *American Economic Review* 96(5): 1418–1448.

Basu, Susanto, and Miles S Kimball. 1997. Cyclical productivity with unobserved input variation. NBER Working Paper 5915.

Benigno, Gianluca, and Luca Fornaro. 2018. Stagnation Traps. *The Review of Economic Studies* 85(3): 1425–1470.

Benigno, Gianluca, Luca Fornaro, and Martin Wolf. 2020. The global financial resource curse. *FRB of New York Staff Report* (915).

Bergeaud, Antonin, Gilbert Cette, and Rémy Lecat. 2016. Productivity Trends in Advanced Countries between 1890 and 2012. *Review of Income and Wealth* 62(3): 420–444.

Berk, Kenneth N. 1974. Consistent autoregressive spectral estimates. *Annals of Statistics* 2(3): 489–502.

Bernanke, Ben S., and Ilian Mihov. 1998. The liquidity effect and long-run neutrality. *Carnegie-Rochester Conference Series on Public Policy* 49(1): 149–194.

Bianchi, Francesco, Howard Kung, and Gonzalo Morales. 2019. Growth, slowdowns, and recoveries. *Journal of Monetary Economics* 101(C): 47–63.

Blanchard, Olivier. 2018. Should We Reject the Natural Rate Hypothesis? *Journal of Economic Perspectives* 32(1): 97–120.

Blanchard, Olivier J., and Lawrence H. Summers. 1986. Hysteresis and the European unemployment problem. In *NBER Macroeconomics Annual 1986*, volume 1, 15–90. Cambridge, Mass.: MIT Press.

Blanchard, Olivier Jean, and Charles M. Kahn. 1980. The Solution of Linear Difference Models under Rational Expectations. *Econometrica* 48(5): 1305–1311.

Calvo, Guillermo A. 1983. Staggered prices in a utility-maximizing framework. *Journal of Monetary Economics* 12(3): 383–398.

Cerra, Valerie, Antonio Fatás, and Sweta Saxena. 2020. Hysteresis and Business Cycles. CEPR Discussion Paper 14531.

Christiano, Lawrence. 2015. Applications of DSGE Model: What Happens if we Switch from Fixed to Flexible Interest Rate Regime? Unpublished. Northwestern University.

Christiano, Lawrence J., Martin Eichenbaum, and Charles L. Evans. 1999. Monetary policy shocks: What have we learned and to what end? In *Handbook of Macroeconomics*, edited by Taylor, J. B., and M. Woodford, volume 1 of *Handbook of Macroeconomics*, chapter 2, 65–148. Amsterdam: Elsevier.

Christiano, Lawrence J., Martin Eichenbaum, and Charles L. Evans. 2005. Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy. *Journal of Political Economy* 113(1): 1–45.

DeLong, Bradford J., and Lawrence H. Summers. 2012. Fiscal Policy in a Depressed Economy. *Brookings Papers on Economic Activity* 1: 233–297.

Durbin, J. 1959. Efficient Estimation of Parameters in Moving-Average Models. *Biometrika* 46(3/4): 306–316.

Eggertsson, Gauti B., and Michael Woodford. 2003. The Zero Bound on Interest Rates and Optimal Monetary Policy. *Brookings Papers on Economic Activity* 2003(1): 139–211.

Farhi, Emmanuel, and Ivan Werning. 2012. Dealing with the Trilemma: Optimal Capital Controls with Fixed Exchange Rates. NBER Working Paper 18199.

Farhi, Emmanuel, and Iván Werning. 2017. Fiscal Unions. *American Economic Review* 107(12): 3788–3834.

Fatás, Antonio. 2000. Do business cycles cast long shadows? Short-run persistence and economic growth. *Journal of Economic Growth* 5(2): 147–162.

Fernald, John G. 2014. A quarterly, utilization-adjusted series on total factor productivity. Federal Reserve Bank of San Francisco Working Paper 2012-19.

Fornaro, Luca. 2015. Financial crises and exchange rate policy. *Journal of International Economics* 95(2): 202–215.

Fornaro, Luca, and Federica Romei. 2019. The Paradox of Global Thrift. *American Economic Review* 109(11): 3745–79.

Galí, Jordi. 2015a. Hysteresis and the European Unemployment Problem Revisited. NBER Working Paper 21430.

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

Galí, Jordi, and Tommaso Monacelli. 2016. Understanding the Gains from Wage Flexibility: The Exchange Rate Connection. *American Economic Review* 106(12): 3829–68.

Garga, Vaishali, and Sanjay R. Singh. Forthcoming. Output hysteresis and optimal monetary policy. *Journal of Monetary Economics* .

Gopinath, Gita, Şebnem Kalemli-Özcan, Loukas Karabarbounis, and Carolina Villegas-Sanchez. 2017. Capital Allocation and Productivity in South Europe. *Quarterly Journal of Economics* 132(4): 1915–1967.

Gordon, Robert J. 2016. *The Rise and Fall of American Growth: The U.S. Standard of Living since the Civil War*. Princeton, N.J.: Princeton University Press.

Gourinchas, Pierre-Olivier. 2018. Monetary Policy Transmission in Emerging Markets: An Application to Chile. In *Monetary Policy and Global Spillovers: Mechanisms, Effects and Policy Measures*, edited by Mendoza, Enrique G., Ernesto Pastén, and Diego Saravia, volume 25 of *Central Banking, Analysis, and Economic Policies Book Series*, chapter 8, 279–324. Santiago: Central Bank of Chile.

Guerrieri, Luca, and Matteo Iacoviello. 2015. OccBin: A toolkit for solving dynamic models with occasionally binding constraints easily. *Journal of Monetary Economics* 70(C): 22–38.

Hume, David. 1752. 'Of Money', 'Of Interest' and 'Of the Balance of Trade'. In *Essays Moral, Political, and Literary*, edited by Miller, Eugene F. Indianapolis, Ind.: Liberty Fund.

Ilzetzki, Ethan, Carmen M. Reinhart, and Kenneth S. Rogoff. 2019. Exchange Arrangements Entering the Twenty-First Century: Which Anchor will Hold? *Quarterly Journal of Economics* 134(2): 599–646.

Imbs, Jean M. 1999. Technology, growth and the business cycle. *Journal of Monetary Economics* 44(1): 65–80.

Jordà, Òscar. 2005. Estimation and Inference of Impulse Responses by Local Projections. *American Economic Review* 95(1): 161–182.

Jordà, Òscar, Moritz Schularick, and Alan M. Taylor. 2016. The great mortgaging: housing finance, crises and business cycles. *Economic Policy* 31(85): 107–152.

Jordà, Òscar, Moritz Schularick, and Alan M. Taylor. 2017. Macrofinancial History and the New Business Cycle Facts. *NBER Macroeconomics Annual* 31(1): 213–263.

Jordà, Òscar, Moritz Schularick, and Alan M. Taylor. 2020. The effects of quasi-random monetary experiments. *Journal of Monetary Economics* 112: 22–40.

Justiniano, Alejandro, Giorgio E. Primiceri, and Andrea Tambalotti. 2013. Is There a Trade-Off between Inflation and Output Stabilization? *American Economic Journal: Macroeconomics* 5(2): 1–31.

Kilian, Lutz. 1998. Small-sample Confidence Intervals for Impulse Response Functions. *Review of Economics and Statistics* 80(2): 218–230.

King, Robert G., and Mark W. Watson. 1997. Testing Long-Run Neutrality. *Federal Reserve Bank of Richmond Economic Quarterly* 83(3): 69–101.

Kuersteiner, Guido M. 2005. Automatic Inference For Infinite Order Vector Autoregressions. *Econometric Theory* 21(1): 85–115.

Laséen, Stefan, and Lars E. O. Svensson. 2011. Anticipated Alternative Policy Rate Paths in Policy Simulations. *International Journal of Central Banking* 7(3): 1–35.

Lewis, Richard, and Gregory C. Reinsel. 1985. Prediction of multivariate time series by autoregressive model fitting. *Journal of Multivariate Analysis* 16(3): 393–411.

Lusompa, Amaze B. 2019. Local projections, autocorrelation, and efficiency. Unpublished. UC Irvine.

Mankiw, N. Gregory. 2001. The inexorable and mysterious tradeoff between inflation and unemployment. *Economic Journal* 111(471): 45–61.

Meier, Matthias, and Timo Reinelt. 2020. Monetary Policy, Markup Dispersion, and Aggregate TFP. University of Mannheim. Unpublished.

Montiel Olea, José Luis, and Mikkel Plagborg-Møller. 2019. Simultaneous confidence bands: Theory, implementation, and an application to SVARs. *Journal of Applied Econometrics* 34(1): 1–17.

Moran, Patrick, and Albert Queralto. 2018. Innovation, productivity, and monetary policy. *Journal of Monetary Economics* 93: 24–41.

Nakamura, Emi, and Jón Steinsson. 2018. Identification in Macroeconomics. *Journal of Economic Perspectives* 32(3): 59–86.

Obstfeld, Maurice, and Kenneth Rogoff. 1995a. Exchange rate dynamics redux. *Journal of Political Economy* 103(3): 624–660.

Obstfeld, Maurice, and Kenneth Rogoff. 1995b. The Mirage of Fixed Exchange Rates. *Journal of Economic Perspectives* 9(4): 73–96.

Obstfeld, Maurice, Jay C. Shambaugh, and Alan M. Taylor. 2004. Monetary sovereignty, exchange rates, and capital controls: The trilemma in the interwar period. *IMF Staff Papers* 51(3S1): 75–108.

Obstfeld, Maurice, Jay C. Shambaugh, and Alan M. Taylor. 2005. The Trilemma in History: Tradeoffs Among Exchange Rates, Monetary Policies, and Capital Mobility. *Review of Economics and Statistics* 87(3): 423–438.

Quinn, Dennis, Martin Schindler, and A. Maria Toyoda. 2011. Assessing measures of financial openness and integration. *IMF Economic Review* 59(3): 488–522.

Ramey, Valerie A. 2016. Macroeconomic Shocks and Their Propagation. In *Handbook of Macroeconomics*, edited by Taylor, John B., and Harald Uhlig, volume 2 of *Handbook of Macroeconomics*, chapter 2, 71–162. Amsterdam: Elsevier.

Schmitt-Grohé, Stephanie, and Martın Uribe. 2003. Closing small open economy models. *Journal of international Economics* 61(1): 163–185.

Schmitt-Grohé, Stephanie, and Martin Uribe. 2016. Downward Nominal Wage Rigidity, Currency. Pegs, and Involuntary Unemployment. *Journal of Political Economy* 124(5): 1466–1514.

Shambaugh, Jay C. 2004. The Effect of Fixed Exchange Rates on Monetary Policy. *Quarterly Journal of Economics* 119(1): 301–352.

Smets, Frank, and Rafael Wouters. 2007. Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach. *American Economic Review* 97(3): 586–606.

Stadler, George W. 1990. Business Cycle Models with Endogenous Technology. *American Economic Review* 80(4): 763–78.

Syverson, Chad. 2011. What determines productivity? *Journal of Economic Literature* 49(2): 326–365.

Uribe, Martin, and Stephanie Schmitt-Grohé. 2017. *Open Economy Macroeconomics*. Princeton, N.J.: Princeton University Press.

Werning, Ivan. 2011. Managing a Liquidity Trap: Monetary and Fiscal Policy. NBER Working Papers 17344, National Bureau of Economic Research, Inc.

Woodford, Michael. 2003. *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton, N.J.: Princeton University Press.

Yellen, Janet L. 2016. Macroeconomic Research After the Crisis. Speech at the conference "The Elusive 'Great' Recovery: Causes and Implications for Future Business Cycle Dynamics," Federal Reserve Bank of Boston, October.

Online Appendix

A. Proofs of consistency for impulse responses

This section provides the basic ideas behind the proofs of consistency for truncated VARs and LPs when the true DGP is an invertible MA(∞). The reader is referred to the references cited for additional details.

A.1. Data generating process and main assumptions

Assume the data generating process for the m–dimensional vector process $\boldsymbol{y}_t$ is:

$$\boldsymbol{y}_t = \sum_{j=0}^{\infty} B_j \boldsymbol{\epsilon}_{t-j}; \quad B_0 = I; \quad \sum_{j=0}^{\infty} ||B_j|| < \infty\,, \tag{37}$$

where $||B_j||^2 = tr(B_j' B_j)$ and $B(z) = \sum_{j=0}^{\infty} B_j z_j$ such that $\det\{B(z)\} \neq 0$ for $|z| \leq 1$. Under these assumptions, this invertible $MA(\infty)$ can also be expressed as:

$$\boldsymbol{y}_t = \sum_{j=1}^{\infty} A_j \boldsymbol{y}_{t-j} + \boldsymbol{\epsilon}_t; \quad \sum_{j=1}^{\infty} ||A_j|| < \infty; \quad \det\{A(z)\} \neq 0 \text{ for } |z| \leq 1\,.$$

Further, we make assumptions 1–4 following Lewis and Reinsel (1985), and Lusompa (2019) (Kuersteiner (2005) makes somewhat stronger assumptions because he later derives testing procedures to determine the optimal lag length). These assumptions are:

Assumption 1 $\{\boldsymbol{y}_t\}$ is generated by Equation 37.

Assumption 2 $E|\epsilon_{it}\epsilon_{jt}\epsilon_{kt}\epsilon_{lt}| \leq \gamma_4 < \infty$ for $1 \leq i, j, k, l \leq m$.

Assumption 3 The truncation lag p is chosen as a function of the sample size T such that $p^2/T \to 0$ as $p, T \to \infty$.

Assumption 4 p is chosen as a function of T such that

$$p^{1/2} \sum_{j=p+1}^{\infty} ||A_j|| \to 0 \quad \text{as} \quad p, T \to \infty.$$

Then, as discussed in the text, Lewis and Reinsel (1985) show:

$$||\hat{A}_j - A_j|| \xrightarrow{p} 0 \quad \text{as} \quad p, T \to \infty.$$

This well-known result says that even when the data are generated by an infinite-order process, the coefficients of the first p terms are consistently estimated. We show next that despite this result, inconsistencies in the estimation of impulse responses can crop up.

A.2. Potential sources of bias in truncated VARs

In finite samples, inconsistent estimates of the impulse response function can arise from at least two sources that we now quantify: (1) the truncation lag is too short given Assumptions 1–4; and (2) the truncation lag is appropriate, but the impulse response is calculated for periods that extend beyond the truncation lag. To investigate the first source of inconsistency, rewrite the $VAR(\infty)$ as

$$\begin{aligned}
\boldsymbol{y}_t &= \sum_{j=1}^{k} A_j \boldsymbol{y}_{t-j} + \boldsymbol{u}_t\,, \\
\boldsymbol{u}_t &= \sum_{j=k+1}^{p} A_j \boldsymbol{y}_{t-j} + \sum_{j=p+1}^{\infty} A_j \boldsymbol{y}_{t-j} + \boldsymbol{\epsilon}_t\,,
\end{aligned}$$

where we assume $k < p$ and p is the truncation lag that meets Assumptions 1–4 of the proof of consistency. Hence rewrite the previous expression as

$$\boldsymbol{y}_t = A(k) X_{k,t-1} + \boldsymbol{u}_t; \quad A(k) = (A_1, \ldots, A_k); \quad X_{k,t-1} = (\boldsymbol{y}_{t-1}, \ldots, \boldsymbol{y}_{t-k})'\,.$$

The least-squares estimate of $A(k)$ is therefore

$$\hat{A}(k) = \left(\frac{1}{T-k} \sum_{p}^{T} \boldsymbol{y}_t X'_{k,t-1} \right) \left(\frac{1}{T-k} \sum_{p}^{T} X_{k,t-1} X'_{k,t-1} \right)^{-1}\,.$$

Hence

$$\hat{A}(k) = A(k) + \left(\frac{1}{T-k} \sum_{p}^{T} \boldsymbol{u}_t X'_{k,t-1} \right) \left(\frac{1}{T-k} \sum_{p}^{T} X_{k,t-1} X'_{k,t-1} \right)^{-1}\,.$$

Given the three components of $\boldsymbol{u}_t$, it is easy to see that the source of inconsistency in estimates of the first k autoregressive terms will come from the component

$$\left(\frac{1}{T-k} \sum_{p}^{T} \sum_{j=k+1}^{p} A_j \boldsymbol{y}_{t-j} X'_{k,t-1} \right) \left(\frac{1}{T-k} \sum_{p}^{T} X_{k,t-1} X'_{k,t-1} \right)^{-1}\,,$$

since the proof of consistency in Lewis and Reinsel (1985) shows that the other two terms vanish asymptotically. The source of inconsistency can be quantified by noticing that

$$\left(\frac{1}{T-k} \sum_{p}^{T} X_{k,t-1} X'_{k,t-1} \right)^{-1} \to \begin{pmatrix} \Gamma(0) & \Gamma(1) & \cdots & \Gamma(k) \\ \Gamma(1) & \Gamma(0) & \cdots & \Gamma(k-1) \\ \vdots & \vdots & \vdots & \vdots \\ \Gamma(k) & \Gamma(k-1) & \cdots & \Gamma(0) \end{pmatrix}^{-1} \to \Gamma_k^{-1}\,,$$

as shown in Lewis and Reinsel (1985), where $E(\boldsymbol{y}_t \boldsymbol{y}'_{t-j}) = \Gamma(j)$ and $\Gamma(-j) = \Gamma(j)'$. Hence, asymptotically, the source of inconsistency is

$$\sum_{k+1}^{p} A_j \left(\Gamma(j-1), \ldots, \Gamma(j-k) \right) \Gamma_k^{-1}\,.$$

However, even when the lag-length p is chosen to be sufficiently large, another source of bias can crop up into the estimation of the impulse response. In particular, following Durbin (1959), we know that

$$B_h = A_1 B_{h-1} + A_2 B_{h-2} + \cdots + A_{h-1} B_1 + A_h \,.$$

If the VAR is truncated at lag k, for $k \leq p$, it is easy to see that the previous expression becomes

$$B_h = A_1 B_{h-1} + A_2 B_{h-2} + \cdots + A_k B_{k-h} + \underbrace{A_{k+1} B_{k-h-1} + \cdots + A_{h-1} B_1 + A_h}_{remainder} \,.$$

and, hence,

$$||\hat{B}_h - B_h|| = ||A_{k+1}(\hat{B}_{k-h-1} - B_{k-h-1}) + \ldots + A_{h-1}(\hat{B}_1 - B_1) + A_h|| \not\to 0\,,$$

since $||\hat{B}_{k+j} - B_{k+j}||$ for $j \geq 1$ is not guaranteed to vanish asymptotically.

Next notice that the impulse response for horizons $h > k$ will be estimated using the recursion

$$\hat{B}_{k+j} = \hat{A}_1 \hat{B}_{k+j-1} + \ldots + \hat{A}_k \hat{B}_j; \quad j = 1, \ldots, H\,.$$

Even if $k = p$, and hence $||\hat{A}_j|| \to A_j$ for $j = 1, \ldots, k$, the fact remains that the *remainder* term

$$A_{k+1} B_{k-h-1} + \cdots + A_{h-1} B_1 + A_h$$

cumulates increasing sums of coefficients that are not estimated in the model. As the Monte Carlo exercise showed earlier, the inconsistency at longer horizons tends to accumulate.

A.3. The consistency of the local projections estimator

In this section we use the same assumptions as in the previous section to establish the consistency of the local projections estimator at any horizon.

Using the VAR(∞) representation of the DGP and recursive substitution, it is easy to see that

$$\boldsymbol{y}_{t+h} = B_{h+1} \boldsymbol{y}_{t-1} + \{C_{h+2} \boldsymbol{y}_{t-2} + C_{h+3} \boldsymbol{y}_{t-3} + \cdots\} + \boldsymbol{\epsilon}_{t+h} + B_1 \boldsymbol{\epsilon}_{t+h-1} + \boldsymbol{\epsilon}_t B_h\,,$$

where

$$\begin{aligned} C_{h+2} &= B_h A_1 + \cdots + B_1 A_h + A_{h+1}\,, \\ C_{h+3} &= B_h A_2 + \cdots + B_1 A_{h+1} + A_{h+2}\,, \\ &\vdots \\ C_{h+k} &= B_h A_{k-1} + \cdots + B_1 A_{h+k-2} + A_{h+k-1}\,. \end{aligned}$$

Now, consider truncating the lag of the local projection at $k = p$, where p meets Assumptions 1–4 of the Lewis and Reinsel (1985) consistency theorem discussed in the previous section.

Then the truncated local projection can be written as

$$\boldsymbol{y}_{t+h} = B_{h+1} \boldsymbol{y}_{t-1} + C_{h+2} \boldsymbol{y}_{t-2} + C_{h+3} \boldsymbol{y}_{t-3} \cdots + C_{h+k} \boldsymbol{y}_{t-k} + \boldsymbol{u}_{t+h}\,,$$

$$\boldsymbol{u}_{t+h} = \boldsymbol{\epsilon}_{t+h} + \{B_1 \boldsymbol{\epsilon}_{t+h-1} + B_2 \boldsymbol{\epsilon}_{t+h-2} + \cdots + B_h \boldsymbol{\epsilon}_t\} + \{C_{h+k+1} \boldsymbol{y}_{t-k-1} + C_{h+k+2} \boldsymbol{y}_{t-k-2} + \cdots\}\,.$$

Let $D = (B_h, C_{h+2}, \ldots, C_{h+k})$ and $X_{t-1} = (\boldsymbol{y}_{t-1}, \ldots, \boldsymbol{y}_{t-k})'$ as defined earlier but where the subscript k is omitted here for simplicity. Then the local projection can be compactly written as

$$\boldsymbol{y}_{t+h} = DX_{t-1} + \boldsymbol{u}_{t+h}\,.$$

The least-squares estimate of D is simply

$$\hat{D} = \left(\frac{1}{T-h-k}\sum_{k}^{T-h} \boldsymbol{y}_{t+h}X'_{t-1}\right)\left(\frac{1}{T-h-k}\sum_{k}^{T-h} X_{t-1}X'_{t-1}\right)^{-1},$$

from where consistency can be determined from the following expression

$$\hat{D} = D + \left(\frac{1}{T-h-k}\sum_{k}^{T-h} \boldsymbol{u}_{t+h}X'_{t-1}\right)\left(\frac{1}{T-h-k}\sum_{k}^{T-h} X_{t-1}X'_{t-1}\right)^{-1}.$$

Lewis and Reinsel (1985) show that $||\Gamma_k^{-1}||_1$ is uniformly bounded where we use the fact that $||AB||^2 \leq ||A||_1^2||B||^2$; as well as $||AB||^2 \leq ||A||^2||B||_1^2$ where $||C||_1^2 = sup_{l\neq 0} l'C'Cl/l'l$, the largest eigenvalue of $C'C$ (see Wiener and Masani, 1958).

Now we turn our focus to the terms

$$\begin{aligned}\frac{1}{T-h-k}\sum_{k}^{T-h} \boldsymbol{u}_{t+h}X'_{t-1} &= \frac{1}{T-h-k}\sum_{k}^{T-h} (\boldsymbol{\epsilon}_{t+h} + B_1\boldsymbol{\epsilon}_{t+h-1} + \cdots + B_h\boldsymbol{\epsilon}_t)X'_{t-1}\\ &= \frac{1}{T-h-k}\sum_{k}^{T-h} (C_{h+k}\boldsymbol{y}_{t-k-1} + C_{h+k+1}\boldsymbol{y}_{t-k-2} + \cdots)X'_{t-1}\,.\end{aligned}$$

It is easy to see that

$$\frac{1}{T-h-k}\sum_{k}^{T-h} \boldsymbol{\epsilon}_{t+h}X'_{t-1} \to 0\,,$$
$$\frac{B_j}{T-h-k}\sum_{k}^{T-h} \boldsymbol{\epsilon}_{t+h-j}X'_{t-1} \to 0\,,$$

since $||B_j|| < \infty$ for $j = 1, \cdots, h$. Hence, the only tricky part is to examine the terms

$$\frac{C_{h+k+j}}{T-h-k}\sum_{k}^{T-h} \boldsymbol{y}_{t-k-(j+1)}X'_{t-1} \qquad \text{for } j = 0, 1, \ldots\,.$$

Note that

$$C_{h+k+j} = B_hA_{k+k} + \cdots + B_1A_{h+k+j-1} + A_{h+k+j} \qquad j = 0, 1, \ldots,$$

hence

$$\begin{aligned}
\sum_{j=0}^{\infty}||C_{h+k+j}|| &= \sum_{j=0}^{\infty}||B_h A_{k+k} + \cdots + B_1 A_{h+k+j-1} + A_{h+k+j}|| \\
&\leq \sum_{j=0}^{\infty}||B_h A_{k+j}|| + \cdots + \sum_{j=1}^{\infty}||B_1 A_{h+k+j-1}|| + \sum_{j=0}^{\infty}||A_{h+j+j}|| \\
&= ||B_h||_1 \sum_{j=0}^{\infty}||A_{k+j}|| + \cdots + ||B_1||_1 \sum_{j=1}^{\infty}||A_{h+k+j-1}|| + \sum_{j=0}^{\infty}||A_{h+k+j}||.
\end{aligned}$$

From the assumptions we know that the $||B_j||_1$ are uniformly bounded, and also that

$$k^{1/2}\sum_{j=0}^{\infty}||A_{k+j}|| \to 0 \quad \Longrightarrow \quad k^{1/2}\sum_{j=0}^{\infty}||C_{h+k+j}|| \to 0\,,$$

and this condition can now be used to show that

$$\frac{\sum_{j=0}^{\infty} C_{h+k+j}}{T-h-k}\sum_{k}^{T-h} \boldsymbol{y}_{t-k-(j+1)}X'_{t-1} \to 0, \quad \text{as} \quad k, T \to \infty\,.$$

Summarizing, these derivations show that the same conditions that ensure consistency of the coefficients estimates in a truncated VAR also ensure consistency of the local projections with truncated lag length. However, because the coefficient for $\boldsymbol{y}_{t-1}$ in the local projection is a direct estimate of the impulse response coefficient, then we directly get a proof of consistency for the coefficients of the impulse response at any horizon regardless of truncation.

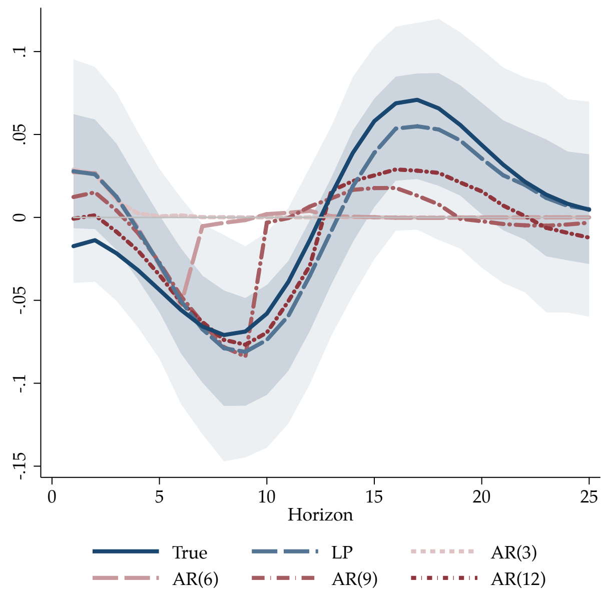
A.4. Monte Carlo results for impulse response estimators

This section provides details of the Monte Carlo experiments reported in the main text in addition to presenting complementary Monte Carlo experiments based on the same simulated data, but presenting the impulse response (rather than the cumulated response itself).

The data are generated as a $MA(25)$ model whose coefficients are generated by the following Gaussian Basis Function: $\theta_j = a\exp(-((j-b)/c)^2)$ for $j = 1, \ldots, 25$ and for $a = -0.5; b = 12;$ and $c = 6$. This results in the impulse and cumulative responses shown in panel (a) of Figure 1. The error terms are assumed to be standard Gaussian. The left hand side variable is expressed in the differences to replicate exactly the estimation of the cumulative response in the empirical section. We simulate samples of size 1,500, but the first 500 observations are then discarded to avoid initialization issues. Using these data, we then estimate $AR(k)$ models for $k = 3, 6, 9, 12$ and local projections using 2 lags.

As a complement to Figure 1, Figure A1 presents the experiments based on the impulse response itself to illustrate the consistency of the $AR(k)$ estimators up to horizon $h \leq k$ but not beyond. The solid blue line is the true response based on our parameter choices for the D.G.P. The dashed blue line with Monte Carlo one and two standard error bands are the local projections using two lags only. The dotted maroon lines are the impulse responses from AR models with 3, 6, 9, and 12 lags as in the Monte Carlo in the main text. As the figure clearly shows, impulse response coefficients are estimated well using the $AR(k)$ models up to horizon $h = k$, as the asymptotic theory just presented showed. In contrast, the local projection estimator does well across all horizons. The cumulative versions of these responses are the experiments reported in Figure 1 in the main text.

Figure A1: *Estimating non-cumulative responses: autoregressive versus local projection biases at long horizons.*



Notes: sample size = 1,000. Monte Carlo replications: 1,000. Error bands in light blue are 1 and 2 standard error bands based on the local projection Monte Carlo average. $AR(k)$ for k = 3, 6, 9, 12 refers to impulse responses from an autoregressive model with k lags. See text.

B. Equilibrium conditions in the baseline model

A perfect foresight equilibrium in the baseline model (subsection 2.1) is given by a sequence of 16 processes $\{C_t, C_{Tt}, C_{Nt}, d_{t+1}, p_t, \Pi_{Tt}, R_t^n, R_t, w_t, L_t, \Delta_{pNt}, \mathcal{E}_t, \Pi_{Nt}, \tilde{p}_{Nt}, \mathcal{K}_{Npt}, \mathcal{Z}_{Npt}\}$ that satisfy the following equilibrium conditions for a given sequence of exogenous processes $\{Y_{Tt}, R_t^*, \Pi_t^*\}$ and initial values $\{d_0, \mathcal{E}_{-1}, p_{-1}, \Delta_{pN-1}\}$,

$$C_{Tt} + d_t = Y_{Tt} + \frac{d_{t+1}}{R_t}, \tag{38}$$

$$C_t = \left(\frac{C_{Tt}}{\omega}\right)^{\omega} \left(\frac{C_{Nt}}{1-\omega}\right)^{1-\omega}, \tag{39}$$

$$p_t = \frac{(1-\omega)C_{Tt}}{\omega C_{Nt}}, \tag{40}$$

$$C_{Tt}^{-1} = \beta \mathbb{E}_t \left\{ \frac{C_{Tt+1}^{-1} R_t^n}{P_{Tt+1}/P_{Tt}} \right\}, \tag{41}$$

$$C_{Tt}^{-1} = \beta \mathbb{E}_t \left\{ C_{Tt+1}^{-1} R_t \right\}, \tag{42}$$

$$R_t = R_t^* + \psi(e^{d_{t+1}-\bar{d}} - 1), \tag{43}$$

$$\tilde{p}_{Nt} = \frac{\mathcal{K}_{Npt}}{\mathcal{Z}_{Npt}}, \tag{44}$$

$$\mathcal{K}_{Npt} = \frac{\epsilon_p}{\epsilon_p - 1}(1-\tau^p) w_t C_{Nt} + \theta_p \beta \frac{C_{Tt}}{C_{Tt+1}} \Pi_{Nt+1}^{\epsilon_p} \mathcal{K}_{Npt+1}, \tag{45}$$

$$\mathcal{Z}_{Npt} = p_t C_{Nt} + \theta_p \beta \frac{C_{Tt}}{C_{Tt+1}} \Pi_{Nt+1}^{\epsilon_p - 1} \mathcal{Z}_{Npt+1}\,, \tag{46}$$

$$1 = \theta_p \Pi_{Nt}^{\epsilon_p - 1} + (1-\theta_p) \tilde{p}_{Nt}^{1-\epsilon_p}\,, \tag{47}$$

$$\frac{\varphi L_t C_{Tt}}{\omega} = w_t\,, \tag{48}$$

$$\frac{1}{\Delta_{p_{Nt}}} L_{Nt} = C_{Nt}\,, \tag{49}$$

$$\Delta_{p_{Nt}} = (1-\theta_p) \tilde{p}_{Nt}^{-\epsilon_p} + \theta_p \Pi_{Nt}^{\epsilon_p} \Delta_{p_{Nt-1}}\,, \tag{50}$$

$$\frac{p_t}{p_{t-1}} = \frac{\Pi_{Nt}}{\Pi_{Tt}}\,, \tag{51}$$

$$\Pi_{Tt} = \frac{\mathcal{E}_t}{\mathcal{E}_{t-1}} \Pi_t^*\,. \tag{52}$$

and one of the following three equations for the respective policy regime:

$$\mathcal{E}_t = 1 \tag{peg}$$

$$\Pi_{Nt} = 1 \tag{float}$$

$$R_t^n = \bar{R}^n e^{\varepsilon_t} \tag{benchmark}$$

C. Solution for the baseline model

The following system of equations solves for equilibrium in five endogenous variables under the peg and benchmark economy respectively:

$$\hat{C}_{Tt} = \hat{C}_{Tt+1} - \hat{R}_t^*\,, \tag{53}$$

$$\hat{Y}_{Nt} = \hat{Y}_{Nt+1} - (\hat{R}_t^n - \hat{\Pi}_{Nt+1})\,, \tag{54}$$

$$\hat{\Pi}_{Nt} = \beta \hat{\Pi}_{Nt+1} + \kappa \hat{Y}_{Nt}\,, \tag{55}$$

$$\hat{\mathcal{E}}_{t+1} - \hat{\mathcal{E}}_t = \hat{R}_t^n - \hat{R}_t^*\,, \tag{56}$$

$$\hat{\mathcal{E}}_{t+1} - \hat{\mathcal{E}}_t = 0\,, \tag{peg}$$

$$\hat{R}_t^n = \varepsilon_t\,. \tag{benchmark}$$

Furthermore, let the shocks to R_t^* or ε_t follow AR(1) process with persistence ρ.

We solve the model backwards. Denote the time at which the economy returns back to initial steady state with $\tilde{t}+1$ such that $\hat{C}_{T,\tilde{t}+1} = \hat{Y}_{N,\tilde{t}+1} = \hat{\Pi}_{N,\tilde{t}+1} = \hat{\mathcal{E}}_{\tilde{t}+1} = \hat{R}_{\tilde{t}+1}^n = \hat{R}_{\tilde{t}+1}^* = 0$. Tradable goods consumption, under a peg, is given by

$$\hat{C}_{T,\tilde{t}-s} = -\sum_{j=0}^{s} \rho^j \hat{R}_{\tilde{t}-j}^*; \quad \forall\, 0 \le s \le \tilde{t}$$

, whereas tradable goods' consumption does not change in response to shock to ε_t in the benchmark economy. The solution of non-tradable output and non-tradable goods inflation is given by

$$\hat{Y}_{N,\tilde{t}-s} = -\alpha_{Y_N,\tilde{t}-s} \hat{R}_{\tilde{t}-s}; \quad \hat{\Pi}_{N,\tilde{t}-s} = -\alpha_{\Pi_N,\tilde{t}-s} \hat{R}_{\tilde{t}-s}\,,$$

where for $s < 0$, $\alpha_{Y_N,\tilde{t}-s} = \alpha_{\Pi_N,\tilde{t}-s} = 0$; $\alpha_{Y_N,\tilde{t}} = 1$, $\alpha_{\Pi_N,\tilde{t}} = \kappa$; and $\forall\, 0 < s \le \tilde{t}$, $\alpha_{Y_N,\tilde{t}-s} =$

$\left(\alpha_{Y_N,\tilde{t}-s+1}+\alpha_{\Pi_N,\tilde{t}-s+1}\right)\rho+1>0$; and $\alpha_{\Pi_N,\tilde{t}-s}=\left(\kappa\alpha_{Y_N,\tilde{t}-s}+\rho\beta\alpha_{\Pi_N,\tilde{t}-s+1}\right)>0$.

For a similar shock process, the drop in non-tradable output is identical across the peg and benchmark economies.[25]

D. Extension: time-varying aggregation weights

We consider the more general extension of the baseline model (subsection 2.1) allowing for time-variation in aggregation weights in the construction of total output. The consumption aggregator is : $C_t=\Psi C_{Tt}^{\omega}C_{Nt}^{1-\omega}$, where $\Psi\equiv\omega^{-\omega}(1-\omega)^{1-\omega}$ is a scaling factor. This implies that domestic CPI is given by $P_t=P_{Tt}^{\omega}P_{Nt}^{1-\omega}$. Total nominal output is $P_{Tt}Y_{Tt}+P_{Nt}Y_{Nt}$. Let total output be denoted with Y_t, and is given by:

$$Y_t=\frac{P_{Tt}Y_{Tt}+P_{Nt}Y_{Nt}}{P_t}=p_t^{\omega-1}Y_{Tt}+p_t^{\omega}Y_{Nt}\,,$$

where $p_t\equiv\frac{P_{Nt}}{P_{Tt}}$. From the optimality conditions, we have that

$$p_t=\frac{(1-\omega)C_{Tt}}{\omega C_{Nt}}\,.$$

In terms of log-deviations from steady state, total output is given by

$$\hat{Y}_t=\left[(\omega-1)p^{\omega-1}+\omega p^{\omega}\right]\ \left(\hat{C}_{Tt}-\hat{Y}_{Nt}\right)+\frac{P_TY_T}{PY}\hat{Y}_{Tt}+\frac{P_NY_N}{PY}\hat{Y}_{Nt}\,.$$

When $\omega\to 0$ (tradable goods share is infinitesimally small),

$$\hat{Y}_t=p^{-1}\ \left(\hat{Y}_{Nt}-\hat{C}_{Tt}\right)+\frac{P_TY_T}{PY}\hat{Y}_{Tt}+\frac{P_NY_N}{PY}\hat{Y}_{Nt}\,.$$

In the baseline model, with exogenous endowment of tradable goods,

$$\hat{Y}_t=p^{-1}\ \left(\hat{Y}_{Nt}-\hat{C}_{Tt}\right)+\frac{P_NY_N}{PY}\hat{Y}_{Nt}\,.$$

Recall that $\hat{Y}_{Nt}$ is identical across the peg and the benchmark economy as proved in Proposition 2. From results in Appendix C, sequence of $\hat{C}_{Tt}<0$ under a peg and equal to 0 under benchmark economy. Hence the response of total output under a peg is downward biased relative to that under the benchmark economy, where the bias is given by: $p^{-1}\sum_{j=0}^{s}\rho^j\hat{R}^*_{\tilde{t}-s};\quad\forall\,0\le s\le\tilde{t}$. The converse result applies when $\omega\to 1$.

E. Extension: endogenous tradable output in the baseline model

We extend the baseline model (subsection 2.1) by allowing tradable output to be produced with a constant returns to scale production function in labor. Labor is fully mobile across the tradable and non-tradable sector. Thus, economy-wide real wages (in units of tradable goods) are constant. The

[25] Furthermore, it follows from the above solution, the drop in non-tradable goods output is larger than the drop in tradable goods consumption due to a deflationary effect of the shock on non-tradable goods prices.

intra-temporal labor supply condition is

$$\nu\left(\frac{L_T}{L}\hat{L}_{Tt}+(1-\frac{L_T}{L})\hat{L}_{Nt}\right)+\hat{C}_{Tt}=0\,,$$

where $\frac{L_T}{L}$ is fraction of total labor force allocated to the tradable goods sector in the steady state. The non-tradable goods price-Phillips curve is

$$\hat{\Pi}_{Nt}=\beta\hat{\Pi}_{Nt+1}+\kappa\hat{Y}_{Nt}-\kappa\hat{C}_{Tt}\,.$$

The rest of the equations are same as described in Appendix C. We show the solution under two assumptions on persistence of the shock process.

E.1. One period unanticipated shock

When the shock is one-period lived at time t_0, then the solution of non-tradable output at time t_0 (when the shock hits) is given by

$$\hat{Y}_{Nt_0}=-\hat{R}_{t_0}\,.$$

The tradable goods output at time t_0 is given by

$$\hat{Y}_{Tt_0}=-\frac{L}{L_T}\nu^{-1}\hat{C}_{Tt_0}-\left(\frac{L}{L_T}-1\right)\hat{Y}_{Nt_0}\,.$$

The tradable goods output production goes up in response to a one-time increase in ε_t in the benchmark economy and $\hat{R}^*_t$ in the peg economy. This arises due to an increase in labor supply in the tradable goods sector following a contraction in demand for labor in the non-tradable sector. While the impulse response of non-tradable output is identical across the peg and the benchmark economy, total output response is biased downwards in the peg economy relative the benchmark economy. The downward bias is given by $-\nu^{-1}\frac{P_TY_T}{PY}\frac{L}{L_T}\hat{C}_{Tt_0}$, where $\hat{C}_{Tt_0}=-\hat{R}^*_t$ for the peg economy. The tradable goods consumption does not change in response to shock to ε_t in the benchmark economy.

E.2. Unanticipated AR(1) shock

Let the shocks to R^*_t or ε_t follow AR(1) process with persistence ρ.

We solve the model backwards. Denote the time at which the economy returns back to initial steady state with $\tilde{t}+1$ such that $\hat{C}_{T,\tilde{t}+1}=\hat{Y}_{N,\tilde{t}+1}=\hat{\Pi}_{N,\tilde{t}+1}=\hat{\mathcal{E}}_{\tilde{t}+1}=\hat{R}^n_{\tilde{t}+1}=\hat{R}^*_{\tilde{t}+1}=0$.

Benchmark Economy

Tradable goods consumption does not change in response to shock to ε_t in the benchmark economy. The response of non-tradable output and non-tradable goods' inflation is same as solved in Appendix C. That is, the solution of non-tradable output and non-tradable goods inflation is given by

$$\hat{Y}_{N,\tilde{t}-s}=-\alpha_{Y_N,\tilde{t}-s}\hat{R}_{\tilde{t}-s};\quad \hat{\Pi}_{N,\tilde{t}-s}=-\alpha_{\Pi_N,\tilde{t}-s}\hat{R}_{\tilde{t}-s}\,,$$

where for $s<0$, $\alpha_{Y_N,\tilde{t}-s}=\alpha_{\Pi_N,\tilde{t}-s}=0$; $\alpha_{Y_N,\tilde{t}}=1$, $\alpha_{\Pi_N,\tilde{t}}=\kappa$; and $\forall\ 0<s\leq\tilde{t}$, $\alpha_{Y_N,\tilde{t}-s}=\left(\alpha_{Y_N,\tilde{t}-s+1}+\alpha_{\Pi_N,\tilde{t}-s+1}\right)\rho+1>0$; and $\alpha_{\Pi_N,\tilde{t}-s}=\left(\kappa\alpha_{Y_N,\tilde{t}-s}+\rho\beta\alpha_{\Pi_N,\tilde{t}-s+1}\right)>0$.

The tradable goods output is given by

$$\hat{Y}_{Tt} = -\left(\frac{L}{L_T} - 1\right)\hat{Y}_{Nt} \quad \forall t\,.$$

Peg Economy

Tradable goods consumption, under a peg, is given by

$$\hat{C}_{T,\tilde{t}-s} = -\sum_{j=0}^{s} \rho^j \hat{R}^*_{\tilde{t}-j}; \quad \forall\, 0 \leq s \leq \tilde{t}\,.$$

Under a peg, the solution of non-tradable output and non-tradable goods inflation is given by

$$\hat{Y}_{N,\tilde{t}-s} = -\sum_{j=0}^{s} \rho^j \hat{R}^*_{\tilde{t}-j}; \quad \hat{\Pi}_{N,\tilde{t}-s} = 0\,.$$

The tradable goods output is given by

$$\hat{Y}_{Tt} = -\frac{L}{L_T}\nu^{-1}\hat{C}_{Tt} - \left(\frac{L}{L_T} - 1\right)\hat{Y}_{Nt} \quad \forall t\,.$$

E.3. Downward bias in estimates from peg

The response of total output is downward biased under a peg relative to that under the benchmark economy. With a persistent shock, there are two sources of bias: 1) increase in tradable goods production is higher under a peg relative to the benchmark economy, and (2) reduction in non-tradable goods output is lower under a peg relative to the benchmark economy. The second source of bias is absent in the case of a one-period shock.

F. A medium scale small-open economy DSGE model with hysteresis

Relative to the baseline model, production of non-tradable goods uses capital as well as labor. Investment is done in the form of non-tradable goods. Capital accumulation is subject to investment adjustment costs. And there are frictions to adjusting capital utilization.

F.1. Household

Consumers

Each household supplies differentiated labor indexed by j. Household j chooses consumption aggregate C_t, risk-free nominal bonds B_t, real foreign bonds d_{t+1}, investment I_t and capital utilization u_t to maximize the utility function, with external habits over consumption,

$$\mathbb{E}_t\Sigma_{s=t}^{\infty}\beta^{s-t}\left[\log(C_{j,s}) - \frac{\varphi}{1+\nu}L_{j,s}^{1+\nu}\right]\,.$$

The composite good C_t is a Cobb-Douglas aggregate of a tradable good C_{Tt} and a non-tradable good C_{Nt} such that $C_t = \Psi(C_{Tt})^{\omega}(C_{Nt} - hC_{Nt-1})^{1-\omega}$ where $\Psi = \omega^{-\omega}(1-\omega)^{1-\omega}$ is a scaling factor,

$0 < \omega < 1$ is the weight on tradable goods, h is the degree of (external) habit formation in non-tradable good consumption, $\nu > 0$ is the inverse Frisch elasticity of labor supply, $\varphi > 0$ is a parameter that pins down the steady-state level of hours, and the discount factor β satisfies $0 < \beta < 1$.

We assume perfect consumption risk sharing across the households.

As a result, household's budget constraint in period t is given by

$$\mathcal{E}_t D_t + B_t + P_{Nt} I_t + P_{Tt} C_{Tt} + P_{Nt} C_{Nt} =$$
$$(1+\tau^w) W_t L_{j,t} + P_{Tt} Y_{Tt} + \mathcal{E}_t \frac{D_{t+1}}{1+r_t} + \frac{B_{t+1}}{1+i_t} + T_t + Z_t + R^K_t u_t K^u_t - P_{Nt} a(u_t) K^u_t \,, \tag{57}$$

where P_{Tt} and P_{Nt} denote the prices of a unit of tradable and non-tradable good, respectively, in units of local currency; D_t is the level of real debt denominated in units of tradable good assumed in period $t-1$ and due in period t; B_t is the level of nominal debt denominated in units of local currency assumed in period $t-1$ and due in period t; I_t is investment in physical capital, W_t is the nominal wage; Labor income $W_t L_{j,t}$ is subsidized at a fixed rate τ_w. T_t are nominal lump-sum transfers from the government; and Z_t are nominal profits from domestic firms owned by the households; and $Y_{Tt} > 0$ is the endowment of tradable goods received by the households. Since households own the capital and choose the utilization rate, the amount of effective capital that the households rent to the firms at nominal rate R^K_t is

$$K_t = u_t K^u_t \,.$$

The (nominal) cost of capital utilization is $P_{Nt} a(u_t)$ per unit of physical capital. As in the literature (Smets and Wouters 2007) we assume $a(1) = 0$ in the steady state and $a'' > 0$. Following Christiano, Eichenbaum, and Evans (2005), we assume investment adjustment costs in the production of capital. Law of motion for capital is as follows:

$$K^u_{t+1} = \left[1 - S\left(\frac{I_t}{(1+g_{ss}) I_{t-1}}\right)\right] I_t + (1-\delta_k) K^u_t \,,$$

where $g_{ss} \equiv \bar{\mu}$ is the steady state growth rate of aggregate productivity A_t.

Utility maximization delivers the first order condition linking the inter-temporal consumption smoothing to the marginal utility of nominal and real bonds

$$1 = \beta \mathbb{E}_t \left[\frac{\Lambda_{t+1}}{\Lambda_t} (1+i_t) \frac{P_{Tt}}{P_{Tt+1}}\right] \,, \tag{58}$$

$$1 = \beta \mathbb{E}_t \left[\frac{\Lambda_{t+1}}{\Lambda_t} (1+r_t)\right] \,, \tag{59}$$

where Λ_t is the marginal utility of tradable goods consumption given by

$$\Lambda_t = \frac{1}{C_{Tt}} \,.$$

Marginal rate of substitution between tradable and non-tradable goods is equal to relative price

$$\frac{P_{Nt}}{P_{Tt}} = \frac{(1-\omega) C_{Tt}}{\omega (C_{Nt} - h C_{Nt-1})} \,. \tag{60}$$

Define relative price of non-tradables as $p_t \equiv \frac{P_{Nt}}{P_{Tt}}$. We assume that law of one price on the tradable good holds. Let $\mathcal{E}_t$ be the nominal exchange rate for home relative to the base, and let P_t^* be the base price of the tradable good denominated in base currency.[26] Then, we have that $P_{Tt} = \mathcal{E}_t \; P_t^*$. From consumption euler equations, we can then derive the interest rate parity condition,

$$(1+i_t) = (1+r_t)\,\frac{P_{Tt+1}}{P_{Tt}} = (1+r_t)\,\frac{\mathcal{E}_{t+1}}{\mathcal{E}_t}\,\frac{P_{t+1}^*}{P_t^*}\,. \tag{61}$$

Henceforth, we normalize $P_t^* = 1$.

To ensure stationarity under incomplete markets, we follow Schmitt-Grohé and Uribe (2003); Uribe and Schmitt-Grohé (2017) and assume that the home real interest rate is related to foreign real interest rate through a debt-elastic interest rate premium,[27]

$$\frac{1+r_t}{1+g_{t+1}} = \frac{1+r_t^*}{1+\bar{g}^*} + \psi(e^{d_{t+1}-\bar{d}} - 1)\,. \tag{62}$$

where g_t is productivity growth at home and g^* is productivity growth rate in the base economy.

The stochastic discount factor in period $t+1$ is given by

$$Q_{t,t+1} = \beta\,\frac{\Lambda_{t+1}}{\Lambda_t}\,\frac{P_{Tt}}{P_{Tt+1}}\,.$$

The household does not choose hours directly. Rather each type of worker is represented by a wage union who sets wages on a staggered basis. Consequently the household supplies labor at the posted wages as demanded by firms.

We introduce capital accumulation through households. Solving household problem for investment and capital yields the Euler condition for capital

$$q_t = \beta\mathbb{E}_t\left[\frac{\Lambda_{t+1}p_{t+1}}{\Lambda_t p_t}\left(r_{t+1}^K u_{t+1} - a(u_{t+1}) + q_{t+1}(1-\delta_k)\right)\right]\,,$$

where $r_{t+1}^K = R_{t+1}^K / P_{Nt+1}$, and the (relative) price of installed capital q_t is given by[28]

$$\begin{aligned} q_t\left[1 - S\left(\frac{I_t}{(1+g_{ss})I_{t-1}}\right) - S'\left(\frac{I_t}{(1+g_{ss})I_{t-1}}\right)\frac{I_t}{(1+g_{ss})I_{t-1}}\right] & \\ + \beta\frac{\Lambda_{t+1}p_{t+1}}{\Lambda_t p_t} q_{t+1}\frac{1}{(1+g_{ss})}\left(\frac{I_{t+1}}{I_t}\right)^2 S'\left(\frac{I_{t+1}}{(1+g_{ss})I_t}\right) &= 1\,. \end{aligned}$$

Choice of capital utilization rate yields

$$\frac{R_t^K}{P_{Nt}} = a'(u_t)\,.$$

[26]It is common in the small open economy literature to treat price level in the base economy P_t^* as synonymous for price level of tradable goods in the base economy P_{Tt}^*.

[27]Due to the presence of endogenous growth in the tradable good sector, the relevant real interest rate parity adjusts for productivity growth rate differentials between home and the base economy. Further note that for stationarity, we assume that $D_{t+1} = d_{t+1}A_{t+1}$.

[28]$q_t \equiv \frac{\Phi_t}{\Lambda_t p_t}$ where Φ_t is the Lagrange multipler on the capital accumulation equation.

Wage setting

Wage Setting follows Erceg, Henderson, and Levin (2000) and is relatively standard.

Perfectly competitive labor agencies combine j type labor services into a homogeneous labor composite L_t according to a Dixit-Stiglitz aggregation

$$L_t = \left[\int_0^1 L_{j,t}^{\frac{\epsilon_w - 1}{\epsilon_w}} dj\right]^{\frac{\epsilon_w}{\epsilon_w - 1}} .$$

where $\epsilon_w > 1$ is elasticity of substitution across labor varieties.

Labor unions representing workers of type j set wages on a staggered basis following Calvo (1983), taking given the demand for their specific labor input

$$L_{j,t} = \left(\frac{W_{j,t}}{W_t}\right)^{-\epsilon_w} L_t, \qquad \text{where } W_t = \left[\int_0^1 W_{j,t}^{1-\epsilon_w} dj\right]^{\frac{1}{1-\epsilon_w}} .$$

In particular, with probability $1-\theta$, the type-j union is allowed to re-optimize its wage contract and it chooses $\tilde{W}$ to minimize the disutility of working for laborer of type j, taking into account the probability that it will not get to reset wage in the future.[29] By the law of large numbers, the probability of changing the wage corresponds to the fraction of types who actually change their wage.

Consequently, the nominal wage evolves according to

$$W_t^{1-\epsilon_w} = (1-\theta_w)\tilde{W}_t^{1-\epsilon_w} + \theta_w W_{t-1}^{\epsilon_w - 1} ,$$

where the nominal wage inflation and non-tradable goods price inflation are related to each other

$$\Pi_t^w = \frac{W_t}{W_{t-1}} = \frac{w_t}{w_{t-1}} \frac{1}{\Pi_{Nt}} \frac{1}{1+g_t}$$

and where $\Pi_{Nt} \equiv \frac{P_{Nt}}{P_{Nt-1}}$ is the inflation rate in non-tradable goods sector, $w_t \equiv \frac{W_t}{P_{Nt}A_t}$ is the productivity adjusted real wage and g_t is the growth rate of A_t.

F.2. Production

The non-tradable consumption good is a Dixit-Stiglitz aggregate over a continuum of products $C_{Nt}(i)$ produced by monopolistically competitive producers indexed by i, with

$$C_{Nt} \equiv (\int_0^1 C_{Nt}(i)^{\frac{(\epsilon_p - 1)}{\epsilon_p}} di)^{\frac{\epsilon_p}{(\epsilon_p - 1)}} .$$

Each firm i in home produces a homogenous good with technology given by $Y_{Nt}(i) = (A_t K_t(i))^\alpha L_{Nt}^{1-\alpha}(i)$, taking the demand for its product as given by

$$C_{Nt}(i) = \left(\frac{P_{Nt}(i)}{P_{Nt}}\right)^{-\epsilon_p} C_{Nt} ,$$

[29] We assume imperfect wage indexation in our nominal wage rigidity assumption. We ignore specifying it here for ease of exposition, but specify those in the equilibrium conditions later on.

where we use the price index of the non-tradable good composite, $P_{Nt} = (\int_0^1 P_{Nt}(i)^{1-\epsilon_p} di)^{\frac{1}{(1-\epsilon_p)}}$. Each firm k is assumed to set prices on a staggered basis following Calvo (1983). With probability $(1-\theta_p)$, a firm adjusts its price independent of previous history.

Firms may not be able to adjust their price in a given period, but they will always choose inputs to minimize total cost each period. The cost minimization yields the input demand functions.

$$W_t = (1-\alpha)\ mc_t(i) A_t^{1-\alpha} \left(\frac{K_t(i)}{L_t(i)}\right)^{\alpha}, \qquad R_t^k = \alpha\ mc_t(i) A_t^{1-\alpha} \left(\frac{K_t(i)}{L_t(i)}\right)^{\alpha-1}.$$

The first order condition implies that the capital labor ratio at the firm level is independent on firm-specific variables:

$$\frac{K_t(i)}{L_t(i)} = \frac{K_t}{L_t} = \frac{\alpha}{1-\alpha}\frac{W_t}{R_t^k}.$$

Thus, (nominal) marginal cost is independent of firm specific variables:

$$P_{Nt} mc_t(i) = P_{Nt} mc_t = \frac{1}{Z_t^{1-\alpha}} \left(\frac{R_t^k}{\alpha}\right)^{\alpha} \left(\frac{W_t}{1-\alpha}\right)^{1-\alpha}.$$

For stationarity along the balanced growth path, we assume that the tradable good endowment grows at the rate of A_t.

F.3. Government

The central bank follows a Taylor rule in setting the nominal interest rate i_t. It responds to deviations in inflation, output and output growth rate from time-t natural allocations.

$$\frac{1+i_t}{1+i_{ss}} = \left(\frac{1+i_{t-1}}{1+i_{ss}}\right)^{\rho_R} \left[\left(\frac{\Pi_{Nt}}{\bar{\Pi}_N}\right)^{\phi_\pi} \left(\frac{Y_{Nt}}{Y_{Nt}^{f,t}}\right)^{\phi_y}\right]^{1-\rho_R} \epsilon_t^{mp}, \tag{63}$$

where i_{ss} is the steady state nominal interest rate, Π_{Nt} is gross producer-price inflation rate, $\bar{\Pi}_N$ is the corresponding steady state inflation target, $Y_{Nt}^{f,t}$ is the time-t natural output, ρ_R determines interest-rate smoothing and $\epsilon_t^{mp} \sim N(0, \sigma_r)$ is the monetary policy shock.

We assume net zero supply of nominal bonds at home and that the government balances budget every period. There is no government spending.

F.4. Market clearing

Nontradable goods are consumed, invested and spent in utilization adjustment with

$$Y_{Nt} = C_{Nt} + I_t + a(u_t) K_t^u.$$

The external borrowing constraint must be satisfied with

$$C_{Tt} + D_t = Y_{Tt} + \frac{D_{t+1}}{1+r_t}.$$

F.5. Aggregate GDP

We construct aggregate GDP using constant aggregation weights implied by the utility function along the balanced growth path.

F.6. Stationary allocation

We normalize the following variables:

$$y_{Tt} = Y_{Tt}/A_t; \quad y_{Nt} = Y_{Nt}/A_t$$

$$c_{Tt} = C_{Tt}/A_t; \quad c_{Nt} = C_{Nt}/A_t$$

$$k_t = K_t/A_t; \quad k_t^u = K_t^u/A_{t-1}; \quad \mathbb{I}_t = I_t/A_t \quad \text{(capital investment)}$$

$$w_t = W_t/(A_t P_{Nt}); \quad r_t^k = R_t^k/P_{Nt}$$

$$\lambda_{Tt} = \Lambda_t A_t$$

Table A1: *Parameters*

(a) Medium-scale DSGE parameters

β Discount factor 0.999	δ_k Capital depreciation rate 0.025	α Capital share 0.28	μ Trend growth rate 2%
θ_w Wage Calvo probability 0.75	ϵ_w Elasticity of substitution across labor types 8	θ_p Price Calvo probability 0.75	ϵ_p Elasticity of substitution across non-tradables 8
ν Inverse Frisch elasticity 1.00	h (external) habit 0.50	η Hysteresis elasticity 0.25	$\frac{a''(1)}{a'(1)}$ Capital utilization cost 4
$S''(1)$ Investment adjustment cost 2	ϕ_π Taylor rule inflation response 1.50	ϕ_y Taylor rule (normalized) output response 0.125	ρ_R Persistence coefficient 0.5

(b) Small-open economy parameters

$\bar{d}$ Parameter of debt-elastic interest rate 2.9	ψ Parameter of debt-elastic interest rate 0.0000335	ω Tradable share in expenditure 0.30

Notes: The table shows the parameter values of the model for the baseline calibration. See text.

F.7. Calibration

To provide an illustration in the main text, we choose parameters from medium-scale DSGE literature (Justiniano et al., 2013) and small-open economy literature (Schmitt-Grohé and Uribe, 2016). We list the calibrated parameters in Table A1.

G. Imbs correction

We follow Imbs (1999) and adjust TFP for utilization of capital and labor inputs. See Paul (2020) for a related construction of utilization-adjusted TFP in the historical data. We assume perfectly competitive factor markets and a technology which is constant returns to scale in effective capital and labor. In aggregate, and for the representative firm, the production function is

$$Y_t = A_t \left(K_t u_t\right)^{\alpha} \left(L_t e_t\right)^{1-\alpha} ,$$

where Y_t is output, K_t is capital stock, L_t is total hours worked, and u_t and e_t denote the respective factor utilizations. A_t is the utilization adjusted TFP. We assume perfect competition in the input and the output markets. Higher capital utilization increases the depreciation of capital $\delta_t = \delta u_t^{\phi}$ where $\phi > 1$. As a result, firms choose capital utilization rate optimally. Labor hoarding is calculated assuming instantaneous adjustment of effort e_t against a payment of a higher wage $w(e_t)$, while keeping fixed employment (determined one period in advance). The firm's optimization problem is given by:

$$\max_{e_t, u_t, K_t} A_t \left(K_t u_t\right)^{\alpha} \left(L_t e_t\right)^{1-\alpha} - w(e_t) L_t - (r_t + \delta u_t^{\phi})) K_t .$$

Households choose consumption, labor supply and effort to maximize their lifetime utility subject to their budget constraint (with complete asset markets)

$$\max_{c_t, L_t, e_t} \sum_{t=0}^{\infty} \beta^t \left[\ln C_t - \frac{(L_t)^{1+\nu}}{1+\nu} - \frac{(e_t)^{1+\nu}}{1+\nu} \right] .$$

Normalizing the long-run capital-utilization and labor-utilization rates to one, the utilization rates can be derived from

$$u_t = \left(\frac{Y_t / K_t}{Y / K} \right)^{\frac{\delta}{r+\delta}} ; \quad e_t = \left(\alpha \frac{Y_t}{C_t} \right)^{\frac{1}{1+\nu}} ;$$

where Y, C, L and K are the steady-state values of output, consumption, labor, and capital.

The Solow residual then can be decomposed into utilization-adjusted TFP and utilization corrections, with

$$TFP_t \equiv \frac{Y_t}{K_t^{\alpha} L_t^{1-\alpha}} = A_t \times u_t^{\alpha} e_t^{1-\alpha} .$$

To construct country-specific steady state values of Y/K, we extract a HP-filter trend from the data series.[30] In the utilization adjusted series used in the main text, we set $\alpha = 0.33$, and $\nu = 1$. Results are robust to constructing country specific values of these parameters.

[30] Our empirical results are robust to computing moving averages over a 10 year window, using time-varying values of α constructed from labor-income data, and reasonable parameters of the aggregate capital depreciation rate. Bergeaud, Cette, and Lecat (2016) constructed capital stock for machines and buildings separately using the perpetual inventory method with data on investment in machines and buildings and different depreciation rates. Our results are robust to choosing different depreciation parameters.

H. Mundell-Fleming-Dornbusch model

Although the arbitrage mechanism behind the trilemma is easily grasped, in this section we investigate the economic underpinnings of our identification strategy more formally with a variant of the well known Mundell-Fleming-Dornbusch model. In particular, we incorporate the extensions to the model discussed in Blanchard (2016) and Gourinchas (2018), which embed various financial spillover mechanisms.

Specifically, consider a framework made of two countries: a small domestic economy and a large foreign economy, which we can call the United States, for now. Foreign (U.S.) variables are denoted with an asterisk. Assume prices are fixed.

Given interest rates, the following equations describe the setup:

$$\begin{aligned}
Y &= A + NX\,, \\
A &= \xi - ci - fE\,, \\
NX &= a(Y^* - Y) + bE\,, \\
Y^* &= A^* = \xi^* - c^* i^*\,, \\
E &= d(i^* - i) + gi^* + \chi\,,
\end{aligned}$$

where $a, b, c, c^*, d, f, g, \chi \geq 0$. Domestic output Y is equal to the sum of domestic absorption A and net exports NX. Domestic absorption depends on an aggregate demand shifter ξ, and negatively on the domestic (policy) nominal interest rate i. f denotes financial spillovers through the exchange rate (e.g., balance sheet exposure of domestic producers in a dollarized world).[31] If $f \geq 0$, then a depreciation of the exchange rate E hurts absorption.

Net exports depends positively on U.S. output Y^*, negatively on domestic output Y, and positively on the exchange rate. U.S. output is determined in similar fashion except that the U.S. is considered a large country, so it is treated as a closed economy. Finally, the exchange rate depends on the difference between domestic and U.S. interest rates and on a risk-premium shock. The term g is intended to capture risk-premium effects associated with U.S. monetary policy.[32]

In order to make the connection between the instrument as we defined it earlier and this stylized model, we now think of Δi^* as the instrument $z_{j,t} \equiv k_{j,t}(\Delta i_{b(j,t),t} - \widehat{\Delta i}_{b(j,t),t})$ described earlier. The proposition below explores the benchmark setting of the trilemma to derive the basic intuition.

The textbook specification with hard pegs

Under the assumption that $f = g = \chi = 0$, many interesting channels are switched off and the model just introduced reduces to the textbook Mundell-Fleming-Dornbusch version. Consider what happens when the U.S. changes its interest rate, ΔR^*. Since $g = 0$, to maintain the peg it must be that $\Delta i = \Delta i^*$. The one-to-one change in the home interest rate has a direct effect on domestic absorption given by $-c\Delta i$.

However, notice that changes in the U.S. rate affect U.S. absorption and in turn net exports.

[31] Jiang, Krishnamurthy, and Lustig (2019) provide a micro-foundation to generate these spillovers associated with the global financial cycle (Rey, 2015).

[32] Itskhoki and Mukhin (2019) argue that such risk-premia violations of UIP are smaller under exchange rate pegs, i.e., g is smaller.

Piecing things together:

$$\begin{aligned}\Delta Y &= \Delta A + \Delta NX\,,\\ \Delta Y &= -\frac{c}{1+a}\Delta i - \frac{c^* a}{1+a}\Delta i^*\,.\end{aligned}$$

As is clear from the expression, Δi^* affects domestic output directly (and not just through Δi), resulting in a violation of the exclusion restriction central to instrumental variable estimation. However, note that this violation is easily resolved by including net exports as a control, or even just base country output, something we do later in the estimation. Moreover, in this simple static model, all effects are contemporaneous. However, in practice the feedback loop of higher U.S. interest rates to lower net exports to lower output will take place gradually, in large part alleviating the exclusion restriction violation.

Financial spillovers with soft pegs

Consider now a more general setting with financial spillovers, that is, $g > 0$ and $f > 0$ and a *soft* peg. That is, the central bank may adjust using interest rates and allow some movement of the exchange rate.

This will affect the pass through of U.S. interest rates to domestic rates since now:

$$\Delta i = \frac{1}{d}\Delta\epsilon + \frac{d+g}{d}\Delta i^*\,,$$

where $\Delta E \in \pm\Delta\epsilon$ refers to some band within which the exchange rate is allowed to fluctuate.

The effect on output from changes in U.S. interest rates is very similar, but with an added term:

$$\Delta Y = \frac{c}{1+a}\Delta i - \frac{c^* a}{1+a}\Delta i^* + (b-f)\Delta\epsilon\,.$$

Under a hard peg policy, with $\Delta\epsilon = 0$, an increase in U.S. interest rates boosts home interest rates but it no longer does so one-to-one, as explained earlier. Partial flexibility in exchange rates under a soft peg, with $|\Delta\epsilon| > 0$, gives some further monetary autonomy to the home economy, and reduces the pass-through to home interest rates, all else equal. This additional flexibility in exchange rates, however, results in other financial and trade spillovers due to dependence of domestic absorption and net exports on the exchange rate as shown by the term $(b-f)\Delta\epsilon$.

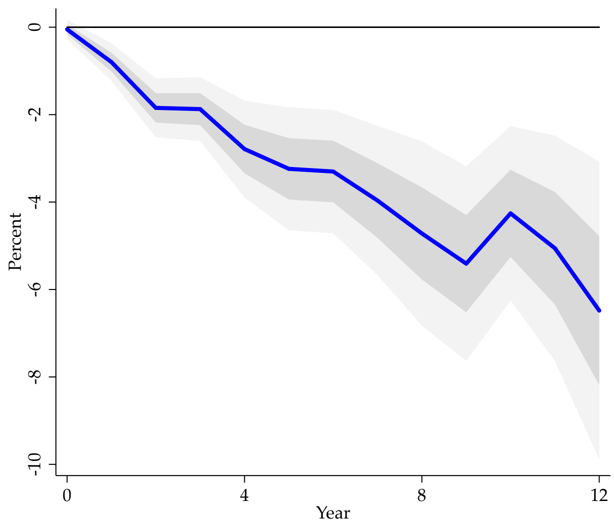
Summarizing our discussion, it is important to recognize that exogenous variation in interest rates (induced either through the trilemma mechanism as just discussed, or through alternative channels) has effects through domestic absorption and through net exports. This secondary channel, if not properly controlled for, generates violations of the exclusion restriction.

I. Additional figures

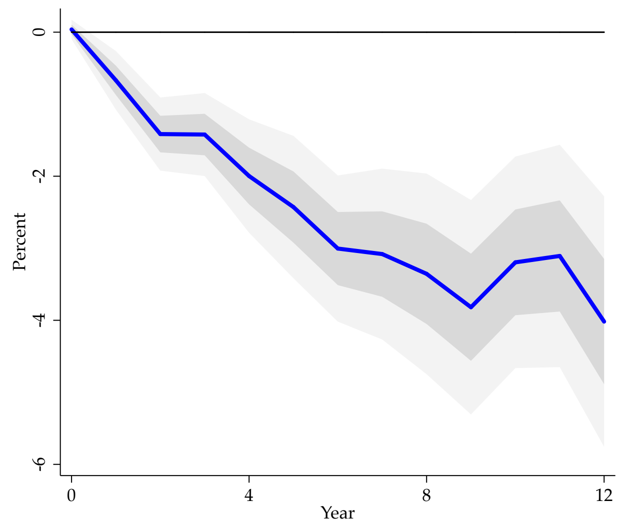
I.1. LPIV responses for GDP per capita

Figure A2: *Baseline response to 100 bps trilemma shock: Real GDP per capita*

(a) *Full sample: 1890–2015*



(b) *Post-WW2 sample: 1948–2015*



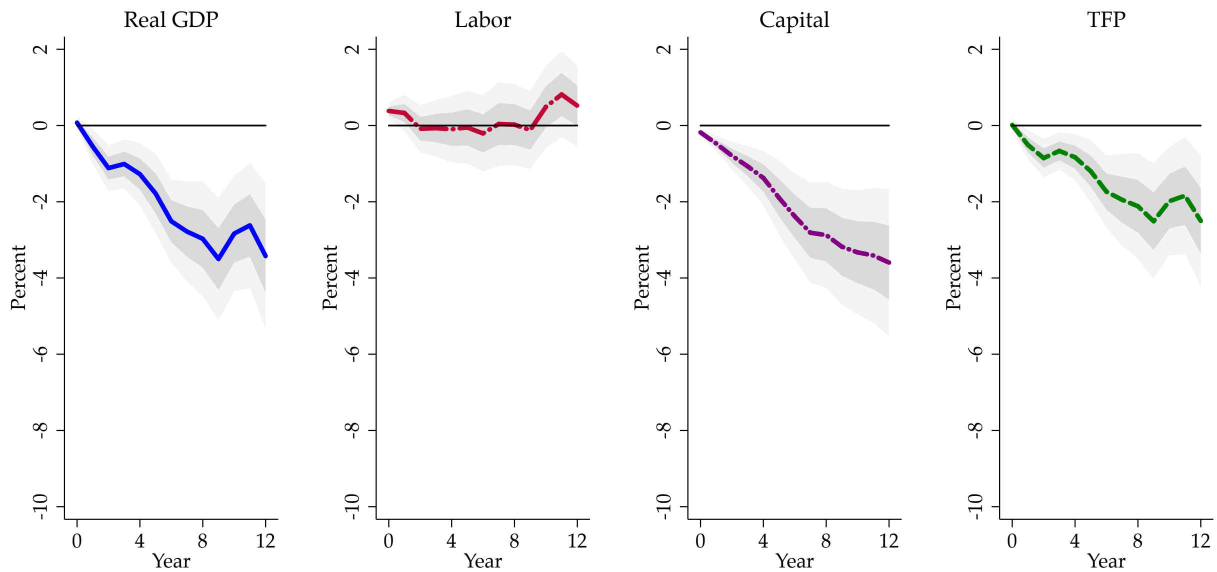
Notes: Response to a 100 bps shock in domestic interest rate instrumented with the trilemma. Responses for pegging economies. Full sample: 1890–2015 (World Wars excluded). Post-WW2 sample: 1948–2015.LP-IV estimates displayed as a solid blue line and and 1 S.E. and 2 S.E. confidence bands constructed using cluster-robust standard errors. See text.

I.2. LPIV responses for components in Post WW2 sample

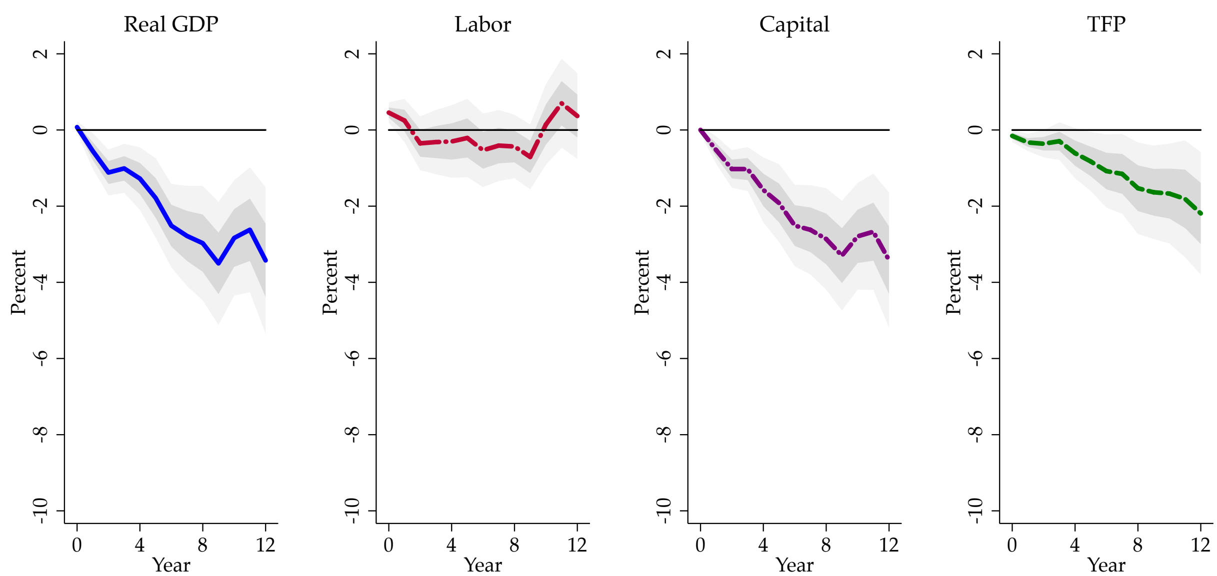
Figure A3 plots the response of components of a Cobb-Douglas production function for the post-WW2 sample.

Figure A3: *Baseline response to 100 bps shock: Real GDP and Solow decomposition. Post ww2 sample, 1948–2015.*

(a) *Estimates using raw data*



(b) *Estimates using Imbs correction for factor utilization*

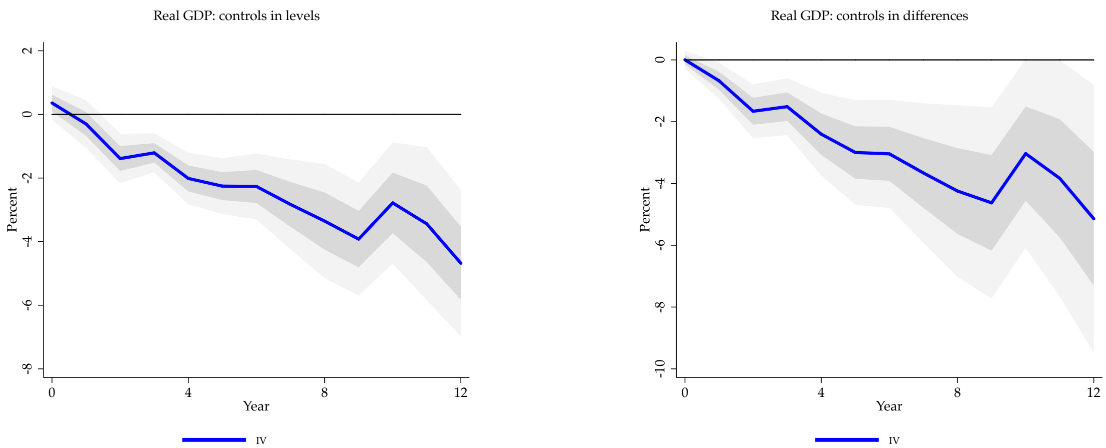


Notes: Response to a 100 bps shock in domestic short-term interest rate instrumented with the trilemma IVs. Post-WW2 sample: 1948–2015. LP-IV estimates displayed as a thick lines and and 1 S.E. and 2 S.E. confidence bands. The upper panel uses raw data on capital stocks and total hours to construct TFP as a residual. The lower panel adjusts the raw data on capital stock and total hours to obtain estimates of actual factor inputs by using the Imbs (1999) correction. See text.

I.3. LPIV responses: controls in levels, differences, and number of lags

We report the robustness of IRFs estimated in the baseline to adding the control variables $\boldsymbol{x}_{j,t}$ in levels instead of first differences in the left panel of Figure A4, as well as to including up to 5 lags of the control variables in the right panel.

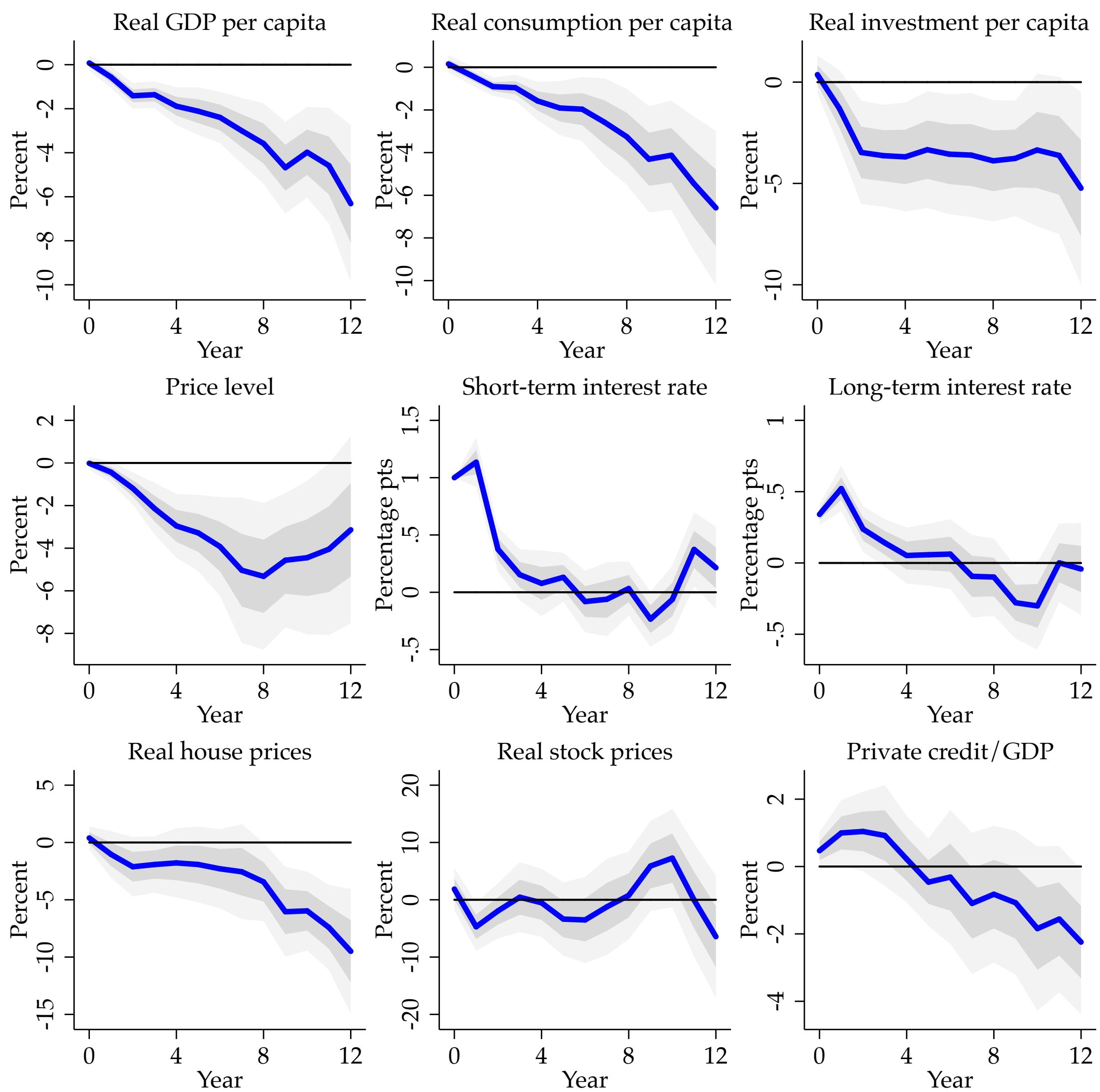
Figure A4: *Response to 100 bps trilemma shock with controls in levels: Real GDP*



Notes: Response to a 100 bps shock in domestic interest rate instrumented with the trilemma. Responses for pegging economies. Full sample: 1890–2015 (World Wars excluded). LP-IV estimates displayed as a solid blue line and 1 S.E. and 2 S.E. confidence bands. See text.

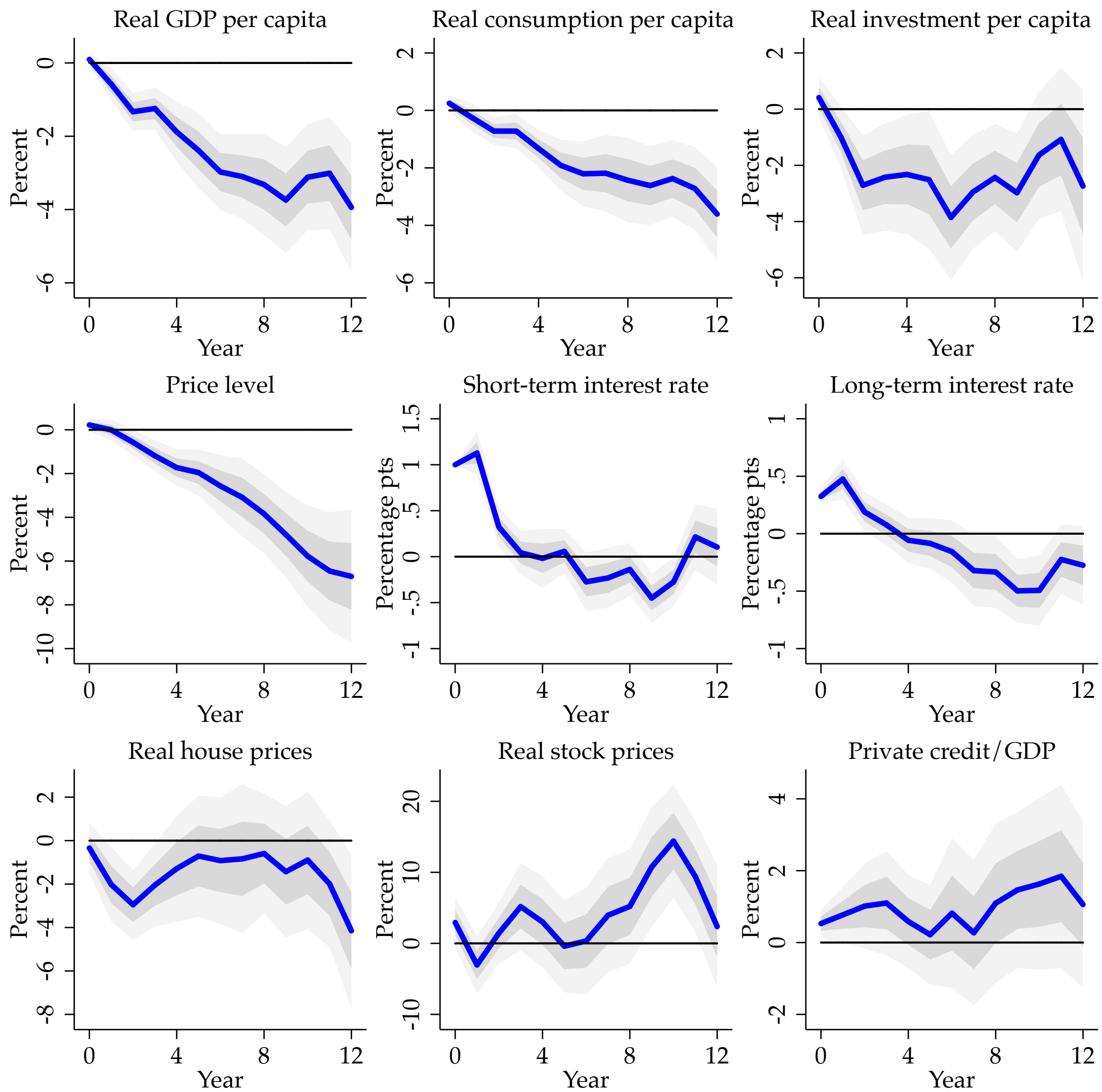
I.4. LPIV responses for various macro-financial variables

Figure A5: *Baseline response to 100 bps trilemma shock: Full Sample (1890-2015)*



Notes: Response to a 100 bps shock in domestic interest rate instrumented with the trilemma. Responses for pegging economies. Full sample: 1890–2015 (World Wars excluded). Confidence bands are one and two standard errors using cluster-robust standard errors. See text.

Figure A6: *Baseline response to 100 bps trilemma shock: Post-WW2 Sample (1948-2015)*

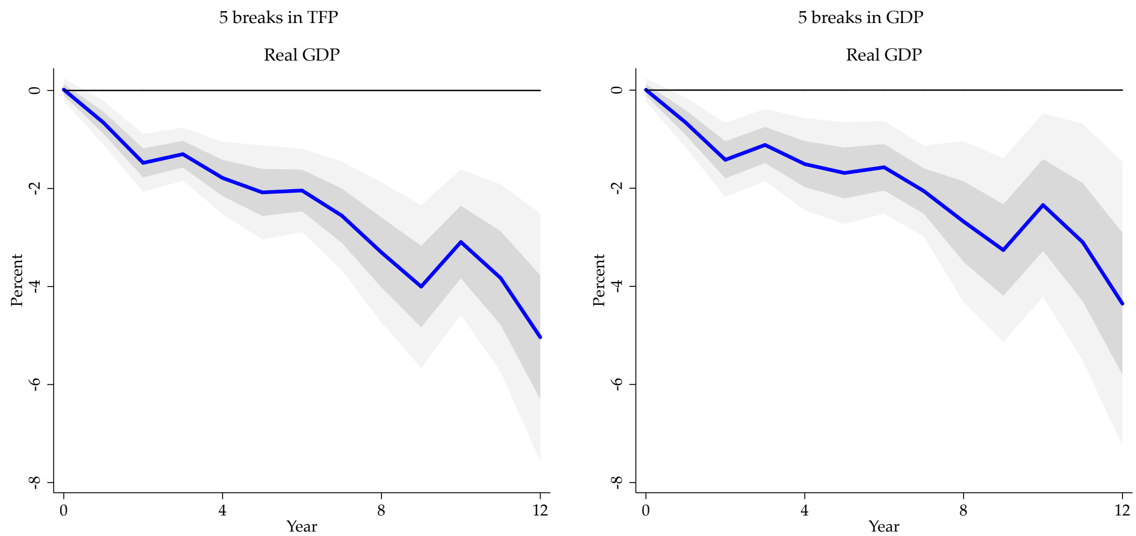


Notes: Response to a 100 bps shock in domestic interest rate instrumented with the trilemma. Responses for pegging economies. PostWW2 sample: 1948–2015 (World Wars excluded). Confidence bands are one and two standard errors using cluster-robust standard errors. See text.

I.5. LPIV responses with structural breaks

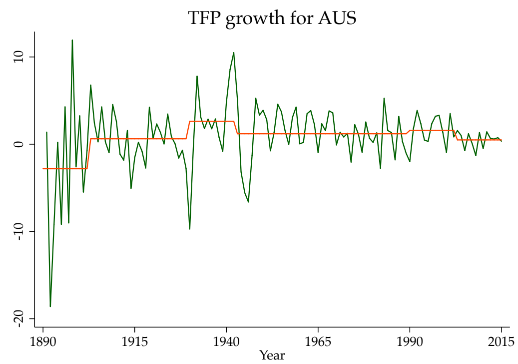
Figure A7: *Response to 100 bps trilemma shock with structural breaks: Real GDP*

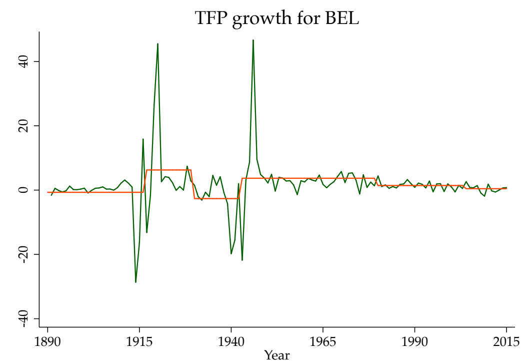
(a) **(b)**

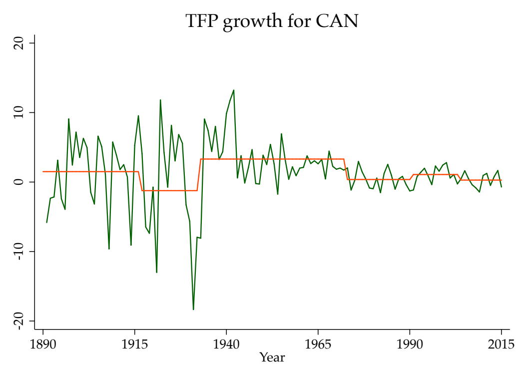


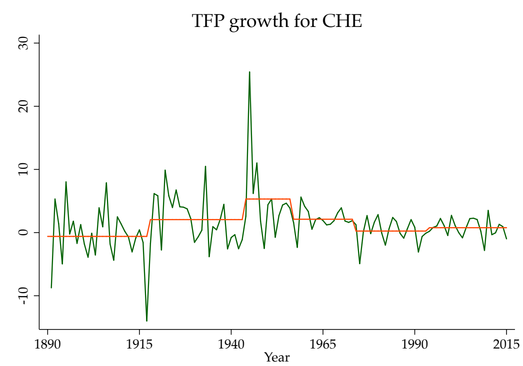
Notes: Response to a 100 bps shock in domestic interest rate instrumented with the trilemma. Responses for pegging economies. Full sample: 1890–2015 (World Wars excluded). LP-IV estimates displayed as a solid blue line and 1 S.E. and 2 S.E. confidence bands. See text.

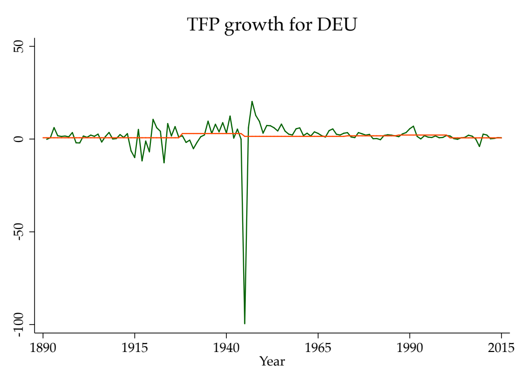
I.6. Structural break dates in TFP growth and GDP growth

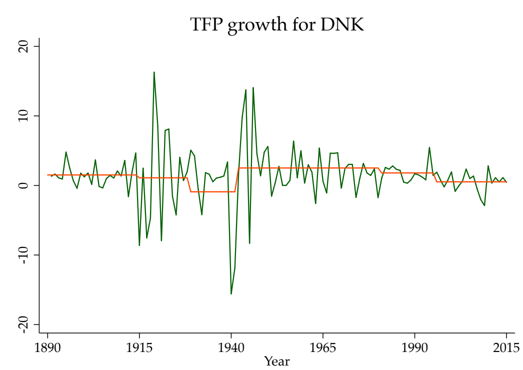


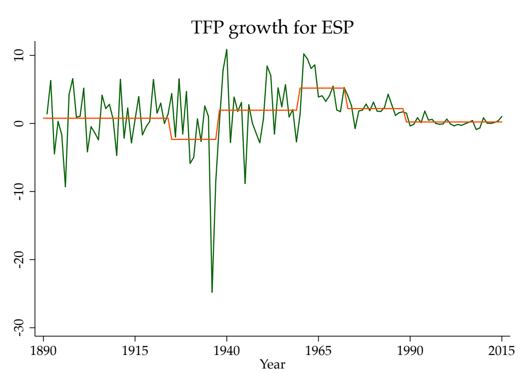


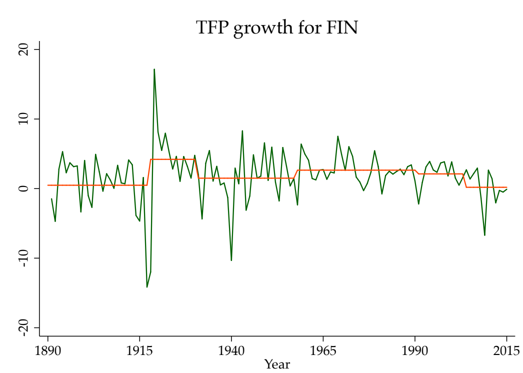




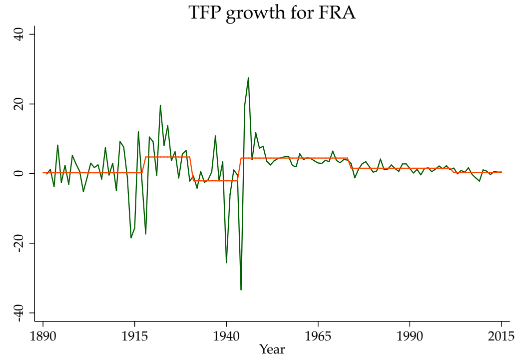




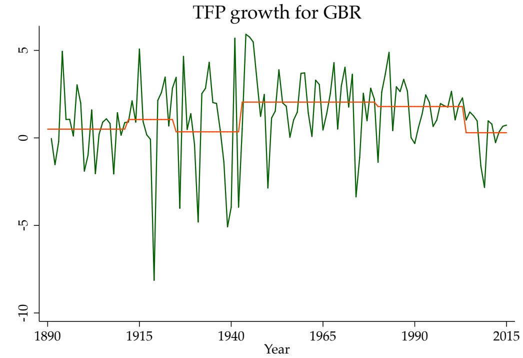




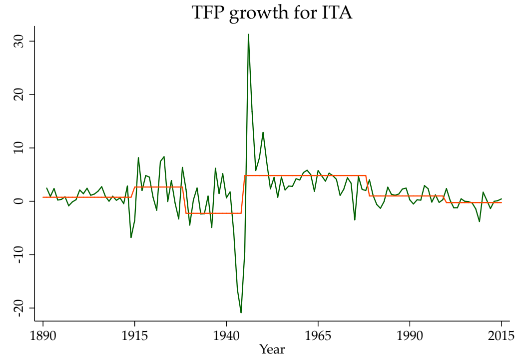
TFP growth for FRA



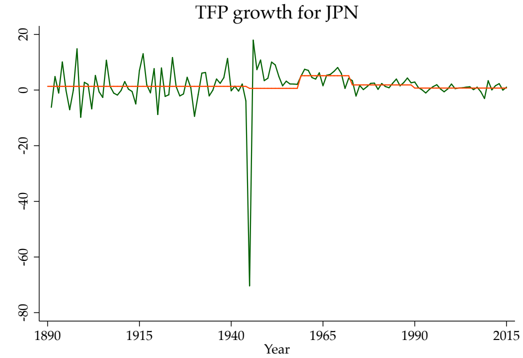
TFP growth for GBR



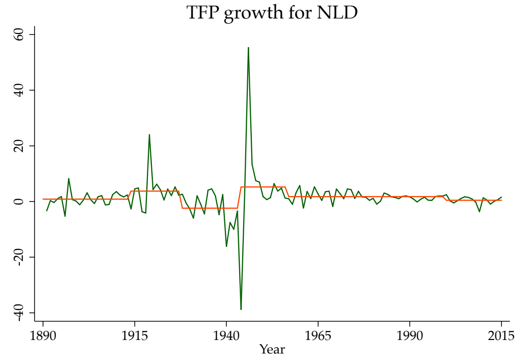
TFP growth for ITA



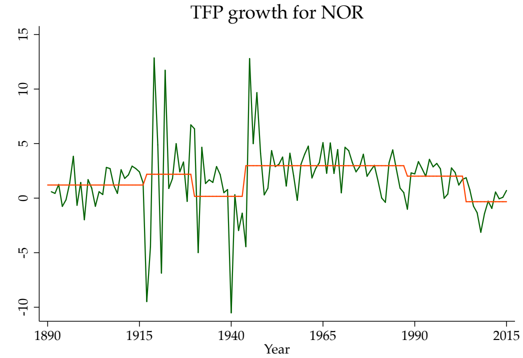
TFP growth for JPN



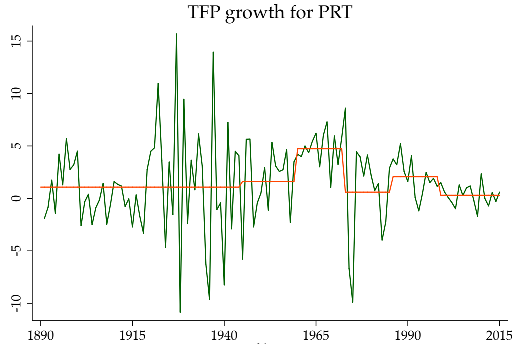
TFP growth for NLD



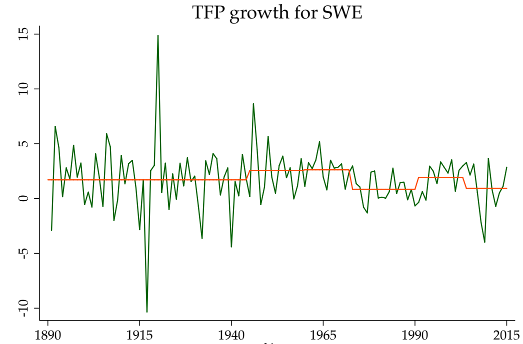
TFP growth for NOR

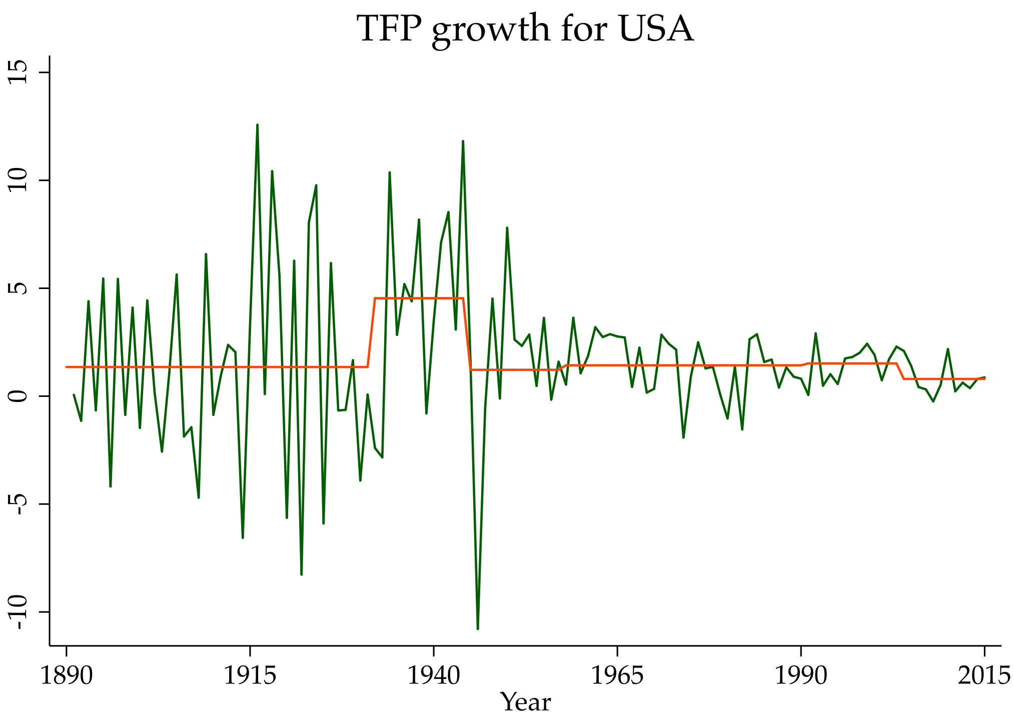


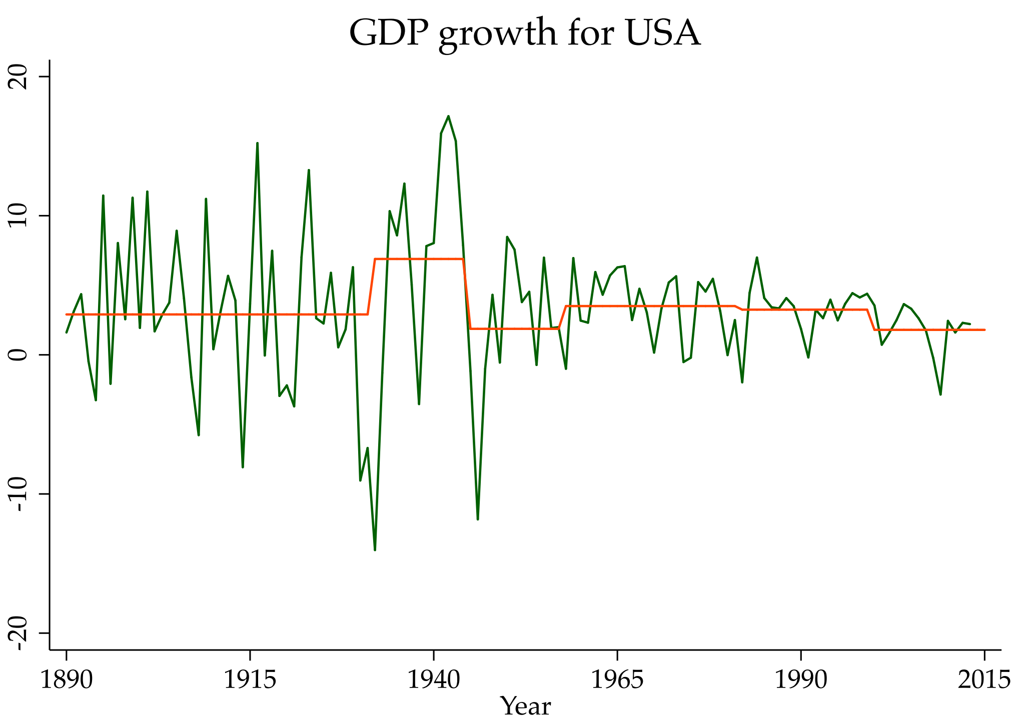
TFP growth for PRT

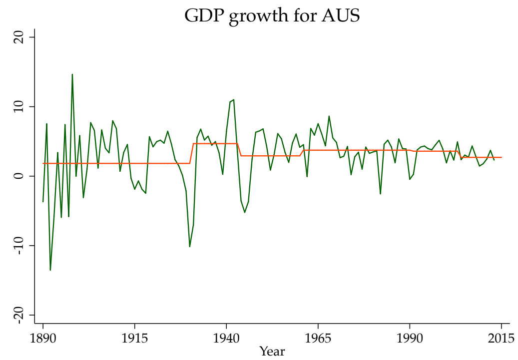


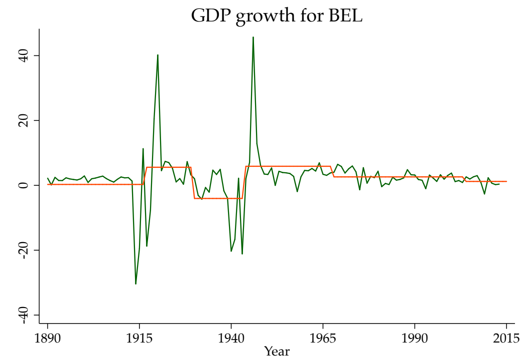
TFP growth for SWE

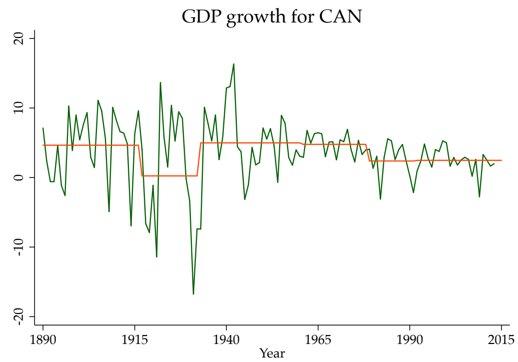


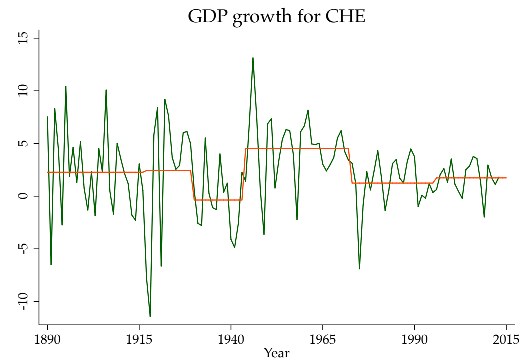

TFP growth for USA
15
10
5
0
-5
-10
1890
1915
1940
1965
1990
2015
Year

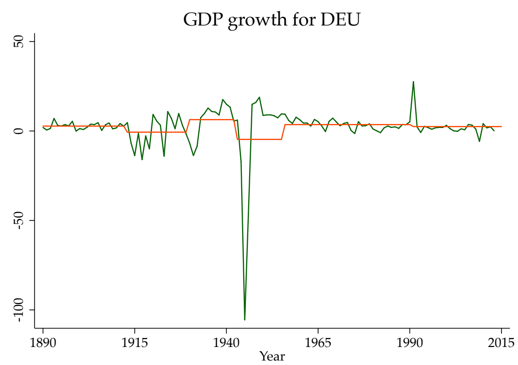

GDP growth for USA
20
10
0
-10
-20
1890
1915
1940
1965
1990
2015
Year

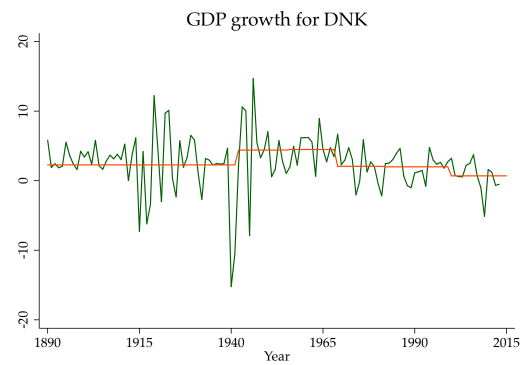

GDP growth for AUS
Year

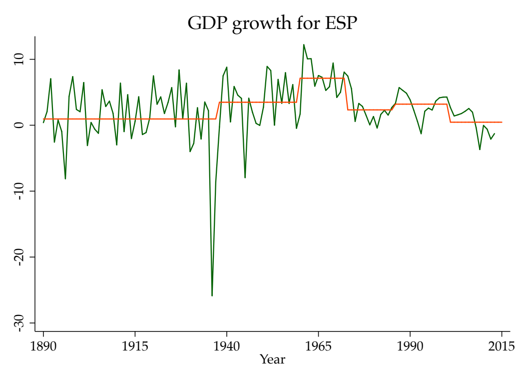

GDP growth for BEL
Year

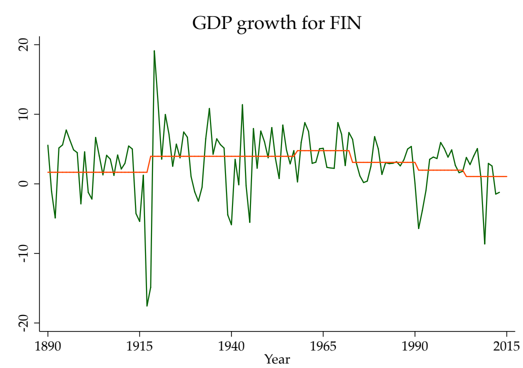

GDP growth for CAN
Year

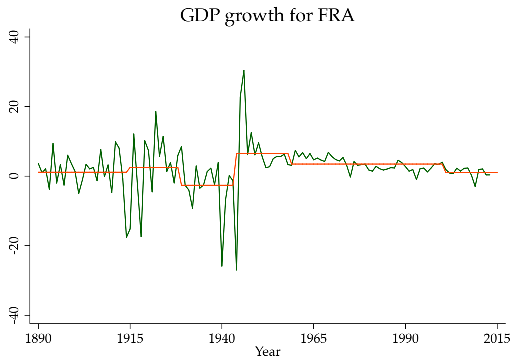

GDP growth for CHE
Year

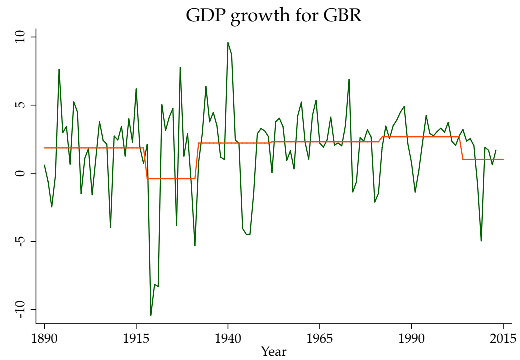

GDP growth for DEU
Year

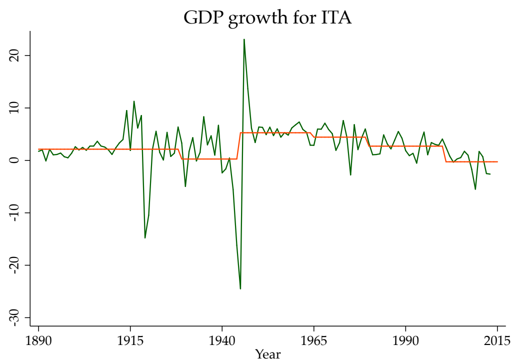

GDP growth for DNK
Year

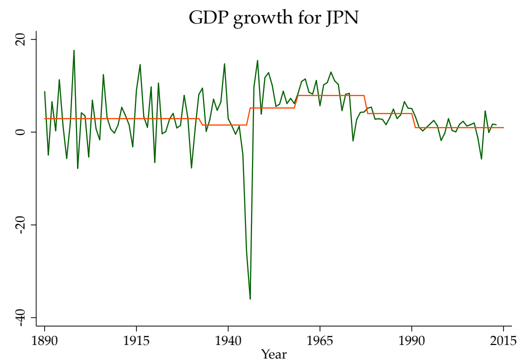

GDP growth for ESP
Year

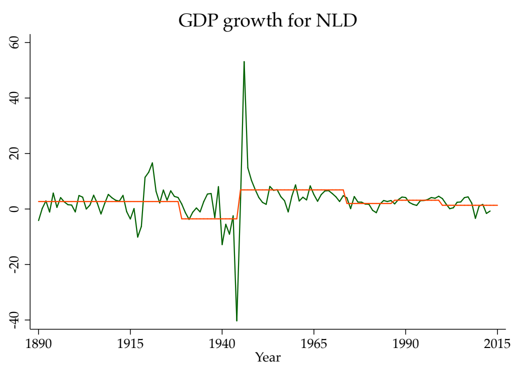

GDP growth for FIN
Year

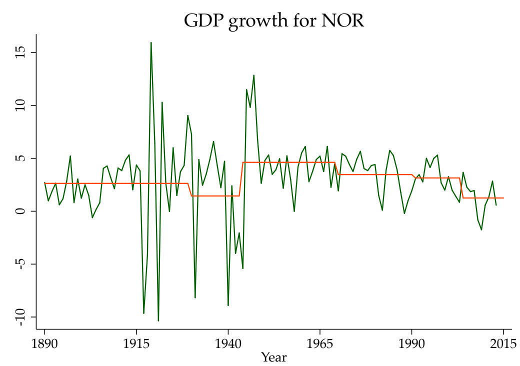

GDP growth for FRA
40
20
0
-20
-40
1890
1915
1940
1965
1990
2015
Year

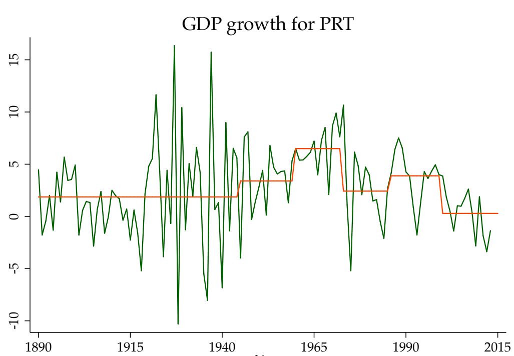

GDP growth for GBR
10
5
0
-5
-10
1890
1915
1940
1965
1990
2015
Year

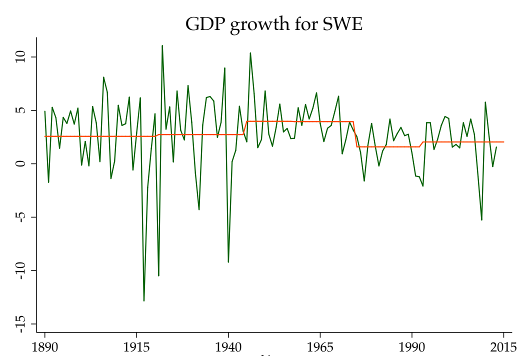

GDP growth for ITA
20
10
0
-10
-20
-30
1890
1915
1940
1965
1990
2015
Year


GDP growth for JPN
20
0
-20
-40
1890
1915
1940
1965
1990
2015
Year


GDP growth for NLD
60
40
20
0
-20
-40
1890
1915
1940
1965
1990
2015
Year


GDP growth for NOR
15
10
5
0
-5
-10
1890
1915
1940
1965
1990
2015
Year


GDP growth for PRT
15
10
5
0
-5
-10
1890
1915
1940
1965
1990
2015
Year


GDP growth for SWE
10
5
0
-5
-10
-15
1890
1915
1940
1965
1990
2015
Year

J. Additional tables

Table A2: *LP-OLS vs. LP-IV. Attenuation bias of real GDP per capita responses to interest rates. Trilemma instrument.*

Responses of real GDP per capita at years 0 to 10 (100 × log change from year 0 baseline).

	(a) Full Sample		*OLS-IV*	*(b) Post-WW2*		*OLS-IV*
Year	LP-OLS	LP-IV	*p-value*	LP-OLS	LP-IV	*p-value*
	(1)	(2)	(3)	(4)	(5)	(6)
$h = 0$	0.05	-0.02	0.52	0.02	0.04	0.76
	(0.03)	(0.11)		(0.02)	(0.07)	
$h = 2$	-0.35**	-1.88***	0.00	-0.37**	-1.41***	0.00
	(0.14)	(0.36)		(0.14)	(0.25)	
$h = 4$	-0.32	-2.73***	0.00	-0.35*	-2.00***	0.00
	(0.22)	(0.53)		(0.21)	(0.39)	
$h = 6$	-0.45	-3.36***	0.00	-0.28	-3.00***	0.00
	(0.37)	(0.70)		(0.31)	(0.51)	
$h = 8$	-0.63*	-4.90***	0.00	-0.27	-3.36***	0.00
	(0.35)	(1.10)		(0.31)	(0.70)	
$h = 10$	-0.62*	-4.40***	0.00	0.06	-3.20***	0.00
	(0.35)	(1.02)		(0.31)	(0.73)	
$h = 12$	-0.62	-6.50***	0.00	0.04	-4.02***	0.00
	(0.40)	(1.68)		(0.36)	(0.87)	
KP weak IV		47.54			62.43	
H_0: $LATE = 0$	0.00	0.00		0.00	0.00	
Observations	963	774		710	585	

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Cluster robust standard errors in parentheses.Full sample: $1890 - 2015$ excluding WW1: $1914 - 1919$ and WW2: $1939 - 1947$. Post WW2 sample: $1948 - 2015$. *Matched sample* indicates LP-OLS sample matches the sample used to obtain LP-IV estimates. *KP weak IV* refers to the Kleibergen-Paap test for weak instruments. H_0: $LATE = 0$ refers to the p-value of the test of the null hypothesis that the coefficients for $h = 0, ..., 10$ are jointly zero for a given subpopulation. $OLS = IV$ shows the p-value for the Hausmann test of the null that OLS estimates equal IV estimates. See text.

Appendix references

Bergeaud, Antonin, Gilbert Cette, and Rémy Lecat. 2016. Productivity Trends in Advanced Countries between 1890 and 2012. *Review of Income and Wealth* 62(3): 420–444.

Blanchard, Olivier. 2016. Currency Wars, Coordination, and Capital Controls. Working Paper 22388, National Bureau of Economic Research.

Calvo, Guillermo A. 1983. Staggered prices in a utility-maximizing framework. *Journal of Monetary Economics* 12(3): 383–398.

Christiano, Lawrence J., Martin Eichenbaum, and Charles L. Evans. 2005. Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy. *Journal of Political Economy* 113(1): 1–45.

Durbin, J. 1959. Efficient Estimation of Parameters in Moving-Average Models. *Biometrika* 46(3/4): 306–316.

Erceg, Christopher J., Dale W. Henderson, and Andrew T. Levin. 2000. Optimal monetary policy with staggered wage and price contracts. *Journal of Monetary Economics* 46(2): 281–313.

Gourinchas, Pierre-Olivier. 2018. Monetary Policy Transmission in Emerging Markets: An Application to Chile. In *Monetary Policy and Global Spillovers: Mechanisms, Effects and Policy Measures*, edited by Mendoza, Enrique G., Ernesto Pastén, and Diego Saravia, volume 25 of *Central Banking, Analysis, and Economic Policies Book Series*, chapter 8, 279–324. Santiago: Central Bank of Chile.

Imbs, Jean M. 1999. Technology, growth and the business cycle. *Journal of Monetary Economics* 44(1): 65–80.

Itskhoki, Oleg, and Dmitry Mukhin. 2019. Mussa puzzle redux. Princeton University. Unpublished.

Jiang, Zhengyang, Arvind Krishnamurthy, and Hanno Lustig. 2019. Dollar Safety and the Global Financial Cycle. Research Paper 19-16, Stanford University Graduate School of Business.

Justiniano, Alejandro, Giorgio E. Primiceri, and Andrea Tambalotti. 2013. Is There a Trade-Off between Inflation and Output Stabilization? *American Economic Journal: Macroeconomics* 5(2): 1–31.

Kuersteiner, Guido M. 2005. Automatic Inference For Infinite Order Vector Autoregressions. *Econometric Theory* 21(1): 85–115.

Lewis, Richard, and Gregory C. Reinsel. 1985. Prediction of multivariate time series by autoregressive model fitting. *Journal of Multivariate Analysis* 16(3): 393–411.

Lusompa, Amaze B. 2019. Local projections, autocorrelation, and efficiency. Unpublished. UC Irvine.

Paul, Pascal. 2020. Historical patterns of inequality and productivity around financial crises. Federal Reserve Bank of San Francisco Working Paper 2017-23.

Rey, Hélène. 2015. Dilemma not Trilemma: The Global Financial Cycle and Monetary Policy Independence. NBER Working Papers 21162, National Bureau of Economic Research, Inc.

Schmitt-Grohé, Stephanie, and Martın Uribe. 2003. Closing small open economy models. *Journal of International Economics* 61(1): 163–185.

Schmitt-Grohé, Stephanie, and Martin Uribe. 2016. Downward Nominal Wage Rigidity, Currency. Pegs, and Involuntary Unemployment. *Journal of Political Economy* 124(5): 1466–1514.

Smets, Frank, and Rafael Wouters. 2007. Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach. *American Economic Review* 97(3): 586–606.

Uribe, Martin, and Stephanie Schmitt-Grohé. 2017. *Open Economy Macroeconomics*. Princeton, N.J.: Princeton University Press.

Wiener, Norbert, and Pesi Masani. 1958. The prediction theory of multivariate stochastic processes, II: The linear predictor. *Acta Mathematica* 99(1): 93–137.