The long-run effects of monetary policy*

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

January 2020

Abstract

Is the effect of monetary policy on the productive capacity of the economy long lived? Yes, in fact we find such impacts are significant and last for over a decade based on: (1) merged data from two new international historical databases; (2) identification of exogenous monetary policy using the macroeconomic trilemma; and (3) improved econometric methods. Notably, the capital stock and total factor productivity (TFP) exhibit hysteresis, but labor does not. Money is non-neutral for a much longer period of time than is customarily assumed. A New Keynesian model with endogenous TFP growth can reconcile all these empirical observations.

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, 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.

†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? At least since 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). Lately, however, many have begun to question this monetary canon. 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 interest rates to the *level* of productivity, whereas more recently, Benigno and Fornaro (2018) link low interest rates with the rate of *growth* of productivity.

We argue that any decisive investigation of monetary neutrality must rest on three pillars. First, we are looking at long-run outcomes and so we need a sample based on longer time series data and, if possible, on a wide panel of countries to obtain more statistical power. Second, it is essential to identify exogenous movements in the interest rate to obtain a reliable measure of monetary effects and to avoid confounding. Third, as we will show below, the empirical method used can make a big difference. Common approaches are designed to maximize short-horizon fit. It is therefore important to use methods that are instead better suited to characterize behavior over longer spans of time. We discuss how we build on each of these three pillars next.

Starting first with 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 database contains major macroeconomic aggregates, such as output, interest rates, credit, and so on. Second, in order to decompose output into its components, we bring in the data from Bergeaud, Cette, and Lecat (2016) and available at `http://www.longtermproductivity.com`.[1] Their data include observations on investment in machines and buildings, number of employees, and hours worked. Using these variables, we can then construct measures of total factor productivity (TFP), as we show later.

With regard to identification, the second pillar, we exploit the trilemma of international finance (see, e.g., Obstfeld, Shambaugh, and Taylor, 2004, 2005; Shambaugh, 2004). The

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

basic idea is that when a country pegs its currency to some base currency, but allows free movement of capital across its borders, it effectively loses full control over its own domestic interest rate: a correlation in home and base interest rates is induced, which is one when the peg is hard and mobility is perfect, but generally less than one otherwise. Suitably constructed shocks to the base interest rate are then exogeous to the home country (Jordà, Schularick, and Taylor, 2016, 2019). Of course, this mechanism depends on a variety of other factors that we examine through the lens of an extended, stylized Mundell-Fleming framework (see, e.g., Gourinchas, 2018). This extension allows for contagion via financial linkages, deviations from uncovered interest rate parity, and partial control over monetary policy through small exchange rate interventions. Using this framework, we examine a rich array of spillover mechanisms that could threaten the exclusion restriction typical of instrumental variable assumptions. By making appropriate adjustments based on these economic mechanisms and novel econometric methods, we show that our results are robust to potential violations of the exclusion restriction (Conley, Hansen, and Rossi, 2012; van Kippersluis and Rietveld, 2018).

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).

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 even twelve years out it has not returned to its pre-shock trend. 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 quickly to the original trend. This finding is distinct from the usual labor hysteresis mechanism 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.

Before providing a framework to make sense of these results, we first discuss the robustness of our findings. In particular, there are obvious differences in the monetary arrangements before and after WW2. Think of the pre-WW1 Gold Standard versus the Bretton Woods era, for example. Moreover, the Global Financial Crisis, second only to the

Great Depression, could be affecting our findings unduly. Thus we examine several sample splits of the data. Fernald (2007), Bergeaud, Cette, and Lecat (2016) and Gordon (2016) document trend breaks in TFP for a variety of countries in our sample. These breaks could affect our estimates and therefore we re-estimate our models using breaks determined with the structural break testing procedure by Bai and Perron (1998).

In any instrumental variable analysis, it is important to assess the exclusion restriction. Absent overidentifying restrictions to formally test the validity of the instrument, we resort to a variety of alternative approaches. First, we allow for spillovers motivated by our Mundell-Fleming framework and captured by fluctuations in base-country as well as global GDP, current account balances, and nominal exchange rates. In particular, we control for past, current, and future values of these variables. Including future values of variables is unusual, but is meant to stack the odds against our findings. We also account for unknown spillovers, following Conley, Hansen, and Rossi (2012) and van Kippersluis and Rietveld (2018), by exploiting different subpopulations present in our sample.

There have been numerous influential papers written based on post-WW2 U.S. data, which naturally form a useful benchmark for our analysis (see, e.g., Ramey, 2016; Nakamura and Steinsson, 2018). In particular, we examine whether monetary shocks identified as in Romer and Romer (2004) produce the same results when applied to the data in Fernald (2014) on utilization adjusted TFP, capital, and total hours worked. We find similar results to those based on our sample and identification approach, with some noteworthy differences that we shall discuss. This exhaustive list of robustness checks give us confidence in our findings.

In order to make economic sense of our findings, we extend a workhorse quantitative macroeconomic model (Christiano, Eichenbaum, and Evans, 2005; Smets and Wouters, 2007) with endogenous hysteresis effects. We choose a parametrically convenient process for endogenous TFP growth that borrows from the formulation in Stadler (1990). We estimate the hysteresis elasticity directly from the cumulative responses of TFP and output using a two-step, classical minimum distance procedure. Our estimated elasticity is similar to that assumed in DeLong and Summers (2012) and comfortably falls within a tight probability interval.

In our model, a contractionary monetary policy shock lowers output temporarily producing a slowdown in TFP growth. Under a standard Taylor rule, this slowdown in TFP growth accumulates to yield permanently lower trend levels of output and capital, while labor returns to the stationary equilibrium quickly. The model can generate empirically consistent medium-run effects on GDP while maintaining other conventional textbook results in monetary economics (as in, e.g., Galí, 2015b).

These hysteresis effects naturally matter for how we build models of monetary economies and design optimal monetary policy. A recent literature provides micro-foundations for endogenous hysteresis effects via TFP growth (Fatas, 2000; Barlevy, 2004; Anzoategui, Comin, Gertler, and Martinez, 2019; Bianchi, Kung, and Morales, 2019). Our findings and framework incorporate a natural reduced-form representation of these mechanisms.[2]

Whilst we confirm the reduced-form implications of these models, a fuller investigation into the empirical basis for appropriate micro-foundations would require an entirely new paper, and most likely additional micro evidence, that we must leave for future research. For example, the design of an optimal rule would depend on the type of frictions assumed. In an environment where the only friction comes from nominal rigidities, a strict inflation-targeting rule, for example, could deliver optimal allocations. As an illustration of how one might modify the policy rule in our environment, we show that a hysteresis-augmented policy rule generates a milder dip and quicker recovery in output at a modest cost to inflation stabilization.

Our paper is related to the seminal work by Cerra and Saxena (2008) who documented that output losses in the aftermath of economic and political crises in low-income and emerging economies are highly persistent. More recently, researchers have also documented similar effects following financial crises for advanced economies (see e.g., Reinhart and Rogoff, 2009; Schularick and Taylor, 2012). Fatás and Summers (2018) document strong hysteresis effects of fiscal consolidations for advanced economies following the global financial crisis.

Over the much longer run, we document a causal effect of monetary policy shocks in a historical panel data for advanced economies. There is mixed evidence of long-run non-neutrality proposition in monetary economics as already mentioned above (King and Watson, 1997). In the structural VAR literature, Bernanke and Mihov (1998) fail to reject long-run neutrality in their identified responses to monetary policy shocks, which Mankiw (2001) interprets as evidence of long-run non-neutrality due to the imprecision of the estimates. More recently, Moran and Queraltó (2018) use three-equation VAR models to emphasize an empirical link between TFP growth and monetary policy shocks. The persistent effects of monetary policy bear implications for measurement of potential output (see, e.g., Reifschneider, Wascher, and Wilcox, 2015; Coibion, Gorodnichenko, and Ulate, 2017). Relatedly, these echo the important early findings by Evans (1992) who cast doubt on the exogeneity of TFP shocks, a key tenet of all mainstream real-business cycle and dynamic stochastic general equilibrium (DSGE) models.

[2]For an alternate modeling of hysteresis, we refer the reader to Farmer and Nicolò (2018); Farmer and Platonov (2019) for models of unemployment hysteresis.

2. Data and series construction

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 (2019), 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 (2019), 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 the appendix in Section B.[3]

3. Identification

The trilemma of international finance gives a theoretically justified source of exogenous variation in interest rates for open pegs, as Jordà, Schularick, and Taylor (2019) showed. The logic is straightforward: under a peg, short-term rates in two countries will be arbitraged. With *strict* interest parity, rates are exactly correlated and arbitrage is perfect, but this is *not* necessary for our approach. Even with frictions or imperfect arbitrage, a non-zero interest rate correlation is induced which is enough for identification purposes, as we show.

Formally, consider a setting with two countries: a large foreign economy that permits its

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

exchange rate to float, and a smaller home economy that pegs its exchange rate to the large, base economy. Under some conditions—to be made explicit momentarily—the interest rate in the base provides a source of exogenous variation, i.e., an instrumental variable, for the interest rate at home.

We divide the 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, and will not peg to any base.

Base and peg country codings can be found in Jordà, Schularick, and Taylor (2019, 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 j is defined to be in a peg at time t and denoted with the dummy variable $q_{j,t} = 1$ if it maintained a peg to its base at dates $t-1$ *and* t. (This conservative definition serves to eliminate the possibility of opportunistic pegging, and 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.[4])

Based on this discussion, we construct the instrument as follows. Let $\Delta i_{j,t}$ denote changes in country j's short-term nominal interest rate at date t, let $\Delta i_{b(j,t),t}$ denote the change in short term interest rate of country j's base country $b(j,t)$, and $\widehat{\Delta} i_{b(j,t),t}$ denote the 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.[5] 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) and denoted $k_{j,t} \in [0,1]$.

The resulting trilemma instrument is therefore defined as $z_{j,t} \equiv k_{j,t}(\Delta i_{b(j,t),t} - \widehat{\Delta} i_{b(j,t),t})$. Identification thus works for the subpopulation with $q_{j,t} = 1$, the pegs. But the subpopulation with $q_{j,t} = 0$, the floats, can also provide useful information to assess the exclusion restriction using the methods described in Conley, Hansen, and Rossi (2012) and van Kippersluis and Rietveld (2018), as we show below.

[4]In the full sample, the index k 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 (2019) for further details on the construction of the instrument.

[5]The list of controls used to construct $\widehat{\Delta} i_{b(j,t),t}$ 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 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.

3.1. Justifying the instrument with an open economy 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. Intuition is best communicated in a static setting, which can be micro-founded with pre-fixed prices and wages as in Obstfeld and Rogoff (1995). However, the appendix (see Section C) extends the model to a dynamic setting with rational expectations as in Gali and Monacelli (2005) to provide more formal justification.

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).[6] 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.[7]

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

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

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 next section 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. 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.

Importantly, because most studies of monetary economies take a closed economy perspective, these international channels are not factored in. For example, this would be true for the commonly used Romer and Romer (2004) shocks.[8] This is quite possibly an under-appreciated source of bias when estimating the effects of monetary policy. Below we develop methods to quantify and correct for this sort of bias.

4. Measuring long-run treatment effects in finite samples

The statistical approach used to measure the propagation of a monetary experiment over long horizons can make a big difference in the impulse response estimates obtained in a finite sample. The impulse response representation of the data is the infinite moving average representation or $MA(\infty)$. Thus, it is natural to make this assumption in thinking about the data generating process since it relates directly to the object of interest. If in addition the $MA(\infty)$ is invertible, it will have a corresponding infinite vector autoregressive representation or $VAR(\infty)$.

[8] In closed economy identification of monetary shocks (for e.g. Romer and Romer, 2004), expectations of path of economic variables is factored in through Fed staff's forecasts. While we do not use similar forecasts, we believe our identification is less likely to be affected by concerns of systematic changes in base country rates. One, we control for base country GDP growth and inflation rate when constructing our instrument. Two, we also control for future path of base country GDP growth when estimating the impulse responses. Three, at a broader level, it is unlikely that a base country systematically changes interest rate taking a peg's economic conditions into account.

In finite samples, there are two approaches to estimate impulse responses. One is to truncate the $VAR(\infty)$ to $VAR(p)$ and derive impulse responses from the truncated model. The other is to estimate impulse responses with local projections with a similar lag truncation. The estimates at long horizons turn out to be more precisely estimated with local projections, as we now explain.

Truncated local projections and a $VAR(p)$ can be thought of as members of the class of semi-parametric sieve estimators (see, e.g., Grenander, 1981; Geman and Hwang, 1982; Bühlmann, 1995). The thought experiment to justify the asymptotic analysis consists of letting the truncation lag in both approaches to increase with the sample size at the same rate, to be made explicit below.

Briefly, we show that impulse responses estimated with finite lag VARs are guaranteed to be consistent up to the truncation lag, under some assumptions. Unless the truncated VAR represents the data generating process (DGP) itself, this results in a quantifiable and potentially large source of bias for horizons beyond the truncation lag. In contrast, local projections do not have this problem.

The rest of the section provides formal arguments justifying this and other statements. Many derivations are relegated to the Appendix for expedient reading. In addition, we provide a simple Monte Carlo experiment to illustrate the differences between these two methods more tangibly.

4.1. Setup

Our top level starting point is to assume that an m-dimensional vector valued process $\{\boldsymbol{y}_t\}$ is generated by the MA(∞) given by

$$\boldsymbol{y}_t = \sum_{j=0}^{\infty} B_0 \boldsymbol{\epsilon}_{t-j}\,, \tag{1}$$

where $\boldsymbol{y}_t = (y_{1t}, \ldots, y_{mt})'$, $B_0 = I_m$, and $\{\boldsymbol{\epsilon}_t\}$ are a sequence of i.i.d. random vectors with mean zero and positive definite covariance matrix Σ.[9] We omit deterministic terms (such as the constant and time trends) for convenience, but without loss of generality. Setting aside identification to simplify the presentation of the main ideas, note that Equation 1 is the impulse response representation of $\{\boldsymbol{y}_t\}$.

Let $||B_j||^2 = tr(B_j' B_j)$, i.e., the Froebenius norm. Further assume that $\sum_{j=0}^{\infty} ||B_j|| < \infty$, and $\det\{B(z)\} \neq 0$ for $|z| \leq 1$. These are the conditions in the Wold decomposition

[9] This is the assumption made, for example, in, e.g., Lewis and Reinsel (1985) and Kuersteiner (2005). However, Gonçalves and Kilian (2007) provide similar results for possibly conditionally heteroskedastic martingale difference sequences.

theorem (Wold, 1938) by which any stationary vector process can be decomposed into its deterministic components and Equation 1. Note that the assumption of stationarity is not required for local projections (Jordà, 2005; Plagborg-Møller and Wolf, 2018).

4.2. Impulse responses from a VAR(p)

Under the stated assumptions, Equation 1 is invertible and can be expressed as the following infinite order vector autoregression or VAR(∞) process

$$\boldsymbol{y}_t = \sum_{j=1}^{\infty} A_j \boldsymbol{y}_{t-j} + \boldsymbol{\epsilon}_t \,, \tag{2}$$

where $\sum_{j=1}^{\infty} ||A_j|| < \infty$ and $A(z) = I - \sum_{j=1}^{\infty} A_j z^j = B(z)^{-1}$ satisfies $\det\{A(z)\} \neq 0$ for $|z| \leq 1$.

Assuming the data are generated by a process such as Equation 1 is quite general and it will allow us to investigate instances when the monetary shocks have long-run effects. Note that invertibility is not a necessary condition for local projections, as discussed in Stock and Watson (2018) and Plagborg-Møller and Wolf (2018).

Using this setting and under a variety of general assumptions, Lewis and Reinsel (1985), Kuersteiner (2005), Gonçalves and Kilian (2007) and references therein establish the consistency and asymptotic normality of the VAR(p) coefficient matrices A_j for $j = 1, \ldots, p$. However, two recurring assumptions merit further discussion. One is the assumption that $p \to \infty$ as $T \to \infty$ at a rate $p^3/T \to 0$. That is, p is allowed to grow with the sample size, but not too fast. As remarked by Kuersteiner (2005), this rate does not provide much guidance as to how to choose p and hence, he provides an optimal procedure to determine p in finite samples.

The second and more relevant assumption often made is that p is chosen as a function of T such that

$$\frac{1}{T^{1/2}} \sum_{j=p+1}^{\infty} ||A_j|| \to 0 \quad \text{as} \quad p, T \to \infty, \tag{3}$$

to ensure consistency of $\hat{A}_1, \ldots, \hat{A}_p$.

In turn, this means that the impulse response function will be consistently estimated up to horizon h. To see this, recall the well-known recursion (Durbin, 1959) relating moving-average and autoregressive coefficient matrices,

$$B_h = A_h + A_{h-1}B_1 + \cdots + A_1 B_{h-1}. \tag{4}$$

Given this recursion and direct application of the continuous mapping theorem, it is easy

to see that consistency of $\hat{A}_1, \ldots, \hat{A}_p$ means that

$$||\hat{B}_j - B_j|| \xrightarrow{p} 0 \quad \text{as} \quad p, T \to \infty \quad \text{for} \quad j = 1, \ldots, p. \tag{5}$$

However, consistency of B_h is not guaranteed for $h > p$. The source of the inconsistency can be quantified to be

$$A_{p+1}B_{h-(p+1)} + \cdots + A_{h-1}B_1 + A_h\,,$$

so that the sign of the inconsistency is, a priori, unknown.

Note that, all along, we have assumed that the truncation lag is optimally chosen—though in practical applications, it is likely that practitioners will chose to small a truncation lag so that

$$\frac{1}{T^{1/2}} \sum_{j=p+1}^{\infty} A_j \to c \neq 0\,,$$

in which case, the inconsistency due to misspecification of the model is of order $T^{-1/2}$. Setting aside this problem, a simple example illustrates inconsistency resulting from optimal truncation alone.

Consider, for expositional purposes, truncation at $p = 1$ and for now assume that this is sufficient to ensure consistency of $\hat{A}_1$. Using expression Equation 4, we next show both the impulse response coefficients of the VAR(∞) (in the first column of Equation 6 below) against the relationships resulting from the truncated VAR(1) (presented in the second column of Equation 6 below), specifically

$$VAR(\infty) \to \begin{cases} B_1 &= A_1 \\ B_2 &= A_1^2 + A_2 \\ B_3 &= A_1^3 + 2A_1A_2 + A_3 \\ B_4 &= A_1^4 + 3A_1^2A_2 + 2A_1A_3 + A_4 \\ \cdots \end{cases}, \qquad VAR(1) \to \begin{cases} B_1 &= A_1 \\ B_2 &= A_1^2 \\ B_3 &= A_1^3 \\ B_4 &= A_1^4 \\ \cdots \end{cases}. \tag{6}$$

Clearly, extending the lag-length of the VAR from p to H (where H is the longest horizon displayed for the impulse response function) would solve the problem (assuming the remaining conditions in the theorem are met). In practice, and due to limited sample sizes, practitioners tend to favor relatively short lag-lengths, which potentially exposes them to undesirably inconsistent estimates of the impulse response even when the coefficients of the VAR are consistently estimated.

4.3. Impulse responses from local projections

Under the same assumptions stated earlier, local projections have better consistency properties than truncated VARs, in general. Consider the local projection

$$\boldsymbol{y}_{t+h} = A_{h,1}\boldsymbol{y}_{t-1} + \cdots + A_{h,p}\boldsymbol{y}_{t-p} + \boldsymbol{u}_{t+h}\,. \tag{7}$$

The assumptions that ensure consistency of the $\hat{A}_j$ for $j = 1, \ldots, p$ also ensure that

$$||\hat{A}_{h,1} - B_{h+1}|| \xrightarrow{p} 0 \quad \text{as} \quad p, T \to \infty\,, \tag{8}$$

as shown in Lusompa (2019). The appendix provides a broad overview of the main arguments for this result. In particular, even if the DGP is a $VAR(\infty)$ and the local projection only includes up to p lags, the LP provides a direct estimate (rather than a nonlinear transformation of estimated parameters) of the impulse response coefficient matrix for any h. Moreover, the appendix shows that consistency of the truncated local projection is achieved under the exact same conditions as consistency of the VAR coefficient matrices for the first p lags.

What does this all mean in practice? Under general conditions that would lead researchers to consistently estimate a finite-order VAR given a sample of data, the impulse response coefficients corresponding to horizons that exceed the truncation lag of the VAR would be expected to exhibit bias. This bias can be quantified theoretically though in practice it will depend on several factors. The bias arises because the truncation of the VAR exposes the calculation of the impulse response coefficient to omitted terms that may not be small, specially when the data are persistent. In those cases, of course, a higher truncation lag would be advisable. In practice, researchers tend to choose relatively parsimonious VAR specifications due to the limited sample sizes and the parameter estimation load required in a VAR. Moreover, Bayesian methods are unlikely to result in better outcomes as shrinkage to the priors commonly used will make the VARs more efficient, but possibly make the bias even worse for impulse responses.

Local projections have several advantages. Because they provide a direct estimate of the impulse response coefficient of interest (rather than obtaining the response coefficients as the result of a nonlinear transformation of estimated parameters), truncation of the lag-length is not as big a concern. Moreover, because local projections can be estimated equation by equation, the parametric load is reduced. In settings where the data are persistent, such as ours, and where interest is in characterizing impulse responses at long horizons, local projections are more appealing.

4.4. Impulse response small-sample biases: Monte Carlo evidence

To further illustrate the potential biases that can arise in practice, the next section provides a simple Monte Carlo exercise to illustrate the main ideas. The DGP is an MA(10) whose coefficients are generated from a Gaussian basis function. Specifically, we set

$$y_t = \mu + \sum_{j=1}^{10} b_j u_{t-j} + u_t \,; \qquad u_t \sim N(0,\, 0.5)\,;$$

$$b_j = \frac{b_j^*}{\sum_j b_j^*}\,; \qquad\qquad b_j^* = \alpha \exp\left\{-\left(\frac{j-\beta}{\delta}\right)^2\right\}, \quad j = 1, \ldots, 40\,.$$

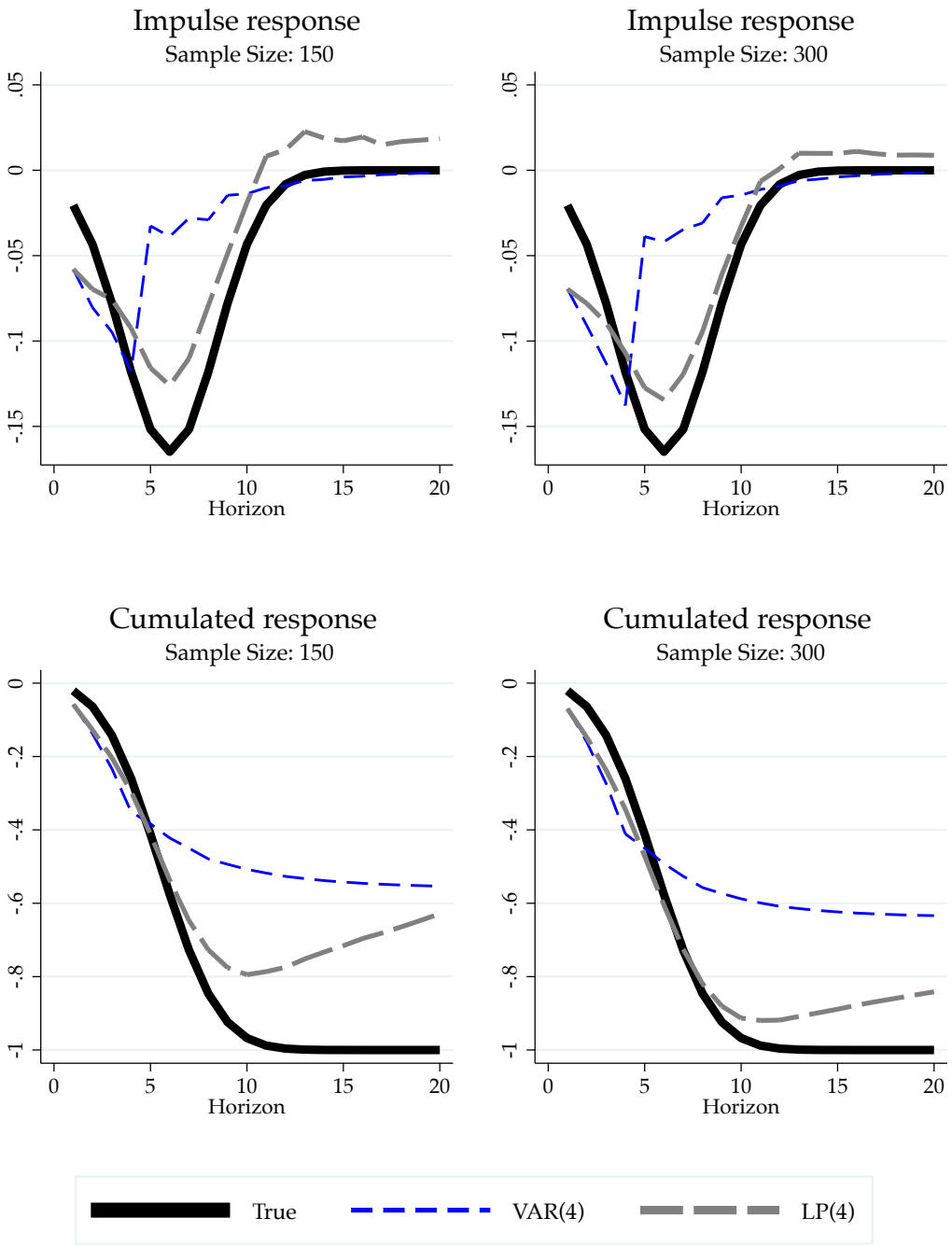
with $\alpha = 1$, $\beta = 6$, and $\delta = 12$. Using these parameter choices, the maximum impact is normalized to have size 1, which occurs 6 periods after impact, and tapers back to 0 with a half-live given by $12\sqrt{\ln 2} \approx 10$. The resulting b_j^* are then scaled so that the long run impact is normalized to 1, resulting in the b_j parameters that we ultimately use in the simulation.

The experiment then consists in simulating the MA(10) using Gaussian shocks $N(0, 0.5)$. The constant term is set to $\mu = 1$. We initialize the process with 500 draws that are then discarded to avoid initialization issues. We consider two sample sizes, 150 and 300 observations. The experiments are repeated 1,000 times and the average over these Monte Carlo replications is then reported. Hence, we estimate impulse responses for this MA(10) process using a VAR(4) and also by local projections with 4 lags, and denoted LP(4). Note that both of these models are misspecified since a lag length of 10 would be appropriate.

The top row of Figure 1 displays the raw impulse responses and the bottom row displays the cumulated responses, which are the responses later used in our analysis. Moreover and without loss of generality, we flip the sign of the responses to match the empirical responses below. Several results are worth noting. First, it is clear that the impulse responses estimated with local projections fit the true impulse responses much better than those from a VAR. Because the MA(10) is designed to have a maximum effect that is delayed by 6 periods after the shock, the VAR(4) has a hard time picking the shape of the response. This is not alleviated by increasing the sample size to 300. As a result, the cumulated responses can be far off, as shown in the figure.

The bottom row of Figure 1 reports the cumulated responses corresponding to the impulse responses calculated in the top row. Because the cumulated impact is normalized to 1, it is easy to assess the bias resulting from the small sample estimates. Regardless of whether one uses the smaller or the large sample size, the VAR(4) based response reduces the long-run impact by about 40% from the true measure. This bias is less than half for the LP(4), even though the LP has the same lag length as the VAR.

Figure 1: *Evaluating VAR versus LP responses in finite samples: A Monte Carlo experiment*



Notes: Impulse responses in the top row calculated with a VAR(4) and local projections with four lags or LP(4). Bottom row displays cumulated responses. Sample size with 150 observations displayed in the left-hand column, sample size with 300 observations displayed on the right-hand column. Averages over 1,000 Monte Carlo replications. See text.

5. Empirical approach: Monetary shocks have long-lived effects

The basic empirical approach relies on local projections (Jordà, 2005) estimated with instrumental variables (LPIV). Several applications of these methods are available in the literature, though a more general discussion of the method can be found in Ramey (2016), Stock and Watson (2018) and Jordà, Schularick, and Taylor (2019).

Based on the latter, we estimate in particular the impulse responses

$$y_{j,t+h} - y_{i,t-1} = \alpha_{i,h} + \widehat{\Delta i}_{j,t}\,\beta_h + \boldsymbol{x}_{j,t}\,\boldsymbol{\gamma}_h + y_{b(j,t,t+h)}\lambda_h + v_{j,t+h}\,, \tag{9}$$

for $h = 0, 1, \ldots, H; i = 1, \ldots, N; t = t_0, \ldots, T$, where $y_{j,t+h}$ is the outcome variable for country j observed h periods from today, $\alpha_{i,h}$ are country fixed effects at horizon h, $\widehat{\Delta i}_{j,t}$ refers to the instrumented change in the short-term government bond (3-months in duration), our stand-in for the policy rate which we instrument with $z_{j,t}$, the trilemma instrument as discussed earlier; and $\boldsymbol{x}_{j,t}$ collects all additional controls including lags of the outcome and interest rates, as well as lagged values of other macro aggregates.[10] Moreover, we control for global business cycle effects through a global real GDP control variable to parsimoniously soak up time effects. To tie our hands as much as possible against finding any long-run effects of monetary policy, we also control for future values of real GDP in the base country at each horizon h, denoted by $y_{b(j,t,t+h)}$. Motivated by the identification discussion in Section 3.1, these set of controls also account for potentially anticipated information on the base-country business cycle. We estimate Equation 9 with instrumental variable methods and report cluster robust standard errors.

Table 1 reports the first-stage regression of the pegging country's short term interest rate $\Delta i_{j,t}$ on the instrument $z_{j,t}$ with controls $x_{j,t}$, country fixed effects and (robust) clustered standard errors. The t-statistic is well above 3 for full sample and post-WW2 samples illustrating that it is not a weak instrument. We refer the reader to Jordà, Schularick, and Taylor (2019) for detailed discussion on the instrument and proceed henceforth assuming the reader is on board regarding instrument relevance and strength.

[10] 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.

Table 1: *Trilemma instrument: First stage evidence*

pegs ($q = 1$)	All years	PreWW2	PostWW2
$z_{j,t}$	0.52^{***}	0.35^{**}	0.56^{***}
t-statistic	[8.62]	[2.05]	[8.97]
Obs	672	148	524

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. See text.

5.1. Main results

The main story is illustrated by the response of real GDP to a shock to domestic interest rates using the trilemma instrument. Before we show the main results, we highlight the value of our instrumental variable with a comparison of the response of real GDP to a shock in the short-term domestic interest rate calculated using selection-on-observables identification versus identification with the trilemma instrument. This is shown in Table 2. The table reports coefficient estimates of the impulse response calculated with each identification approach for the full and post-WW2 samples. *LP-OLS* refers to identification via selection, *LP-IV* to the trilemma instrument identification. The samples are restricted to pegging economies to match the samples in both cases.[11]

Table 2 is organized as follows. We provide the coefficient estimates by row, with a test of the null hypothesis that *LP-OLS* and *LP-IV* estimates are equlal for the full and post-WW2 samples. 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. 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 3.69 percent over 12 years. A similar effect is found when we restrict the sample to post-WW2. The drop 12 years after impact is 3.49 percent. This is a far cry from traditional notions of long-run neutrality found in the literature.

What is the source of this persistent decline? We decompose GDP into its components, namely, hours worked (employees times number of hours per employee); capital stock (measured capital in machines and buildings); and the Solow residual (using a Cobb-Douglas production function) labeled as total factor productivity (TFP). (Using the Imbs

[11] The plots and inference are robust to using real GDP per capita. See Table A1

Table 2: *LP-OLS vs. LP-IV. Attenuation bias of real GDP responses to interest rates. Trilemma instrument. Matched samples*

Responses of real GDP 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)
$h = 0$	0.07**	0.01	0.50	0.03	0.08	0.44
	(0.03)	(0.08)		(0.02)	(0.07)	
$h = 2$	-0.31**	-1.37***	0.00	-0.29***	-1.20***	0.00
	(0.14)	(0.34)		(0.11)	(0.31)	
$h = 4$	-0.22	-1.89***	0.00	-0.15	-1.24***	0.01
	(0.23)	(0.48)		(0.20)	(0.40)	
$h = 6$	-0.12	-2.67***	0.00	-0.07	-2.55***	0.00
	(0.36)	(0.59)		(0.29)	(0.59)	
$h = 8$	-0.34	-4.83***	0.00	0.08	-2.93***	0.00
	(0.35)	(1.03)		(0.35)	(0.71)	
$h = 10$	0.13	-4.22***	0.00	0.34	-2.80***	0.00
	(0.37)	(1.03)		(0.42)	(0.80)	
$h = 12$	0.35	-3.82***	0.00	0.56	-3.31***	0.00
	(0.40)	(0.89)		(0.43)	(0.92)	
KP weak IV		46.71			50.56	
H_0: $LATE = 0$	0.00	0.00		0.00	0.00	
Observations	835	765		655	585	

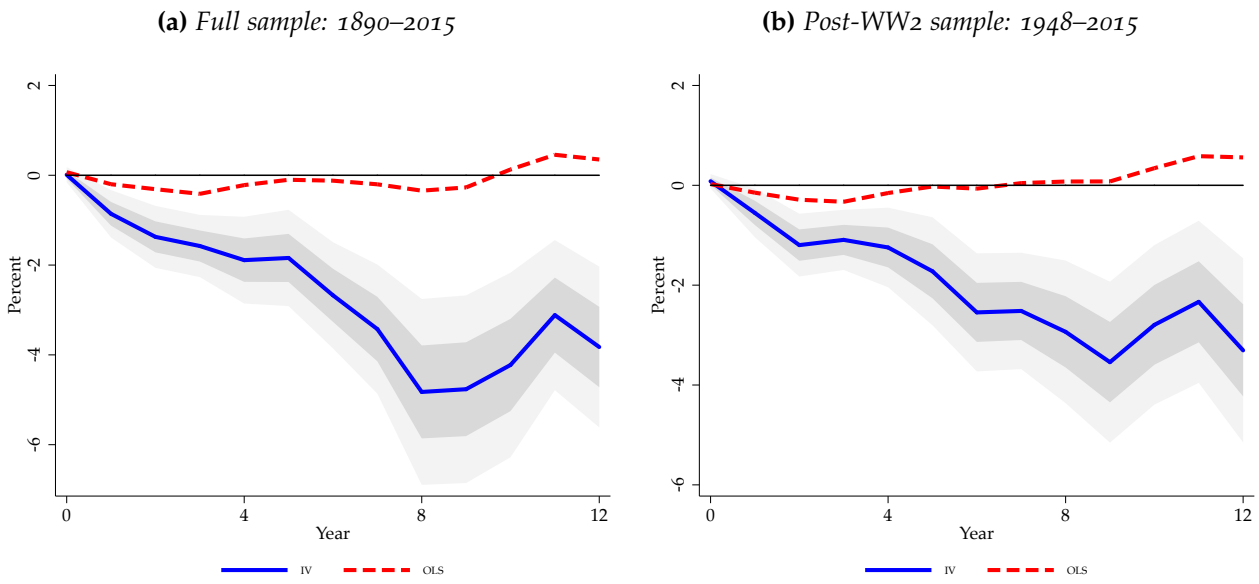
Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Cluster robust standard errors in parentheses.Full sample: 1890–2015 excluding WW1: 1914–1919 and WW2: 1939–1947. Post WW2 sample: 1948–2015. *Matched sample* indicates LP-OLS sample matches the sample used to obtain LP-IV estimates. *KP weak IV* refers 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.

(1999) correction, in Appendix Section B we further decompose the Solow residual into factor-utilization plus a residual utilization-adjusted TFP.)

Figure 3 displays the responses of each of these components to the same shock to the domestic short-term interest rate instrumented with the trilemma, both for the full and the post-WW2 samples. Figure 3a displays the responses of total hours worked, capital and raw TFP without error bands to provide a clearer sense of the dynamic paths. Figure 3b displays each of the components with the one and two standard deviation error bands. In Appendix Section D.2, we provide corresponding figure for post-WW2 sample.

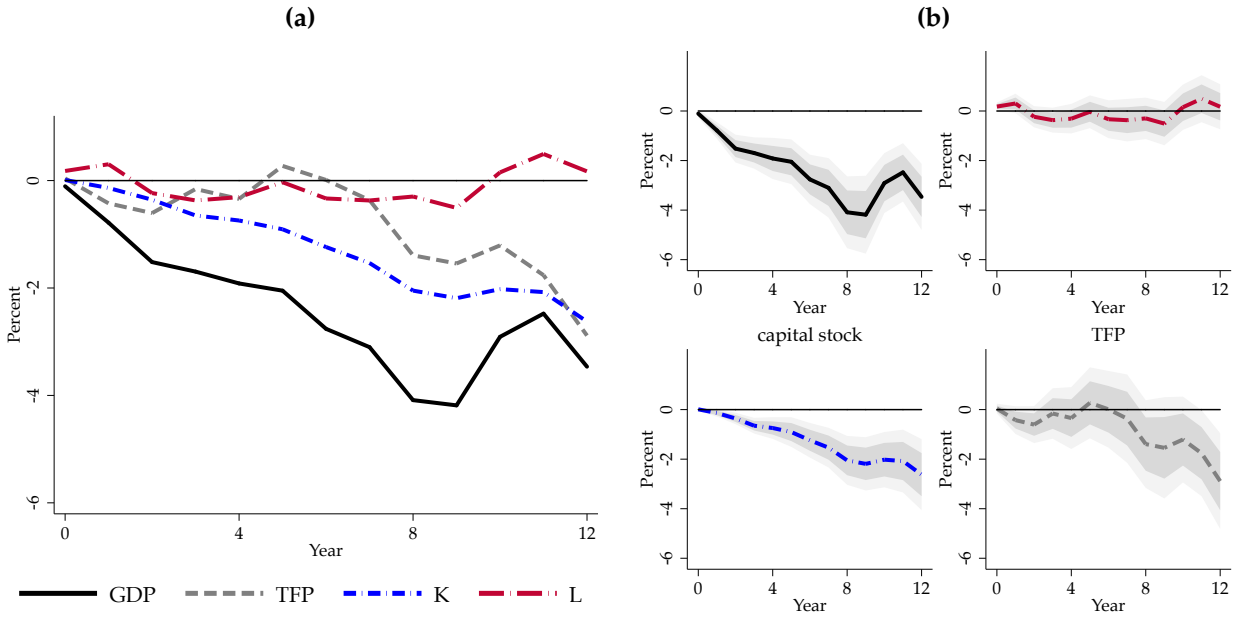
Several features deserve mention. Figure 3 shows that there are similar declines in capital and raw TFP whereas total hours worked exhibits a much flatter pattern. Because

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



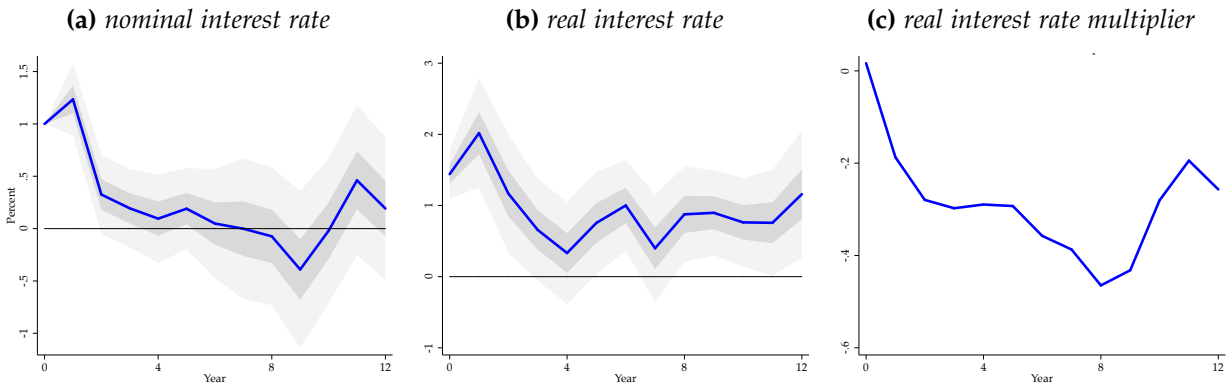
Notes: Response to a 100 bps shock in domestic interest rate instrumented with the trilemma. Responses for pegging economies. Full sample: 1890–2015 (World Wars excluded). Post-WW2 sample: 1948–2015.LP-IV estimates displayed as a solid blue line and and 1 S.D. and 2 S.D. confidence bands constructed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

Figure 3: *Baseline response to 100 bps trilemma shock: Real GDP and components*



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 thick line and and 1 S.D. and 2 S.D. confidence bands. See text.

Figure 4: *Baseline response to 100 bps trilemma shock: a) Short term nominal interest rate, b) real interest rate, and c) real interest rate multiplier*



Notes: Response to a 100 bps shock in domestic interest rate instrumented with the trilemma. a) Responses of short term nominal interest rate for pegging economies. b) Responses of short term real interest rate for pegging economies. Inflation expectations constructed from the impulse response of consumer price level index to the same trilemma shock. c) Response of cumulative change in real GDP divided by cumulative change in short term nominal interest rate for pegging economies. Sample: 1890–2015 (World Wars excluded). LP-IV estimates displayed as a solid blue line and and 1 S.D. and 2 S.D. confidence bands. See text.

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 TFP variable, and then capital, with total hours worked mostly flat.

To be sure, the response of hours worked conforms well with the textbook response to a monetary shock. Total hours fall in the short-run, but then recover quickly and remain mostly flat. Capital accumulation also follows textbook dynamics in the short-run. The response is initially muted but builds up over time. But unlike a textbook new Keynesian model (Galí, 2015b), capital does not appear to recover even 12 years after the shock. Similarly, TFP falls gradually rather than suddenly. Over time, the decline in TFP accelerates, ending at a level 3.32 percent lower in full sample by year 12 relative to year 0.

One 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 of this proposition can be done in two steps.

Figure 4 shows that the short-term real interest rate does indeed take approximately 8 years to return to zero deviation, while the nominal interest rate returns to zero deviation after 4 years. The response of nominal interest rate is typical of what has been reported often in the literature (see, e.g., Christiano, Eichenbaum, and Evans, 1999; Ramey, 2016). Secondly, we can calculate the responses of the main variables normalized by the response of interest rates over time to sterilize the dynamics of interest rates themselves. This is no different than calculating a multiplier (see, e.g., Ramey, 2016; Ramey and Zubairy, 2018).

Thus, in Figure 4c we show the ratio of the cumulative change in GDP to the area under the real interest rate path in Figure 4b. In a given period, the difference in the level of real interest rate relative to the counterfactual path measures the tightness of monetary policy. We consider these cumulative gaps as a measure of overall monetary policy tightness (the area under the solid line in Figure 4b). By year 12, the multiplier is -0.4 in the full sample. In post-WW2 sample, this multiplier is similar (See Section D.2 in the appendix). These numbers indicate that the monetary policy shocks, measured as cumulative response of real interest rate, are somewhat larger but comparable with those obtained in the literature using VAR and Romer and Romer (2004) shocks for the U.S.[12]

6. Robustness and discussion

Our baseline specification included lags and current values of global GDP growth and base country GDP growth, and future values of base country 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, in IRF estimation. 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.

6.1. 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 or the spillovers highlighted in Section 3.1. 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 business cycle variation. In addition to the spillover control strategy used in our baseline specification, we now assess such spillover effects more formally, using two separate approaches: a) a control function approach developed in Jordà, Schularick, and Taylor (2019), and b) controlling for foreign variables at each horizon to remove any spillover channels from the interest rate channel.

[12] In their online appendix (Table A.1), motivated by the findings of Coibion (2012), Nakamura and Steinsson (2018) report the multiplier of monthly industrial production to monetary shocks over 36 months to lie between 1 and 2 using the local projections specification with VAR shocks and Romer-Romer shocks.

Synthetic control function approach

In this section we use a direct spillover correction based on Conley, Hansen, and Rossi (2012); van Kippersluis and Rietveld (2018) that takes advantage of the subpopulation of floats (for which our instrument is theoretically invalid). The solution borrows from Jordà, Schularick, and Taylor (2019), where the reader can find the detailed derivations. Here we simply sketch the main ideas with a simple example.

Suppose the equation to be estimated is:

$$\Delta y = \Delta i\, \beta + z\phi + v\,,$$

where Δy is the outcome variable, Δi is an endogenous variable, and z is an instrumental variable that violates the exclusion restriction as long as $\phi \neq 0$. However, we may assume that $E(z\, v) = 0$, as is common in regression analysis when z is exogenous. Although for the subpopulations of floats z is not a valid instrument, it can be exploited to provide an estimate of the spillover effect ϕ using an OLS estimate of the previous regression and the first stage (invalid) regression

$$\Delta i = z\, b + \eta\,.$$

Without loss of generality, we can assume that $\beta = \lambda\phi$ where λ is an unknown parameter than can be calibrated through economic arguments since it is simply a statement of how large is the spillover effect relative to the effect of the endogenous variable. We choose a conservative set of values for $\lambda \in [1, 8]$. In our case, the endogenous variable is the change in the home nominal interest rate and the instrument is, essentially, the surprise in the base country rate. As a result, one can show that

$$\hat{\phi}(\lambda) = \frac{(\hat{\phi}_{OLS} + \hat{b}\,\hat{\beta}_{OLS})}{1 + \lambda\,\hat{b}}\,,$$

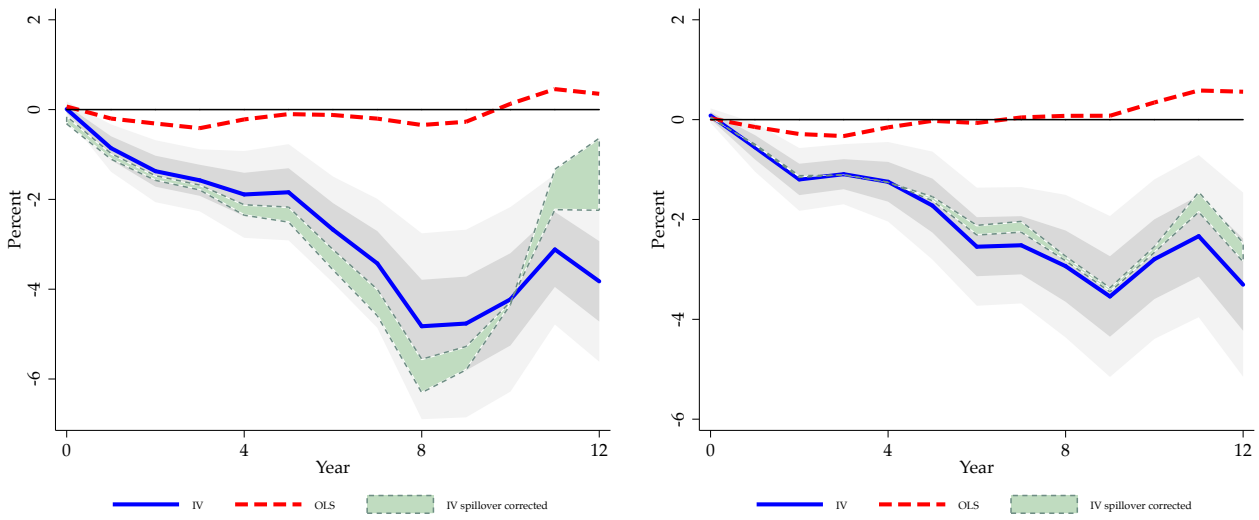
and hence estimate by instrumental variables, on the subpopulation of pegs, the following spillover corrected regression

$$(\Delta y - z\hat{\phi}(\lambda)) = \Delta i\, \beta + v + z(\hat{\phi}(\lambda) - \phi(\lambda))\,,$$

where it is clear that the error term is the sum of two terms whose mean converges to zero asymptotically. 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. Direct spillover corrections have negligible effects on our estimates.

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 interest rate instrumented with the trilemma. Responses for pegging economies. 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 and 1 S.D. and 2 S.D. confidence bands, LP-IV spillover corrected estimates displayed as a light green shaded area with dashed border, using $\lambda \in [1, 8]$. See text and Jordà, Schularick, and Taylor (2019).

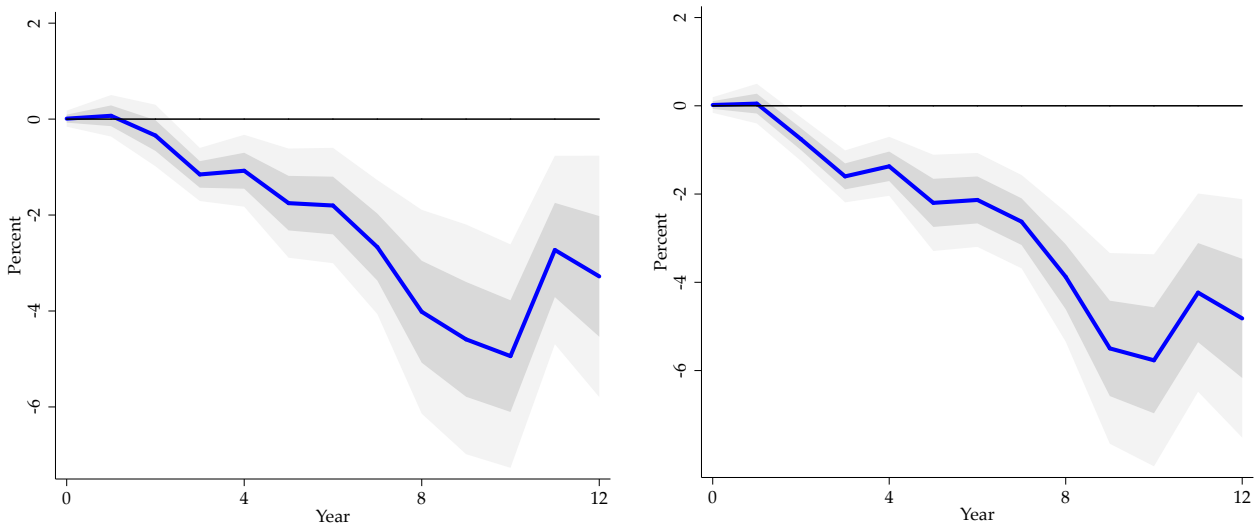
Controlling for base country GDP growth, current account and exchange rate

A second approach that attempts to provide validity to the exclusion restriction is directly controlling for a primary channel through which the spillover effects may originate. A reduction in demand my arise from other trading partners when the tightening affects other countries as well. Or the interest rate channel induced contraction in pegging economy, by reducing demand for other countries' output. may be subject to *spillbacks*. If the transmission is primarily driven by increased trade, controlling for global GDP growth can potentially absorb these effects allowing us to remove domestic demand effects with respect to international spillbacks.

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. Another implication from the Mundell-Fleming-Dornbusch model is that there are financial spillovers that may amplify the effects through the exchange rate channel. To account for these effects, at each horizon h, 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 at each horizon.

Figure 6: *Response to 100 bps trilemma shock with additional controls: Real GDP*

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



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.D. and 2 S.D. confidence bands. See text.

Figure 6a plots the IRFs to trilemma identified shock. Controlling for variables motivated by Section 3.1 does not affect our main result - monetary shocks have a very persistent effect on real GDP.

6.2. Accounting for structural breaks

Fernald (2014) and Gordon (2016) have convincingly argued that there are structural breaks in TFP growth in the U.S. economic trajectory. It is plausible that there are structural breaks in other economies' TFP growth rates. If these structural breaks implying slowdown in TFP growth occur around the identified monetary shocks, it could bias our results leading us to attribute the persistent effects incorrectly to monetary shocks. 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 & Perron (1998). We report these estimated structural break dates in the appendix D.7. Then in our baseline specification, we allow output growth to lie in either of the five regimes at horizon zero as well as horizon h. Our specification is conservative since we allow horizon h regime changes in the estimation along with horizon 0 regimes.

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.[13]

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

6.3. Additional robustness exercises

We conduct variety of additional robustness exercises to check the validity of our results. These are shown in the appendix and we describe them briefly now.

In Section D.3, we show the IRFs of real GDP to monetary shock when we introduce the control variables in levels instead of first differences. We also show the responses when we increase the number of lags in our estimation to five. Results are robust to either of these substitutions.

In Section D.4, we show trajectory of various macroeconomic and financial variables to the identified monetary shock. In particular, we trace out path of CPI-based measure of price level, consumption per capita, investment per capita, long-term nominal rate, loans to GDP ratio, real and nominal house price indices, and real and nominal stock price indices.

In Section E, we report the estimation results when the dependent variable is real GDP per capita instead of real GDP. Our results are robust to not including data post-2007 (i.e., after the global financial crisis). Finally, while we used an unbalanced sample in our baseline estimation (i.e., estimates of IRFs at shorter horizons are estimated with more data than those at longer horizons), our results are robust to using a fixed estimation sample at all horizons.

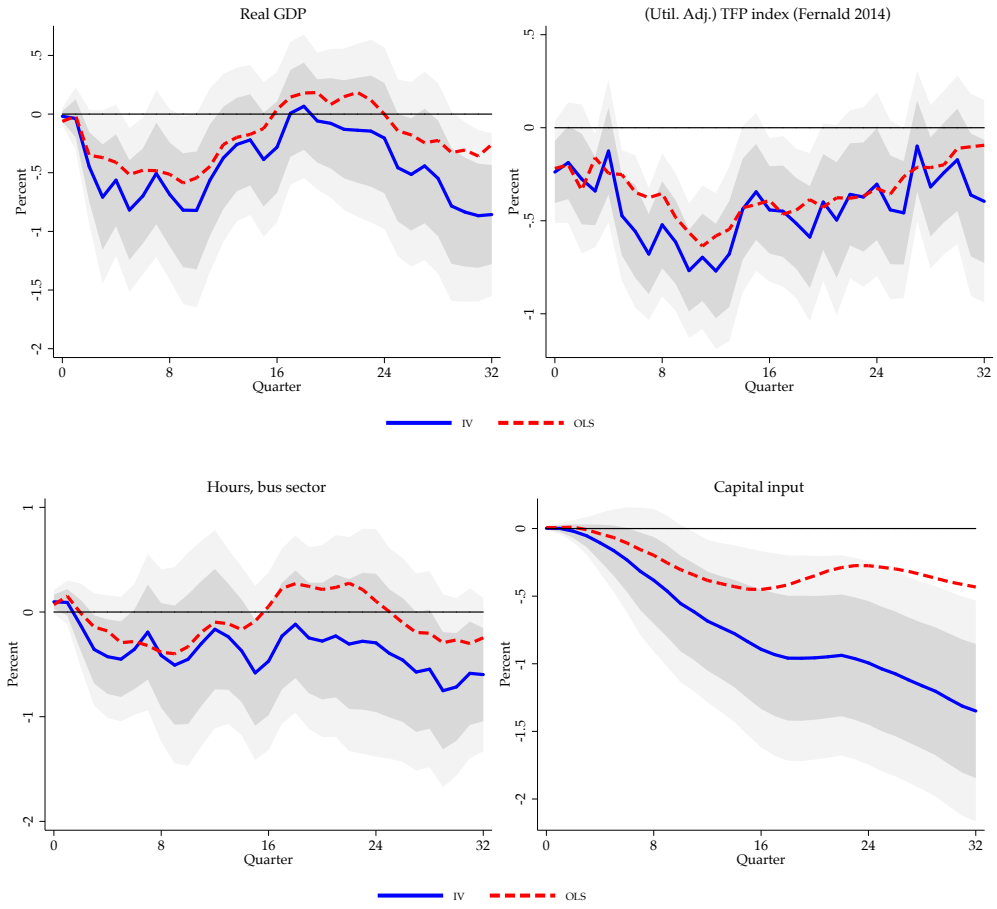
6.4. Is the U.S. different?

The U.S. for most of our historical sample is not a pegging economy (apart from the Gold standard years). It is the quintessential base country for many economies in our sample. For this reason, the trilemma instrument mechanically sets aside any information coming from the U.S. during estimation. It is natural to wonder the extent to which U.S. data conforms with the patterns presented so far.

In this section, we examine U.S. data post-WW2. This allows us to incorporate three useful modifications. First, we use higher frequency quarterly data. Second, we use the alternative utilization adjusted series for TFP constructed by Fernald (2014). Third, in order to achieve identification we rely on the instrumental variable constructed by Romer and Romer (2004) based on the Federal Reserve staff's implied forecast errors for the policy rate, and extended to recent years by Wieland and Yang (2016).

It turns out that results based on U.S. data largely confirm the results we reported in the previous section for non-U.S. pegging economies. Figure 7 plots the the path of real GDP along with its three components: total hours worked, capital, and utilization adjusted TFP. The responses are qualitatively similar to those in the long-run panel, although the amplitudes are more muted. Quarterly data is naturally noisier than yearly data, but

Figure 7: *Baseline response to 100 bps Romer and Romer (2004) shock : U.S. postwar data*



Notes: Response to a 100 bps shock in federal funds rate rate instrumented with policy forecast residuals (Romer and Romer, 2004; Wieland and Yang, 2016). Responses of real GDP, utilization-adjusted TFP (Fernald, 2014), capital stock and hours worked for U.S. economy. Quarterly sample: 1969-Q1: 2007-Q4. Quarterly data series are taken from Fernald (2014).

smoothing over a temporary recovery in GDP 4 to 5 years after the shock, real GDP ends nearly one percent lower 8 years after impact. Utilization adjusted TFP and hours exhibit a similar U-shaped pattern, with TFP nearly back to zero by year 8. Strikingly, capital accumulation exhibits a protracted decline over the entire period, ending about 1.25 percent lower after eight years. For comparison, we plot the various components on the same graph in the appendix (see Section D.9).[14]

[14]We estimate the following specification for the U.S. economy:

$$y_{t+h} - y_{t-1} = \alpha_h + \widehat{\Delta i}_t \, \beta_h + \boldsymbol{x}_t \, \boldsymbol{\gamma}_h + v_{t+h} \, ; \qquad h = 0, 1, \ldots, H; \, i = 1, \ldots, N; \, t = t_0, \ldots, T \, ;$$

where y_{t+h} is the outcome variable at horizon h, $\widehat{\Delta i}_t$ is the instrumented change in the Federal Funds rate, $\boldsymbol{x}_t$ is the set of controls that includes contemporary and four lags of log real GDP, log CPI, and changes in federal funds rate. We do not include the contemporary variable when it is same as the dependent variable. We report robust standard errors.

The data for utilization-adjusted TFP series are based on sectoral data for the U.S. that account for heterogeneity across workers and types of capital. Fernald (2014) notes that there are various other corrections that are not conducted in the quarterly series due to the lack of rich-industry level data.

The finding that monetary policy shocks can affect utilization-adjusted TFP echo the Evans (1992)'s critique of using Solow residuals as productivity shocks in RBC models. While the construction of quarterly utilization-adjusted TFP series is detailed and thorough, our analysis suggests caution against using the quarterly-adjusted residuals as "pure" TFP shocks (i.e., perfectly orthogonal to demand shocks).[15,16]

7. A model of hysteresis

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 a textbook New Keynesian model with endogenous productivity growth in a stylized manner. The baseline framework is that of a medium-scale DSGE model as in Christiano, Eichenbaum, and Evans 2005; Smets and Wouters 2007; or Justiniano, Primiceri, and Tambalotti 2013.

Households demand final consumption good and supply differentiated labor to a labor union. Final consumption good is packaged by a perfectly competitive retailer using a Dixit-Stiglitz aggregate of a continuum of intermediate goods. Each intermediate good is produced by a monopolistically-competitive firm that sets prices in a Calvo fashion. These firm hire labor unit packaged by labor unions that also set wages in a staggered fashion. Government runs a balanced budget every period and central bank sets short term nominal interest rate on a riskfree bond (in net zero supply) following a rule that we describe shortly. Goods market and bond markets clear every period. We leave the formal model to the appendix, and focus on the key departure that allows us to introduce productivity hysteresis with a parametrically convenient process.

[15] Ramey (2016) also documents that utilization adjusted TFP series fail Granger causality tests (See Table 11 in her paper). We are grateful to Valerie Ramey for alerting us to that analysis.

[16] In the appendix D.10, we show that similar results are obtained with samples beginning until at least 1973Q2. However, these persistent effect results are not robust to considering further shorter samples for the U.S. economy. This is consistent with the findings of Coibion, Gorodnichenko, and Ulate (2017).

7.1. Hysteresis 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 Z_t than is conventional. In particular, we assume that the law of motion for Z_t is:

$$\log Z_t = \log Z_{t-1} + \mu_t + \eta \log \left(Y_{t-1} / Y_{t-1}^{f,t-1} \right) ,$$

where μ_t is the exogenous component of the TFP growth rate, that may be subject to trend shocks. Y_t is aggregate output at time t. $Y_{t-1}^{f,t-1}$ is the flexible price level of output in period $t-1$ conditional on Z_{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 as the *hysteresis elasticity* (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.[17] 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, 2016). 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 series.

This functional dependence creates a role for hysteresis stabilization by central banks in a reduced-form manner (Yellen, 2016). Long-run effects of monetary policy shocks depend on the value of η. Theoretically, there is no a priori reason to expect η to be positive. While a "cleansing" effect of recessions may induce counter-cyclicality, recessions may reduce funding access to firms to conduct R&D, skill development, and learning-by-doing. The sign on the cyclicality of TFP misallocation is also ambiguous and depends on the assumptions in a model. To clarify, in this paper, we are only able to discuss the sign and the magnitude of η in response to temporary monetary shocks. We will provide below estimates for η in response to monetary shocks from our estimated empirical IRFs.

[17] A similar setup was used by Stadler (1990) in his seminal work.

7.2. 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.

$$\frac{1+i_t}{1+i_{ss}} = \left(\frac{1+i_{t-1}}{1+i_{ss}}\right)^{\rho_R} \left[\left(\frac{\pi_t}{\pi_{ss}}\right)^{\phi_\pi} \left(\frac{Y_t}{Y_t^{f,t}}\right)^{\phi_y}\right]^{1-\rho_R} \epsilon_t^{mp}, \tag{10}$$

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.

We assume government balances budget every period, where total lumpsum taxes go into giving a production subsidy to intermediate good producers, a wage subsidy to workers and other government spending.

7.3. Simulations

As the DSGE model is intentionally standard, we take parameters from the literature Justiniano, Primiceri, and Tambalotti (2013). We report these in Table 3. The steady state parameters imply (annualized) real interest rate of 2.40%, and an investment-GDP ratio of 17%. For the monetary shock process, we chose the following parameters $\sigma_r = 0.02$ and $\rho_R = 0.8$. The short term nominal interest rate increases by about 10 basis points on impact in the three simulations that we report here.

The new parameter in our model, relative to the business cycles literature, is η: the hysteresis elasticity. Table 4 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 IRF of TFP and real GDP to monetary policy shock. Using the estimated coefficients (see Figure 3), we then estimate η as the ratio of the two IRFs. 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 0.10 which is also within the assumptions of DeLong and Summers (2012).

Figure 8 plots the model-implied impulse responses for output, capital stock, real interest rate and inflation after a monetary policy shock. Solid blue line reports the IRFs for endogenous growth model with $\eta = 0.10$, and dashed blue line reports IRFs for the comparable exogenous growth benchmark i.e. $\eta = 0$. The IRFs for output and capital stock

Table 3: *Parameters*

(a) Steady-state parameters					
	β	δ_k	α	μ	
	Discount factor	Capital depreciation rate	Capital share	Trend growth rate	
	0.999	0.025	0.28	2%	
(b) Parameters characterizing the dynamics					
ν	λ_p	λ_w	θ_p	θ_w	h
Inverse Frisch elasticity	Price s.s. markup	Wage s.s. markup	Price Calvo probability	Wage Calvo probability	(Internal) habit
1.00	0.15	0.15	0.750	0.750	0.5
$\frac{a''(1)}{a'(1)}$	$S''(1)$	ϕ_π	ϕ_y	$1-\frac{1}{\lambda_g}$	η
Capital utilization cost	Investment adjustment cost	Taylor rule inflation response	Taylor rule (normalized) output response	Government spending share	Hysteresis elasticity
4	2	1.50	0.125	0.20	0.05

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

Table 4: *Point estimates for hysteresis elasticity η*

pegs ($q = 1$)	1890–2015	1948–2015
η	0.28	0.42
[95% confidence band]	$[0.10, 0.45]$	$[0.33, 0.50]$

Notes: The point estimates are estimated with a two-step classical minimum distance approach using the IRF of TFP and the IRF of GDP to the monetary policy shock in the second step. Figure 3 report the IRFs for GDP and TFP.

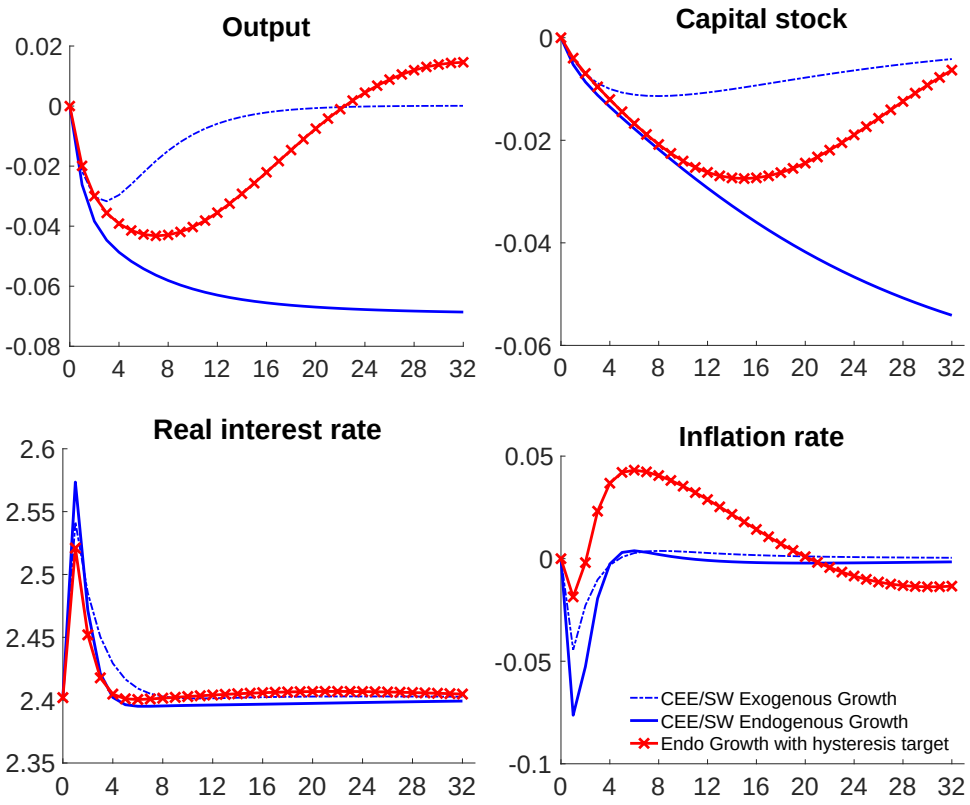
are plotted in percent deviations from an exogenous trend. For real interest rate, we plot the actual path of real interest rate. Inflation is in percent deviation from steady state level. Time is in quarters.

The model replicates the estimated empirical patterns. There is a persistent decline in capital stock, output and TFP. Furthermore, the endogenous growth model exhibits considerable amplification to the transitory shock because of the large hysteresis elasticity.

We define the accumulated gaps in TFP growth rate as the *hysteresis*. We next show the path of output, and capital when the central bank sets interest rates following an augmented Taylor rule:

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

Figure 8: *Response of Output, Capital Stock, Real interest rate and Inflation rate to a 10 bps increase in nominal interest rate*



Notes: The figure plots the model-implied IRFs for output, capital stock, real interest rate and net inflation rate to a transitory shock to the assumed Taylor rule. Solid line reports the IRFs for endogenous growth model with $\eta = 0.10$, and dashed 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 about 10 basis points. The IRFs for output and capital stock are plotted in percent deviations from an exogenous trend. For real interest rate, we plot the actual path of real interest rate. Net inflation rate is in percent deviation from steady state.

where H_t is hysteresis and follows the law motion given by:

$$H_t = H_{t-1} + g_t - g_t^f \,.$$

We set $\phi_H = 0.2$ to plot the graphs. Solid red line with crosses in Figure 8 plots the evolution of output, capital stock, and interest rates under this hysteresis targeting rule. The real interest rate increases by less with hysteresis targeting rule relative to a standard Taylor rule (Equation 14). This is because, under a hysteresis targeting rule, the central bank accommodates above-target inflation at a later time in order to target zero output hysteresis. Expectations of a high inflation rate induce less contraction in the economy as well as a sharp recovery of GDP to the pre-shock trend.[18]

[18] Alternately, a higher weight on inflation stabilization in the policy rule can also reduce the permanent effects of a given monetary shock. An advantage of our simple framework is that a strict inflation targeting prescription of textbook models (Galí, 2015b) is an optimal policy to deal with temporary demand shocks. See Garga and Singh (2016) for detailed analysis from a micro-founded model of TFP growth.

8. Conclusion

This paper challenges the widely accepted benchmark of long-run money neutrality.

To test for a causal relationship in long-run data over more than 140 years in 17 advanced economies we use a trilemma instrument: standard international macroeconomic theory and recent evidence shows that exogenous monetary policy shocks in open pegs can be isolated using unanticipated monetary policy shifts in base (anchor currency) countries.

We show the instrument is strong, and argue that it is valid, but we also account for potential biases. To do this, we develop further the approach of local projections and, in particular, we explain why LPs are preferable to VARs in settings like this where long-horizon responses need to be estimated.

In our annual panel data, using the trilemma-identified monetary policy shocks, we do indeed find evidence of powerful hysteresis forces. Monetary policy shocks have long-run effects on output, capital, and TFP, over a horizon of more than a decade. Notably, the capital stock is found to be permanently lower after a temporary contractionary monetary policy shock identified using instrumental variables.

We reconcile these findings to theory using a simple extension of the Smets and Wouters (2007) medium-scale DSGE model, where TFP growth depends on deviations of output from its flexible-price counterpart. Using the model, we show that a central bank that also targets hysteresis in its policy rule can offset the long-run effects of monetary policy. The analysis of optimal policy in this environment is an important goal for future research.

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.

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.

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

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.

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

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. MIT Press.

Bühlmann, Peter. 1995. Moving-average representation of autoregressive approximations. *Stochastic Processes and their Applications* 60(2): 331–342.

Cerra, Valerie, and Sweta Chaman Saxena. 2008. Growth dynamics: the myth of economic recovery. *American Economic Review* 98(1): 439–457.

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.

Coibion, Olivier. 2012. Are the effects of monetary policy shocks big or small? *American Economic Journal: Macroeconomics* 4(2): 1–32.

Coibion, Olivier, Yuriy Gorodnichenko, and Mauricio Ulate. 2017. The Cyclical Sensitivity in Estimates of Potential Output. Working Paper 23580, National Bureau of Economic Research.

Conley, Timothy G., Christian B. Hansen, and Peter E. Rossi. 2012. Plausibly Exogenous. *The Review of Economics and Statistics* 94(1): 260–272.

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.

Evans, Charles L. 1992. Productivity shocks and real business cycles. *Journal of Monetary Economics* 29(2): 191–208.

Farmer, Roger E. A., and Giovanni Nicolò. 2018. Keynesian economics without the Phillips curve. *Journal of Economic Dynamics and Control* 89: 137–150.

Farmer, Roger E. A., and Konstantin Platonov. 2019. Animal spirits in a monetary model. *European Economic Review* 115: 60–77.

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

Fatás, Antonio, and Lawrence H. Summers. 2018. The permanent effects of fiscal consolidations. *Journal of International Economics* 112: 238–250.

Fernald, John G. 2007. Trend breaks, long-run restrictions, and contractionary technology improvements. *Journal of Monetary Economics* 54(8): 2467–2485.

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

Galí, Jordi. 2015a. Hysteresis and the European unemployment problem revisited. Economics Working Papers 1488, Department of Economics and Business, Universitat Pompeu Fabra.

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

Gali, Jordi, and Tommaso Monacelli. 2005. Monetary policy and exchange rate volatility in a small open economy. *The Review of Economic Studies* 72(3): 707–734.

Garga, Vaishali, and Sanjay R. Singh. 2016. Output Hysteresis and Optimal Monetary Policy. Brown University. Unpublished.

Geman, Stuart, and Chii-Ruey Hwang. 1982. Nonparametric Maximum Likelihood Estimation by the Method of Sieves. *The Annals of Statistics* 10(2): 401–414.

Gonçalves, Sílvia, and Lutz Kilian. 2007. Asymptotic and Bootstrap Inference for AR(∞) Processes with Conditional Heteroskedasticity. *Econometric Reviews* 26(6): 609–641.

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. Central Bank of Chile.

Grenander, Ulf. 1981. *Abstract Inference*. Wiley Inter-Science. New York, N.Y.: Wiley.

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.

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.

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. 2019. The effects of quasi-random monetary experiments. *Journal of Monetary Economics.* Forthcoming.

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.

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.

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.

Moran, Patrick, and Albert Queraltó. 2018. Innovation, productivity, and monetary policy. *Journal of Monetary Economics* .

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

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

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.

Plagborg-Møller, Mikkel, and Christian K. Wolf. 2018. Instrumental Variable Identification of Dynamic Variance Decompositions. Princeton University. Unpublished.

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 0, 71–162. Amsterdam: Elsevier.

Ramey, Valerie A., and Sarah Zubairy. 2018. Government Spending Multipliers in Good Times and in Bad: Evidence from US Historical Data. *Journal of Political Economy* 126(2): 850–901.

Reifschneider, Dave, William Wascher, and David Wilcox. 2015. Aggregate supply in the United States: recent developments and implications for the conduct of monetary policy. *IMF Economic Review* 63(1): 71–109.

Reinhart, Carmen M., and Kenneth S. Rogoff. 2009. The Aftermath of Financial Crises. *American Economic Review* 99(2): 466–472.

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

Romer, Christina D., and David H. Romer. 2004. A New Measure of Monetary Shocks: Derivation and Implications. *American Economic Review* 94(4): 1055–1084.

Schularick, Moritz, and Alan M. Taylor. 2012. Credit Booms Gone Bust: Monetary Policy, Leverage Cycles, and Financial Crises, 1870–2008. *American Economic Review* 102(2): 1029–1061.

Shambaugh, Jay C. 2004. The Effect of Fixed Exchange Rates on Monetary Policy. *The 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.

Stock, James H., and Mark W. Watson. 2018. Identification and Estimation of Dynamic Causal Effects in Macroeconomics Using External Instruments. *Economic Journal* 128(610): 917–948.

Syverson, Chad. 2011. What determines productivity? *Journal of Economic literature* 49(2): 326–65.

van Kippersluis, Hans, and Cornelius A. Rietveld. 2018. Beyond plausibly exogenous. *The Econometrics Journal* 21(3): 316–331.

Wieland, Johannes F., and Mu-Jeung Yang. 2016. Financial Dampening. NBER Working Papers 22141, National Bureau of Economic Research, Inc.

Wold, Herman. 1938. *A study in the analysis of stationary time series*. Uppsala, Sweden: Almqvist & Wiksell.

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.

Appendix

A. Proof of consistency for impulse responses from MA(∞) DGP

This section provides the basic ideas behind the proofs of consistency for truncated VARs and LPs when the true DGP is MA(∞). The reader is referred to the references cited for more details. Here we sketch the main intuition.

A.1. Data generating process and main assumptions

Assume the data generating process for the m–dimensional vector process $\boldsymbol{y}_t$ is as described in section 4, namely

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

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:

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

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 1.

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:

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

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 can be quantified: (1) the truncation lag is too short given Assumptions 1–4; and (2)

even if the truncation lag is appropriate, the impulse response is calculated for periods beyond the truncation lag. To investigate the first source of bias, 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 will come from the term

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

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

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

That is, even if $k = p$, and hence $||\hat{A}_j|| \to A_j$ for $j = 1, \ldots, k$, the fact remains that the omitted term

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

is still unaccounted for.

(Of course, when $k < p$ then one has to add to this problem the bias generated from having too short a VAR relative to the sample, thus potentially compounding the bias problem.)

In practice, it seems advisable to choose relatively long-lags to accurately describe the impulse response. Note that there is a variance-bias trade-off. We are not claiming a lower RMSE, but rather, since interest is in uncovering the behavior of the impulse response at long horizons, we are interested in trading precision for lower bias.

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

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

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

Define $D = (B_h, C_{h+2}, \ldots, C_{h+k})$ 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

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

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

$$\begin{aligned} \frac{1}{T-h-k} \sum_{k}^{T-h} \boldsymbol{\epsilon}_{t+h} X_{t-1}' &\to 0, \\ \frac{B_j}{T-h-k} \sum_{k}^{T-h} \boldsymbol{\epsilon}_{t+h-j} X_{t-1}' &\to 0, \end{aligned}$$

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

$$C_{h+k+j} = B_h A_{k+k} + \cdots + B_1 A_{h+k+j-1} + A_{h+k+j} \qquad j = 0, 1, \ldots,$$

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.

B. Imbs correction

The Imbs (1999) correction follows Burnside and Eichenbaum (1996) in endogenizing the capital utilization rate in a partial equilibrium model. We assume perfectly competitive factor markets and a technology which is constant returns to scale in effective capital and labor.[19]

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)

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

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

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

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

[19] We assume that these variables are stationary. In section 7, we construct capital-utilization in a general equilibrium model along with endogenous growth.

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 steady state values of Y, C, L, and K, we extract a HP-filter trend from the data series. We will show later that 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. We will show robustness of our results to choosing different depreciation parameters.

We present the responses of the various components of the Solow residual in Figure A9 for both the full and post-WW2 samples. Utilization rates of labor and capital exhibit cyclical dynamics: falling in the short-run and then recovering back to zero. In a sense, this pattern is mechanical. It reflects the restriction imposed on the calculation that utilization rates must return to zero eventually (Imbs, 1999).

Figure A10 presents the evolution of capital to utilization adjusted TFP ratio in response to the trilemma shock. The hump-shaped response replicates the conventional justification for including investment adjustment costs in typical DSGE models (Christiano, Eichenbaum, and Evans, 2005).[20]

C. A dynamic open economy model: Implications for the trilemma

We now write down a currency union model of a small open economy based on Gali and Monacelli (2005). There is a continuum of measure zero small open economies.[21] A representative household in an economy has preferences over a consumption aggregator and derives disutility from supplying labor. Final consumption is a CES aggregator over domestic and foreign goods with home bias measured by $1-\alpha \in [0,1]$. The foreign good is a Dixit-Stiglitz aggregator of goods supplied by a continuum of foreign economies with elasticity of substitution among varieties of countries is given by $\gamma > 0$. $\eta > 0$ measures substitutability between domestic and foreign goods. And $\epsilon > 0$ measures elasticity of substitution among varieties produced by one country. Each differentiated product within a country is produced by a monopolistically competitive producer that takes demand as given and sets prices in a staggered fashion ala Calvo price setting. Since we will consider fixed nominal exchange rate, producer or local currency pricing assumptions are equivalent. We assume complete financial markets. We denote the world variables with asterisk. The home economy is too small to influence any world economy variables.

Let $\hat{Y}_t$ be home output, $\hat{C}_t$ be home consumption, $\hat{\pi}_t^H$ be domestic producer price inflation, $\hat{i}_t$ be home nominal interest rate, and $\hat{S}_t$ effective terms of trade (price of foreign goods in terms of home goods) and $\hat{E}_t$ is the nominal exchange rate. Following the seminal work of Gali and Monacelli (2005), we can derive the equilibrium as system of these six variables $\{\hat{Y}_t, \hat{C}_t, \hat{\pi}_t^H, \hat{i}_t, \hat{S}_t, \hat{E}_t\}$ that satisfy the following five equations, plus a sixth equation corresponding to the domestic policy rule, and given foreign output and inflation $\{Y_t^*, \hat{\pi}_t^*\}$, domestic productivity shocks $\hat{\epsilon}_t$, markup shocks $\hat{\mu}_t$,

[20] We later show the robustness of using a more commonly used utilization adjusted TFP series constructed by Fernald (2014) for the post-WW2 US, which requires more detailed data than is available for the historical panel of 17 countries. Reassuringly, the results for utilization adjusted TFP using the two methods are qualitatively similar.

[21] De Paoli (2009)'s model treats foreign as one large economy. Similar results are obtained in that treatment.

government spending shocks $\hat{g}_t$, and natural interest rates $\hat{r}_t^n$:

$$\hat{Y}_t = \mathbb{E}_t \hat{Y}_{t+1} - \frac{1}{\sigma_\alpha}(\hat{i}_t - \mathbb{E}_t \hat{\pi}_{t+1}^H - \hat{r}_t^n)\,, \tag{AD}$$

$$\hat{\pi}_t^H = \beta \mathbb{E}_t \hat{\pi}_{t+1}^H + \kappa((\sigma_\alpha + \varphi)\hat{Y}_t + (\sigma - \sigma_\alpha)\hat{Y}_t^* - (1+\varphi)\hat{e}_t + \hat{\mu}_t)\,, \tag{AS}$$

$$\hat{Y}_t = \hat{C}_t + \frac{\alpha\omega}{\sigma}\hat{S}_t + \hat{g}_t\,, \tag{RC}$$

$$\hat{C}_t = \hat{Y}_t^* + \frac{1-\alpha}{\sigma}\hat{S}_t\,, \tag{RS}$$

$$\Delta\hat{S}_t = \Delta\hat{E}_t + \hat{\pi}_t^* - \hat{\pi}_t^H\,, \tag{TT}$$

where $\kappa = \frac{(1-\beta\theta)(1-\theta)}{\theta}$, where $\kappa(1+\sigma\varphi)$ is slope of the Phillips curve. It is microfounded from a Calvo assumption where $1-\theta$ is the probability of firm price reset (and $\theta \to 0$ implies $\kappa \to \infty$). Here, σ^{-1} is the intertemporal elasticity of substitution, ϵ is intratemporal elasticity of substituion, φ^{-1} is Frisch elasticity of labor supply, and, $\sigma_\alpha = \frac{\sigma}{(1-\alpha)+\alpha\omega}$, $\omega = \sigma\gamma + (1-\alpha)(\sigma\eta - 1)$, and $\Theta = (\sigma\gamma - 1) + (1-\alpha)(\sigma\eta - 1) = \omega - 1$. When $\sigma = \eta = \gamma = 1$, we have $\omega = 1$. When $\alpha \to 0$, we have $\sigma_\alpha \to \sigma$. In the closed economy limit $\alpha \to 0$, the eqn (AD) collapses to the AD equation in a closed economy NK model (Galí, 2015). The natural rate of interest is given by $\hat{r}_t^n = \sigma_\alpha(1+\alpha\Theta)\mathbb{E}_t\{\Delta\hat{g}_{t+1}\} + \sigma_\alpha\alpha\Theta\mathbb{E}_t\{\Delta\hat{Y}_{t+1}^*\}$. The domestic component of the natural rate of interest is given by $\sigma_\alpha(1+\alpha\Theta)\mathbb{E}_t\{\Delta\hat{g}_{t+1}\}$, and the foreign component to the natural rate of interest $\sigma_\alpha\alpha\Theta\mathbb{E}_t\{\Delta\hat{Y}_{t+1}^*\}$ is driven by changes in foreign output growth.

A hard peg policy $\Delta\hat{E}_t = 0$ in the presence of partially flexible prices introduces dynamics in real exchange rate. As shown by Benigno (2004); Benigno, Benigno, and Ghironi (2007); Corsetti, Kuester, and Müller (2013), the past level of the terms of trade (and hence inflation rate) is a state-variable under fixed nominal exchange rate. One can analytically solve for law of motion of terms of trade and obtain the following proposition.[22]

Proposition. *In the Cole-Obstfeld case (log utility and unit trade elasticity, i.e., $\sigma = \eta = \gamma = 1$), a monetary policy shock in world economy identifies an exogenous monetary policy shock in a small open economy that pegs its currency to the union.*

Under the Cole-Obstfeld assumption, there is no movement in net exports of a country. Furthermore, there are no terms of trade movements in response to a monetary policy shock in foreign economy in this case. As a result, a monetary policy shock in the foreign economy acts like a monetary policy shock in the home economy. The spillover effects of trade cancel out, and we can identify a monetary multiplier as in the Mundell-Fleming model presented in the main text.

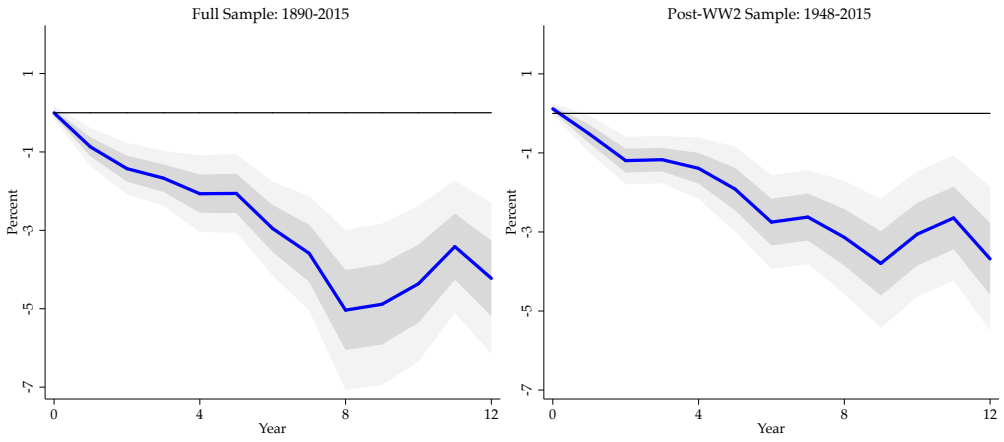
Using calibrated parameters of Gali and Monacelli (2005), we simulate the response of home and foreign economy to a monetary policy shock in the foreign economy in Figure A11. The foreign large economy is modeled as a closed economy in a textbook manner (Galí, 2015). A contractionary shock to the Taylor rule in foreign economy increases foreign real interest rate. As a result, foreign output and inflation fall. Since the home economy follows a hard peg with foreign, the nominal interest rate at home follows identical path as in the foreign economy. There are no changes in terms of trade or net exports. The IRFs of home output and inflation are exactly like that of closed economy's response to a domestic monetary policy shock. The only difference is that there is no systematic response of home's central bank and it maintains a fixed nominal exchange rate.

[22] See also Farhi and Werning (2012); Nakamura and Steinsson (2014); Itskhoki and Mukhin (2019).

D. Additional figures

D.1. Trilemma graphs for GDP per capita

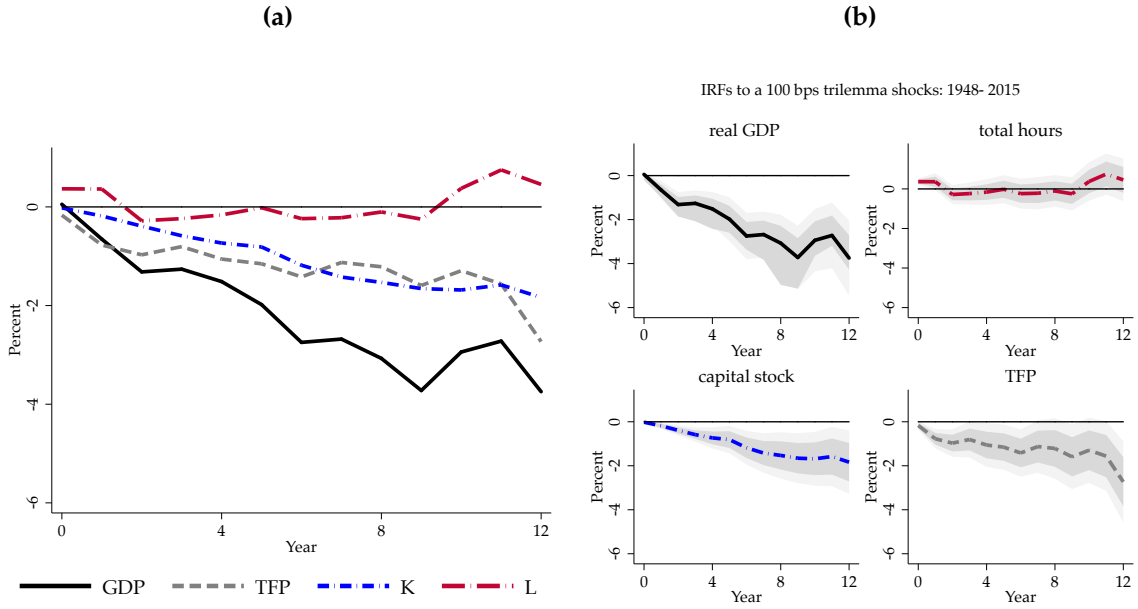
Figure A1: *Baseline response to 100 bps trilemma shock: Real GDP per capita*



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.D. and 2 S.D. confidence bands constructed using cluster-robust standard errors. See text.

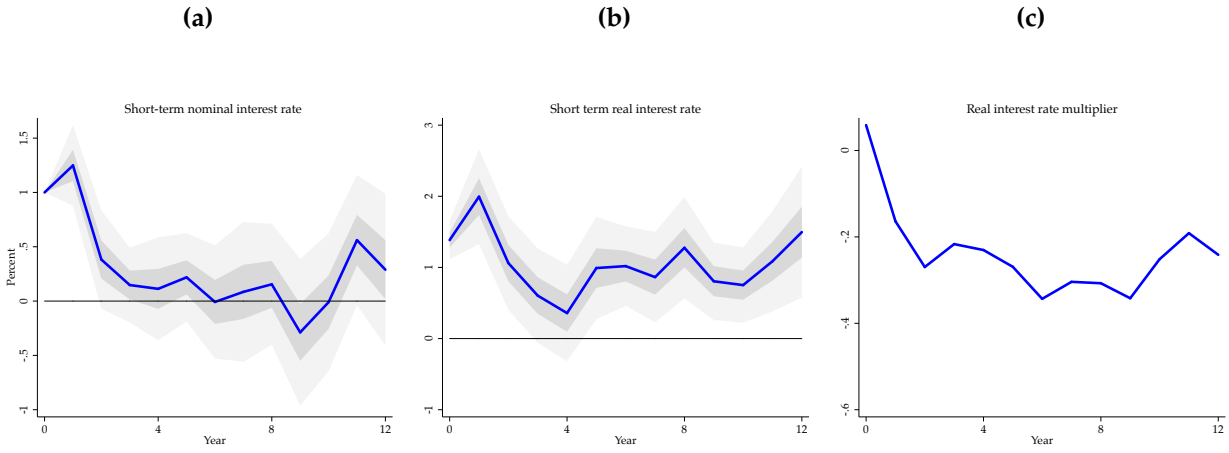
D.2. Trilemma graphs for Post WW2 sample

Figure A2: *Baseline response to 100 bps trilemma shock: Real GDP and components (Post ww2 sample)*



Notes: Response to a 100 bps shock in domestic interest rate instrumented with the trilemma. Responses for pegging economies. Post-WW2 sample: 1948–2015. See text.

Figure A3: *Baseline response to 100 bps trilemma shock: a) Short term nominal interest rate, b) real interest rate, and c) real interest rate multiplier*

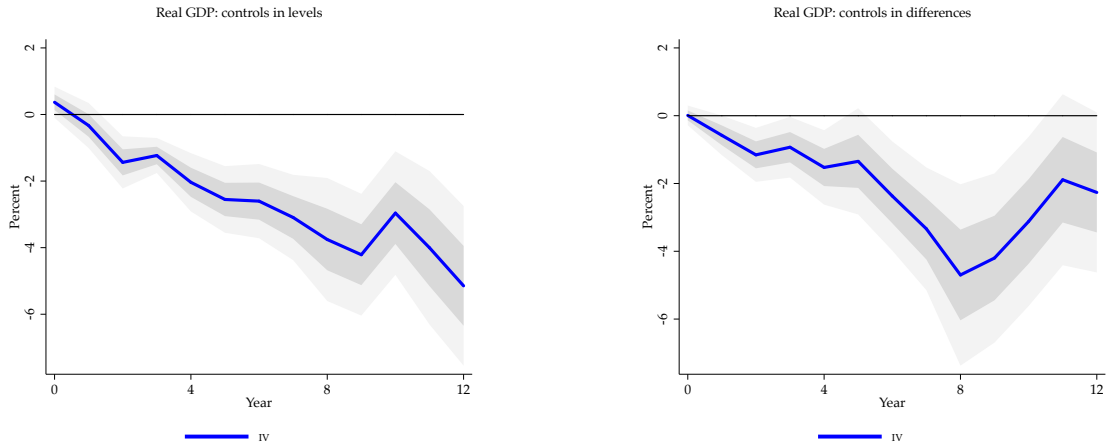


Notes: Response to a 100 bps shock in domestic interest rate instrumented with the trilemma. a) Responses of short term nominal interest rate for pegging economies. b) Responses of short term real interest rate for pegging economies. Inflation expectations constructed from the impulse response of consumer price level index to the same trilemma shock. c) Response of cumulative change in real GDP divided by cumulative change in short term nominal interest rate for pegging economies. Sample: 1948–2015. See text.

D.3. Trilemma graphs: 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.

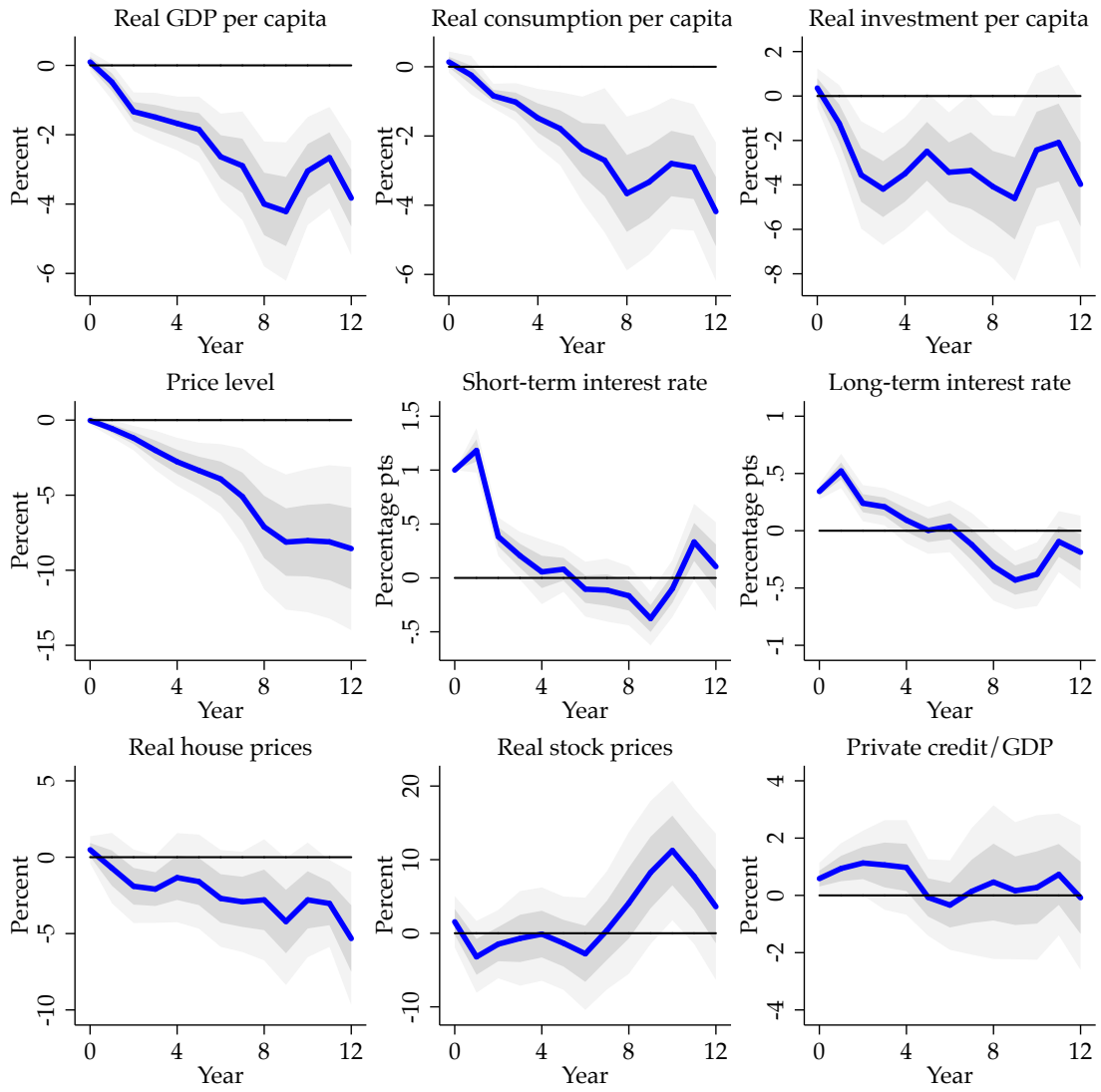
Figure A4: *Response to 100 bps trilemma shock with controls in levels: Real GDP*



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-OLS estimates displayed as a dashed red line, LP-IV estimates displayed as a solid blue line and 1 S.D. and 2 S.D. confidence bands. See text and Jordà, Schularick, and Taylor (2019).

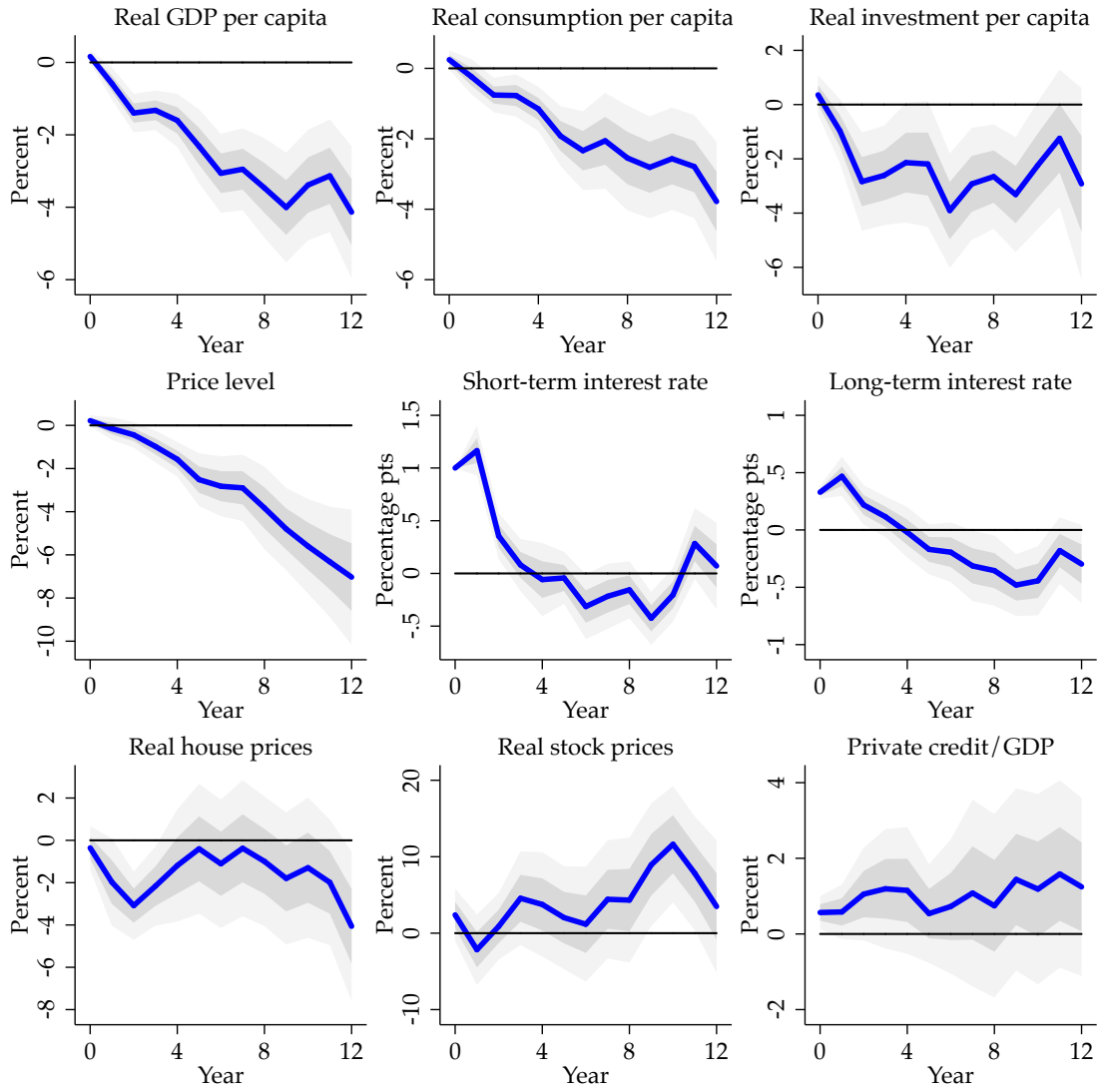
D.4. Trilemma graphs for various macro-financial variables

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 deviations using cluster-robust standard errors. See text.

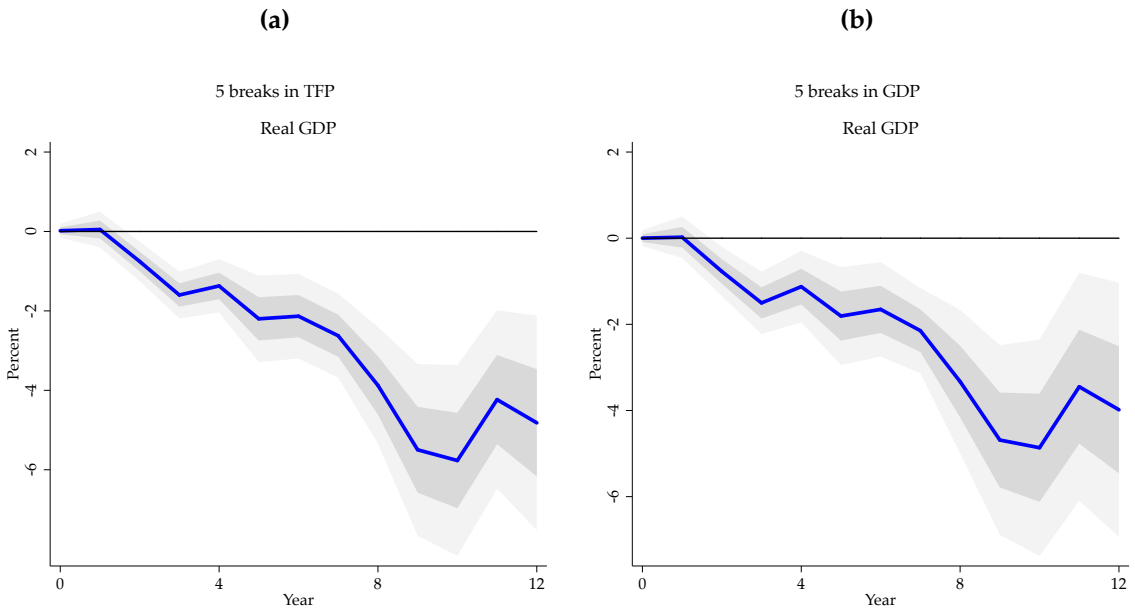
Figure A6: *Baseline response to 100 bps trilemma shock: Post-WW2 Sample (1948-2015)*



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 deviations using cluster-robust standard errors. See text.

D.5. Trilemma graphs with structural breaks

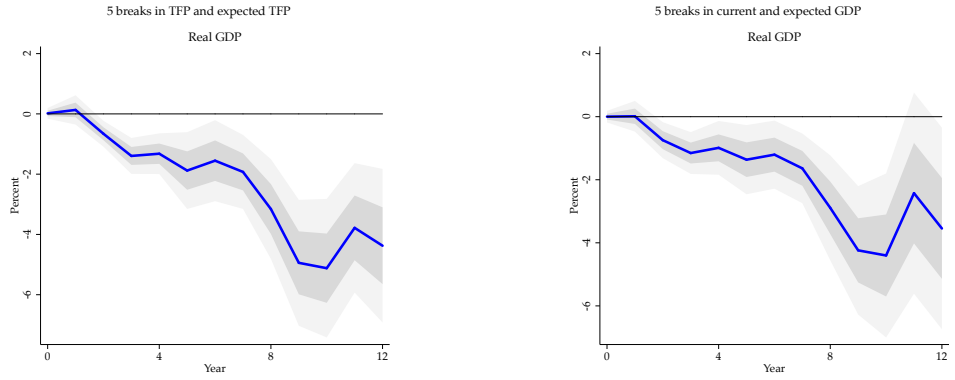
Figure A7: *Response to 100 bps trilemma shock with structural breaks: Real GDP*



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-OLS estimates displayed as a dashed red line, LP-IV estimates displayed as a solid blue line and 1 S.D. and 2 S.D. confidence bands. See text and Jordà, Schularick, and Taylor (2019).

We next show an alternate selection of structural breaks where we allow agents to anticipate structural breaks at upto 12 years ahead starting at time 0.

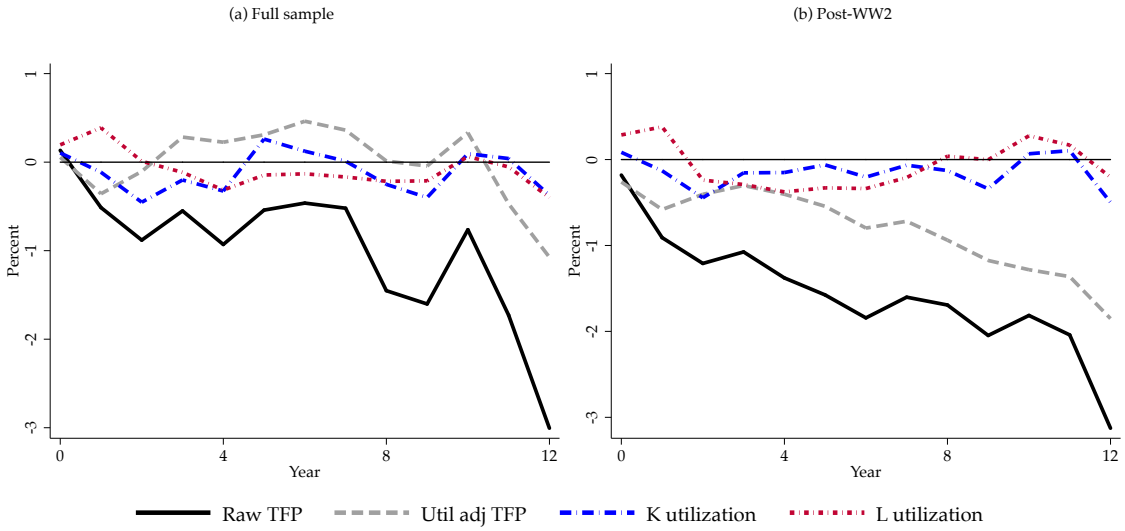
Figure A8: *Response to 100 bps trilemma shock with anticipated structural breaks: Real GDP*



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-OLS estimates displayed as a dashed red line, LP-IV estimates displayed as a solid blue line and 1 S.D. and 2 S.D. confidence bands. See text and Jordà, Schularick, and Taylor (2019).

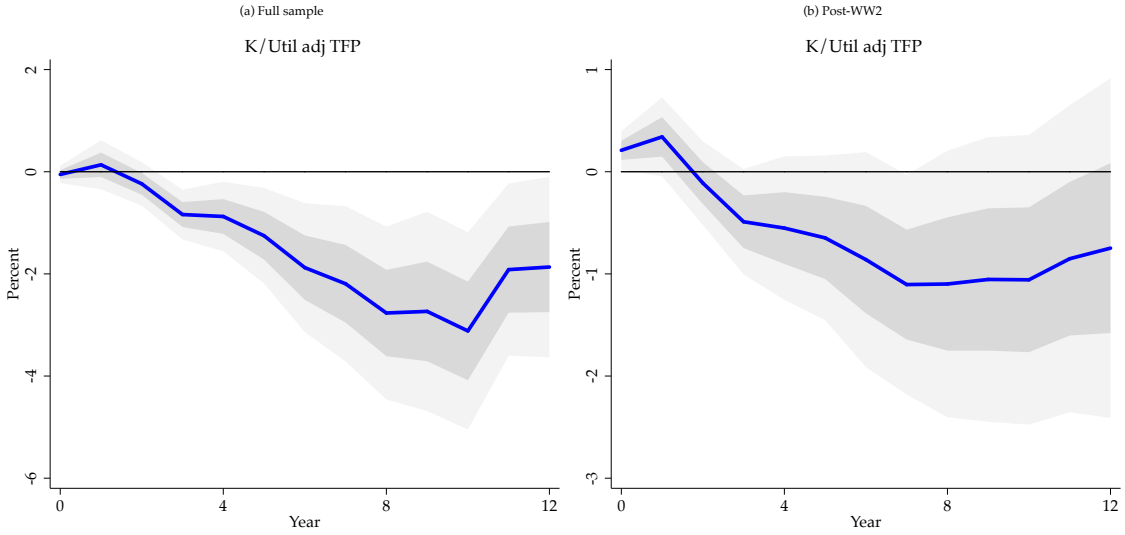
D.6. Trilemma graphs for TFP and utilization adjusted TFP

Figure A9: *Baseline response to 100 bps trilemma shock: TFP and utilization*



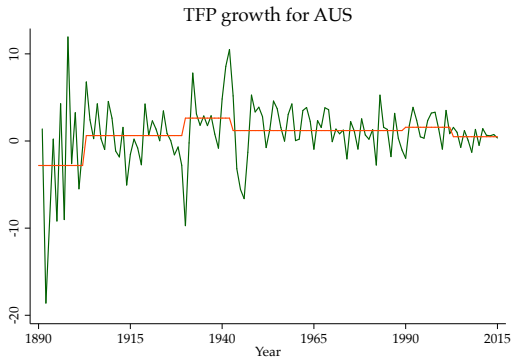
Notes: Response to a 100 bps shock in domestic interest rate instrumented with the trilemma. Responses for pegging economies. Full sample: 1890–2015 (World Wars excluded). Post-WW2 sample: 1948–2015. See text.

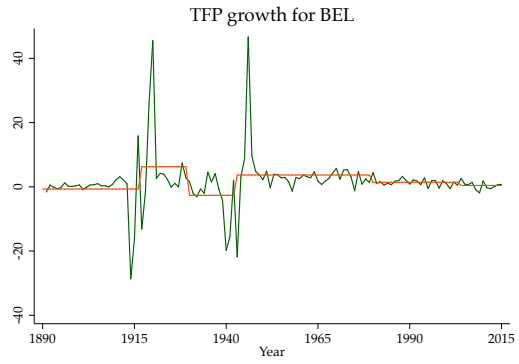
Figure A10: *Baseline response to 100 bps trilemma shock: Capital to utilization adjusted TFP*

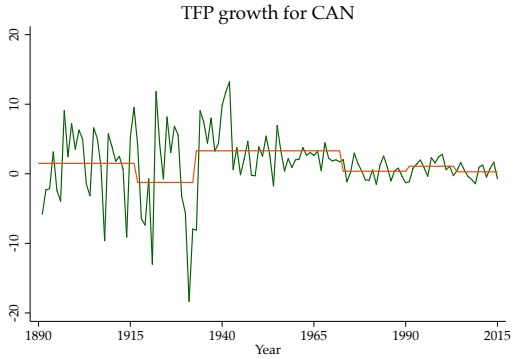


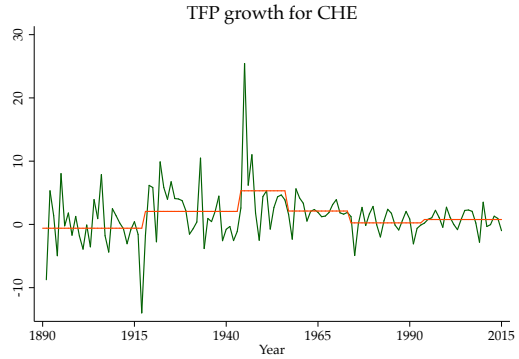
Notes: Response to a 100 bps shock in domestic interest rate instrumented with the trilemma. Responses for pegging economies. Full sample: 1890–2015 (World Wars excluded). Post-WW2 sample: 1948–2015. See text.

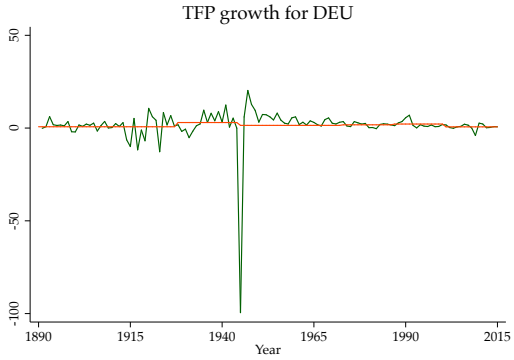
D.7. Structural break dates in TFP growth and GDP growth

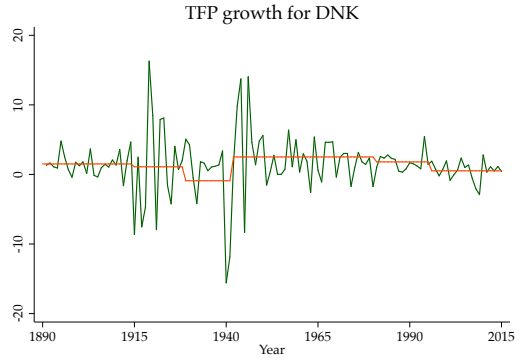


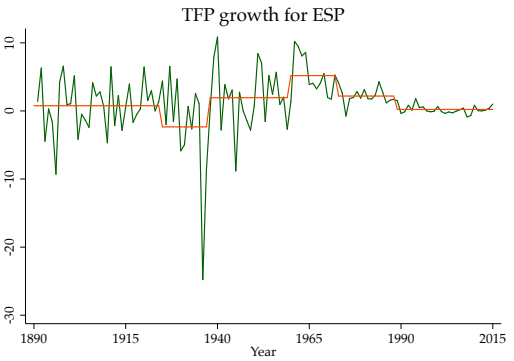


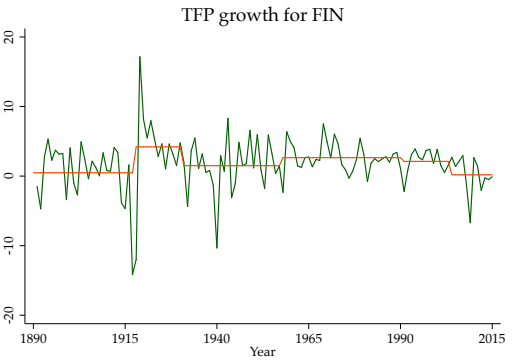


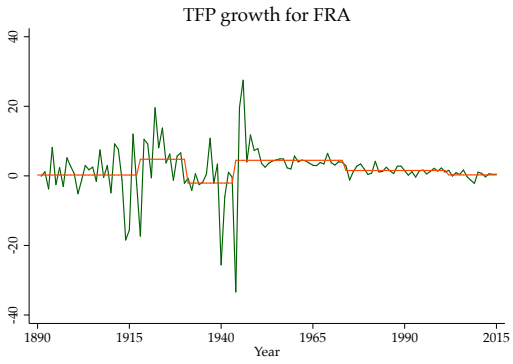


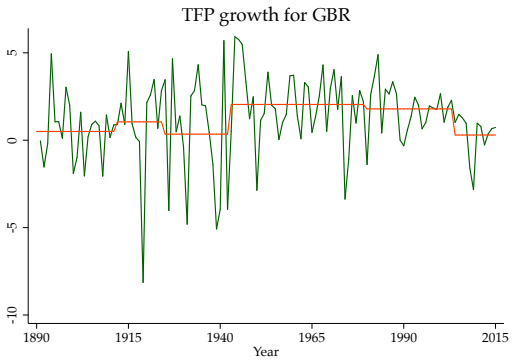


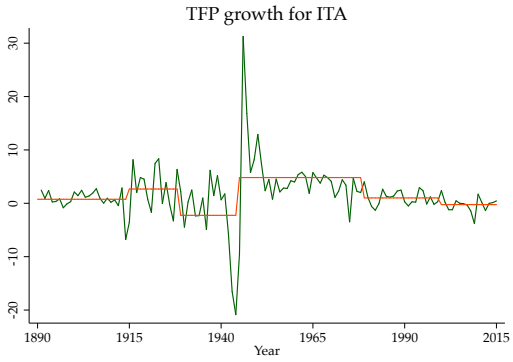


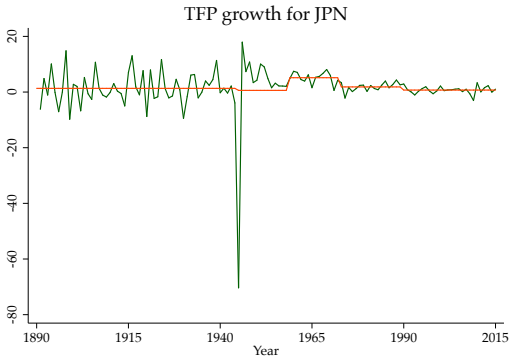


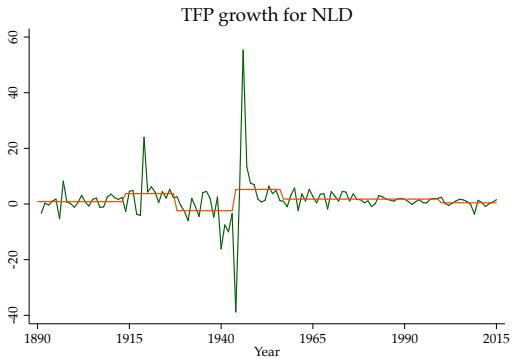


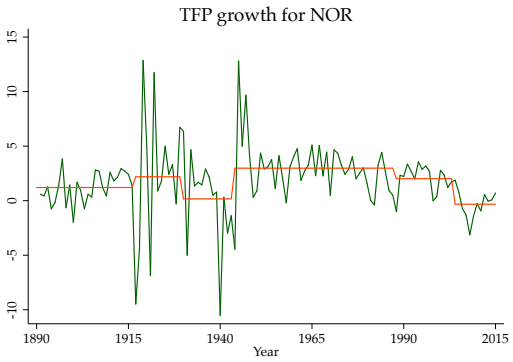

TFP growth for FRA
40
20
0
-20
-40
1890
1915
1940
1965
1990
2015
Year

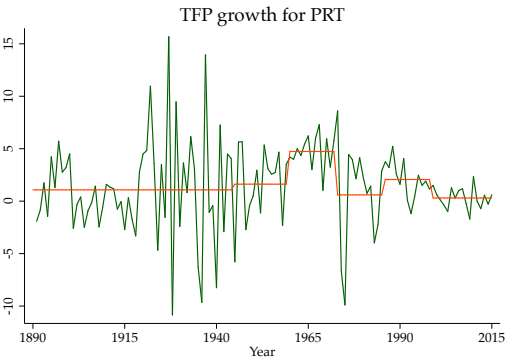

TFP growth for GBR
5
0
-5
-10
1890
1915
1940
1965
1990
2015
Year

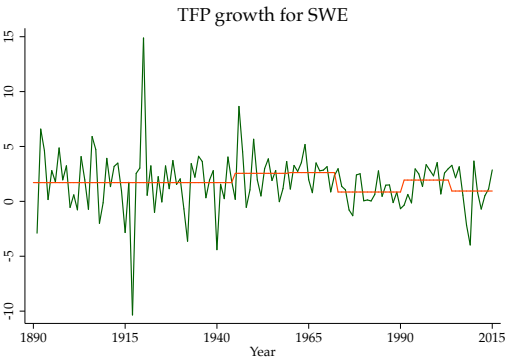

TFP growth for ITA
30
20
10
0
-10
-20
1890
1915
1940
1965
1990
2015
Year

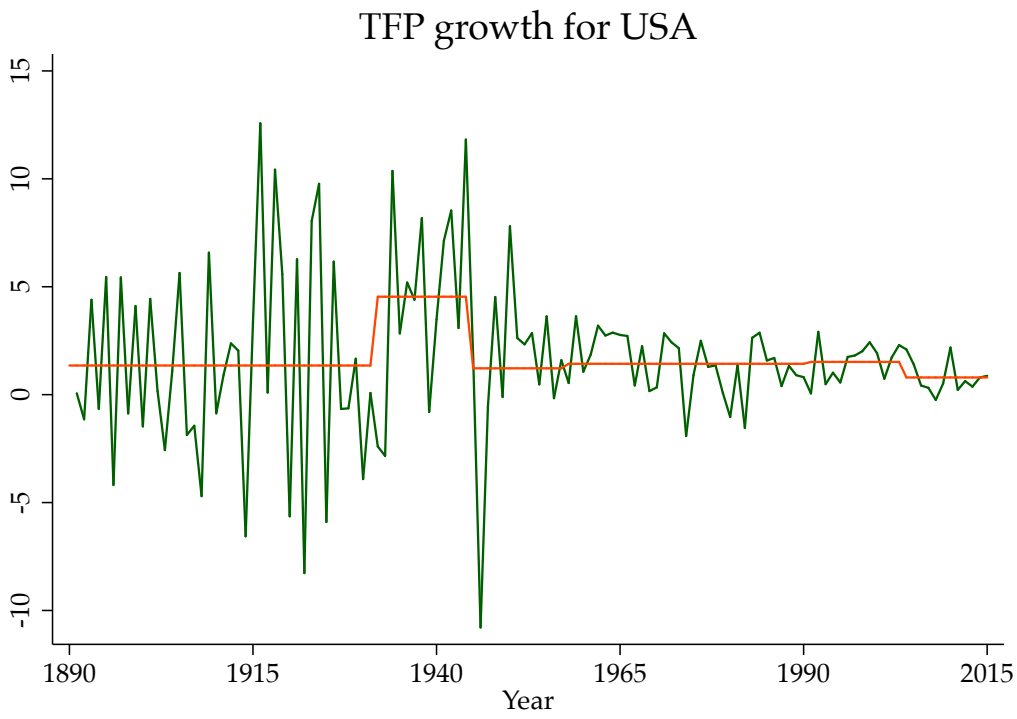

TFP growth for JPN
20
0
-20
-40
-60
-80
1890
1915
1940
1965
1990
2015
Year

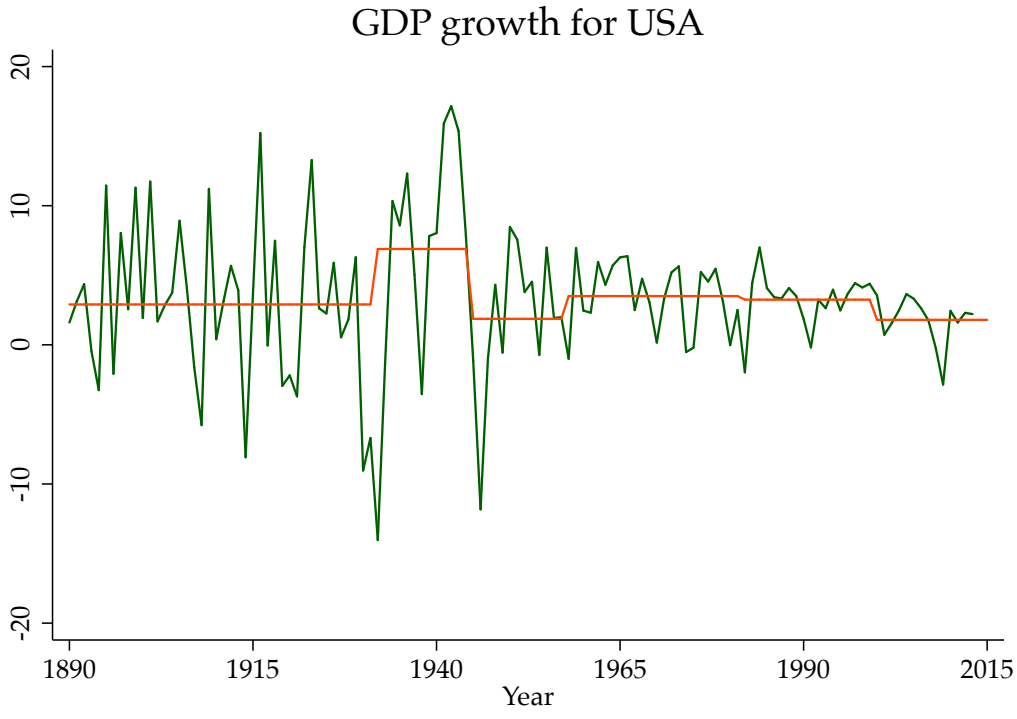

TFP growth for NLD
60
40
20
0
-20
-40
1890
1915
1940
1965
1990
2015
Year


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

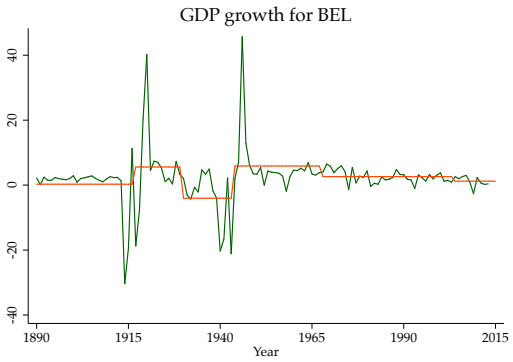

TFP growth for PRT
15
10
5
0
-5
-10
1890
1915
1940
1965
1990
2015
Year

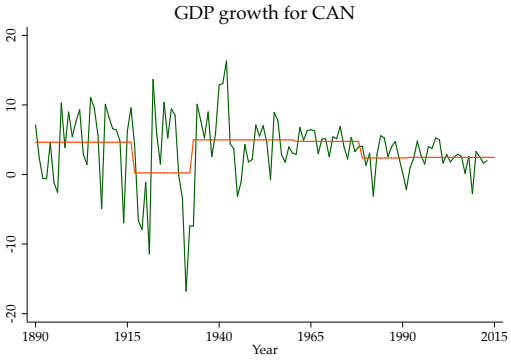

TFP growth for SWE
15
10
5
0
-5
-10
1890
1915
1940
1965
1990
2015
Year

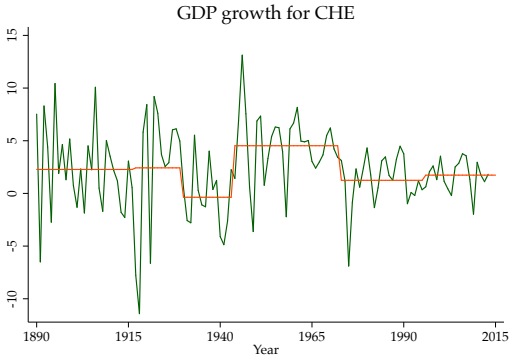

TFP growth for USA
15
10
5
0
-5
-10
1890
1915
1940
1965
1990
2015
Year

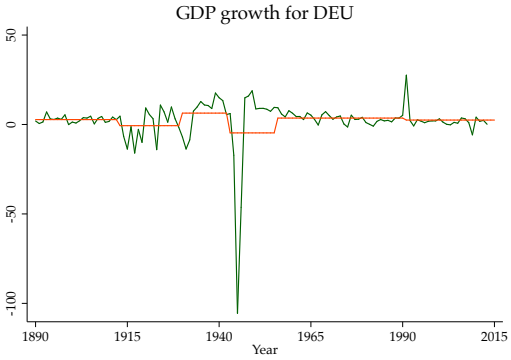

GDP growth for USA
20
10
0
-10
-20
1890
1915
1940
1965
1990
2015
Year

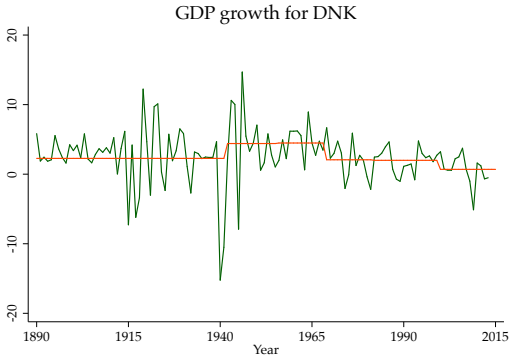

GDP growth for AUS
20
10
0
-10
-20
1890
1915
1940
1965
1990
2015
Year


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

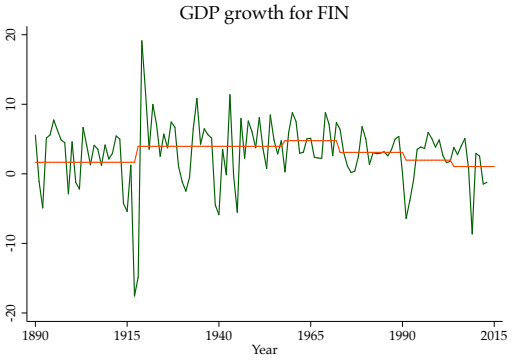

GDP growth for CAN
20
10
0
-10
-20
1890
1915
1940
1965
1990
2015
Year

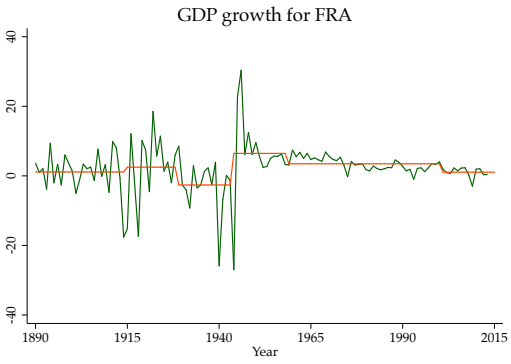

GDP growth for CHE
15
10
5
0
-5
-10
1890
1915
1940
1965
1990
2015
Year

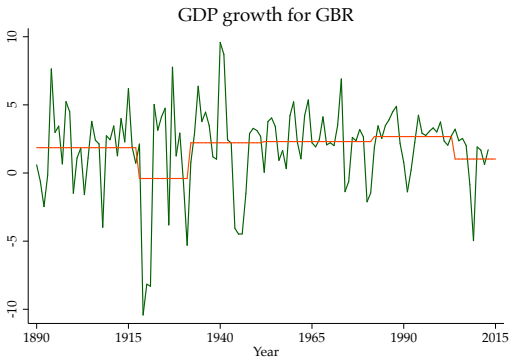

GDP growth for DEU
50
0
-50
-100
1890
1915
1940
1965
1990
2015
Year


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

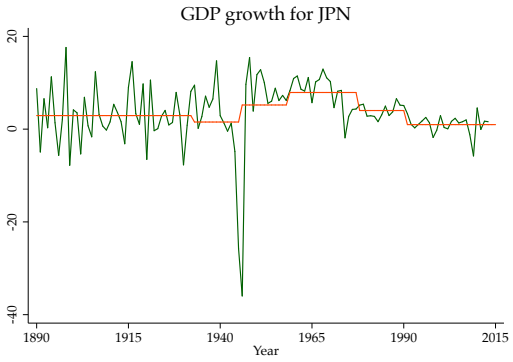

GDP growth for ESP
10
0
-10
-20
-30
1890
1915
1940
1965
1990
2015
Year

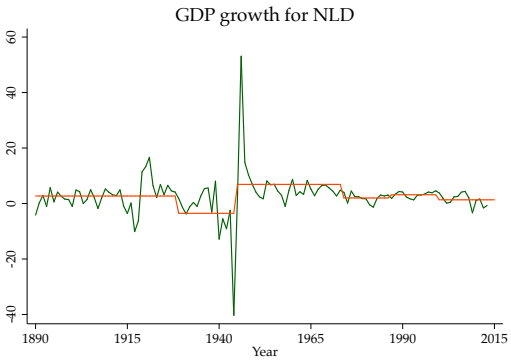

GDP growth for FIN
20
10
0
-10
-20
1890
1915
1940
1965
1990
2015
Year

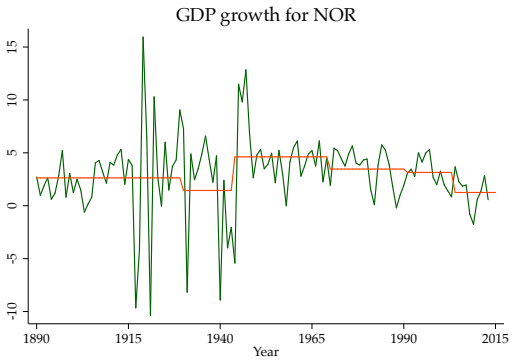

GDP growth for FRA
40
20
0
-20
-40
1890
1915
1940
1965
1990
2015
Year

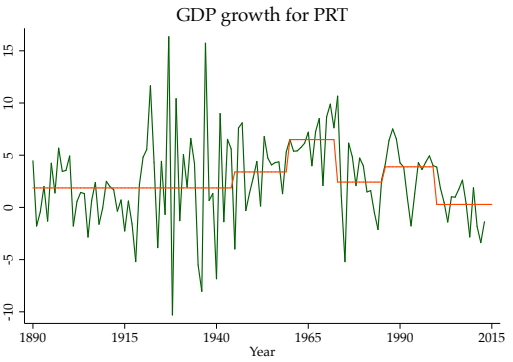

GDP growth for GBR
10
5
0
-5
-10
1890
1915
1940
1965
1990
2015
Year

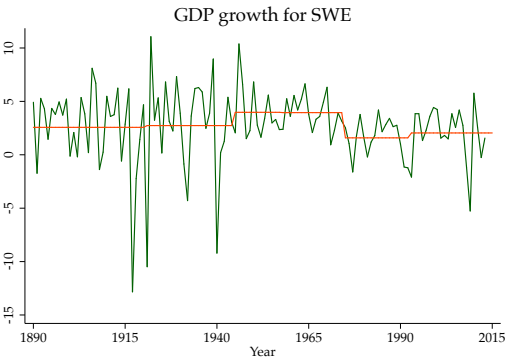

GDP growth for ITA
20
10
0
-10
-20
-30
1890
1915
1940
1965
1990
2015
Year


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

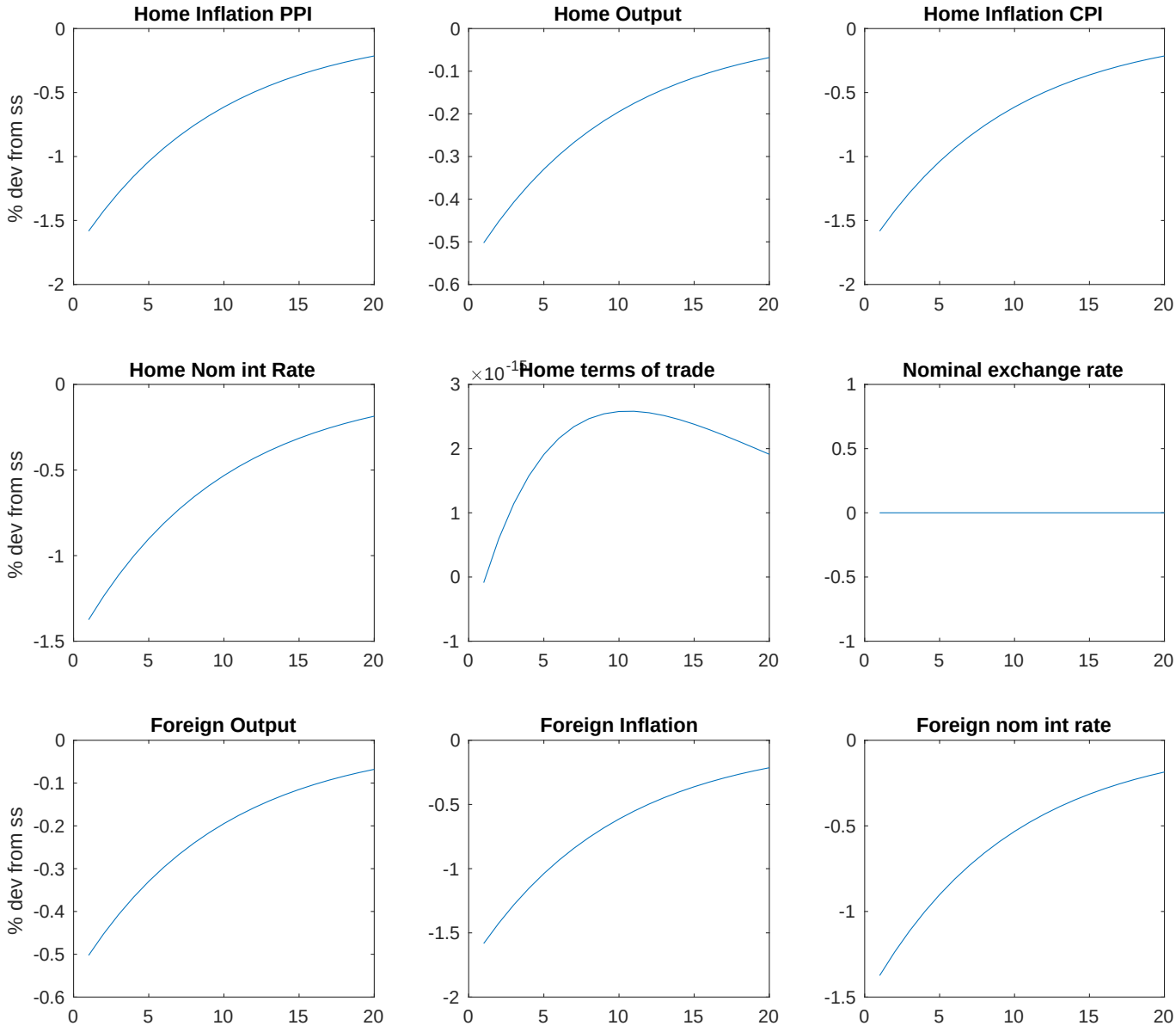

GDP growth for NOR
15
10
5
0
-5
-10
1890
1915
1940
1965
1990
2015
Year


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

D.8. Graph for Gali and Monacelli (2005) open economy model

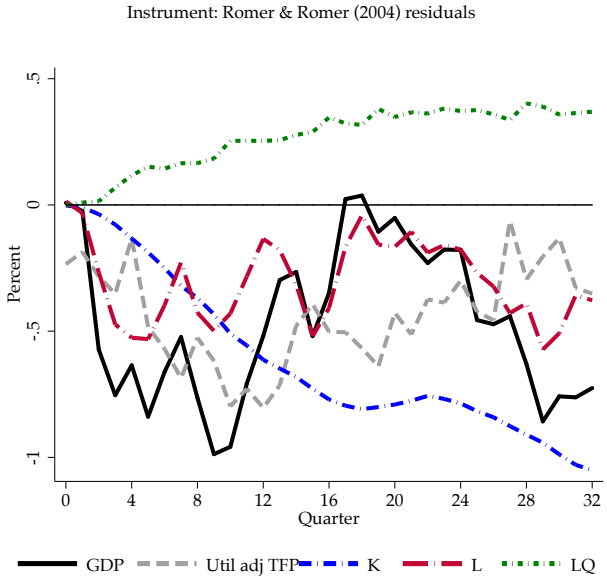
Figure A11: *Response to monetary policy shock in the base country*



Notes: Impulse response to a monetary policy shock in foreign economy using Gali and Monacelli (2005) parameters.

D.9. Graphs for Romer and Romer (2004): Disaggregation of GDP

Figure A12: *Baseline response to 100 bps Romer and Romer (2004) shock: U.S. postwar data*

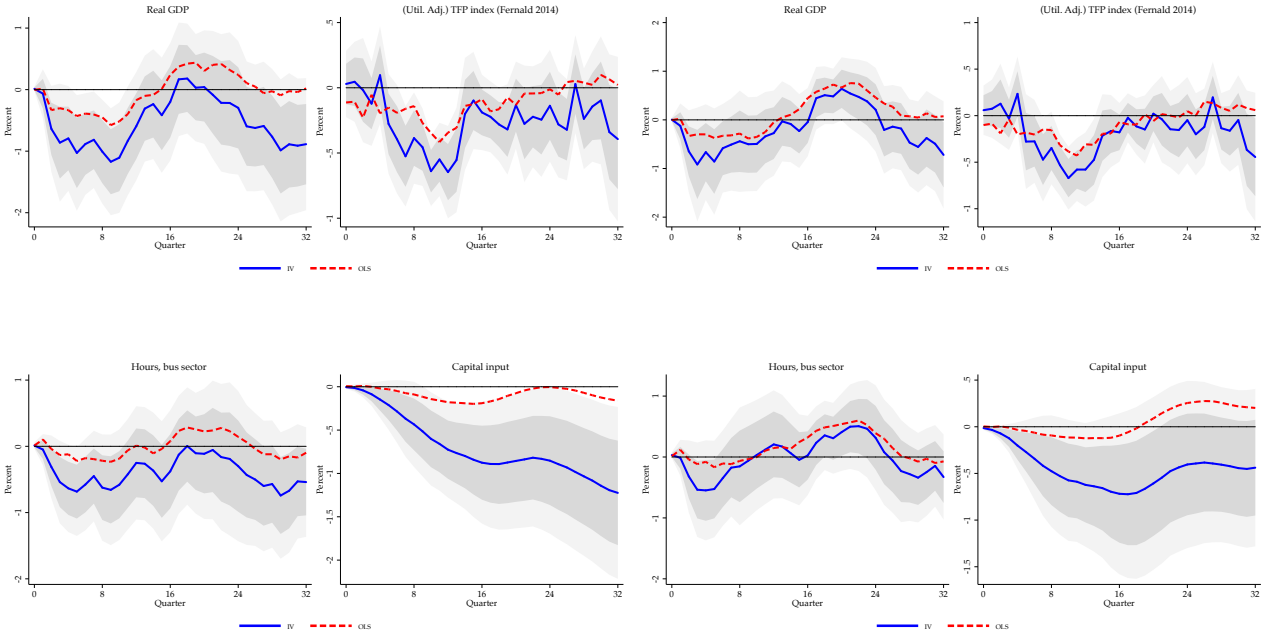


Notes: Response to a 100 bps shock in overnight federal funds rate rate instrumented with policy forecast residuals (Romer and Romer, 2004; Wieland and Yang, 2016). Responses of real GDP, capital stock, total hours worked, labor composition/quality, and raw TFP for U.S. Quarterly sample: 1969-Q1: 2007-Q4. Quarterly data series are taken from Fernald (2014). See text.

D.10. Graphs for Romer and Romer (2004) Surprises: Different sample selection for US

Figure A13: *Response to 100 bps Romer and Romer (2004) shock*

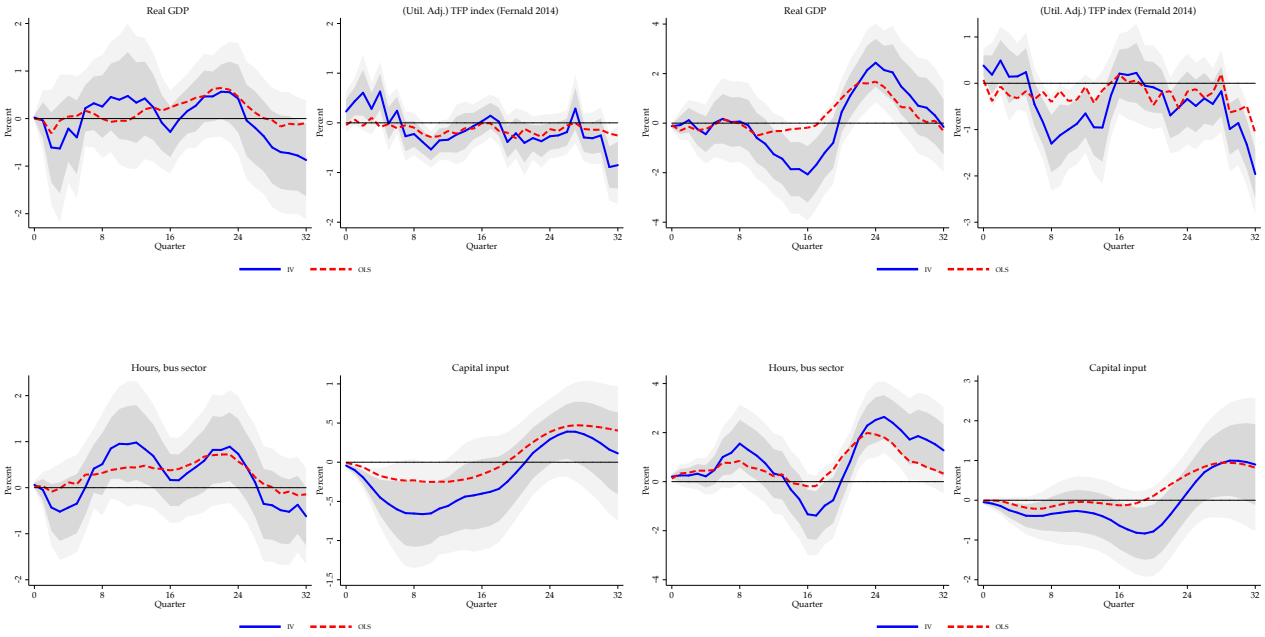
(a) *Full sample: 1969Q2: 2008Q3* **(b)** *Sample: 1973Q2: 2008Q3*



Notes: Response to a 100 bps shock in federal funds rate rate instrumented with policy forecast residuals (Romer and Romer, 2004; Wieland and Yang, 2016). Responses of real GDP, utilization-adjusted TFP (Fernald, 2014), capital stock and hours worked for U.S. economy. Quarterly samples: a) 1969-Q2: 2008-Q3, b) 1973-Q2: 2008-Q3. Quarterly data series are taken from Fernald (2014).

Figure A14: *Response to 100 bps Romer and Romer (2004) shock*

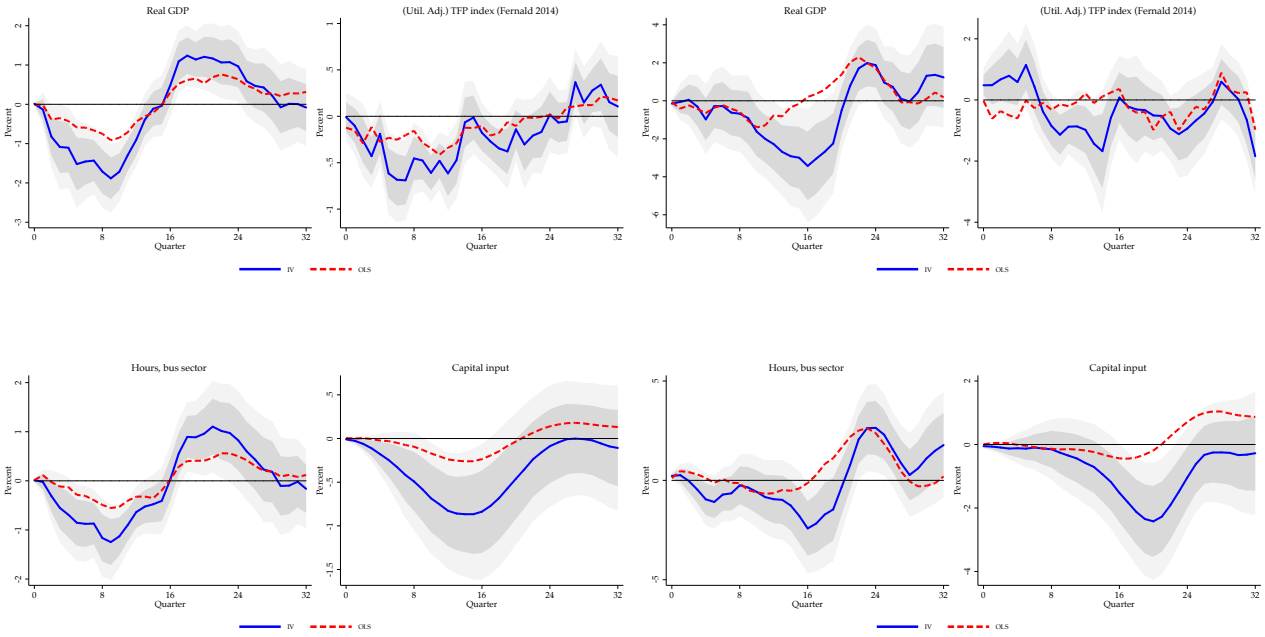
(a) *Sample: 1979Q3: 2008Q3* **(b)** *Sample: 1984Q1: 2008Q3*



Notes: Response to a 100 bps shock in federal funds rate rate instrumented with policy forecast residuals (Romer and Romer, 2004; Wieland and Yang, 2016). Responses of real GDP, utilization-adjusted TFP (Fernald, 2014), capital stock and hours worked for U.S. economy. Quarterly samples: a) 1979-Q3: 2008-Q3, b) 1984-Q1: 2008-Q3. Quarterly data series are taken from Fernald (2014).

Figure A15: *Response to 100 bps Romer and Romer (2004) shock*

(a) *Sample: 1969Q2: 2002Q4* **(b)** *Sample: 1987Q1: 2008Q3*



Notes: Response to a 100 bps shock in federal funds rate rate instrumented with policy forecast residuals (Romer and Romer, 2004; Wieland and Yang, 2016). Responses of real GDP, utilization-adjusted TFP (Fernald, 2014), capital stock and hours worked for U.S. economy. Quarterly samples: a) 1969-Q2: 2002-Q4, b) 1987-Q1: 2008-Q3. Quarterly data series are taken from Fernald (2014).

E. Additional tables

Table A1: *LP-OLS vs. LP-IV. Attenuation bias of real GDP per capita responses to interest rates. Trilemma instrument. Matched samples*

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)
$h = 0$	0.07**	-0.01	0.35	0.03	0.11*	0.22
	(0.03)	(0.08)		(0.02)	(0.07)	
$h = 2$	-0.33**	-1.55***	0.00	-0.32**	-1.40***	0.00
	(0.15)	(0.29)		(0.13)	(0.30)	
$h = 4$	-0.22	-2.07***	0.00	-0.29	-1.64***	0.00
	(0.24)	(0.40)		(0.22)	(0.41)	
$h = 6$	-0.10	-2.76***	0.00	-0.19	-2.93***	0.00
	(0.36)	(0.57)		(0.30)	(0.59)	
$h = 8$	-0.30	-4.71***	0.00	-0.01	-3.13***	0.00
	(0.34)	(0.98)		(0.33)	(0.67)	
$h = 10$	0.10	-4.27***	0.00	0.21	-3.03***	0.00
	(0.34)	(0.96)		(0.37)	(0.75)	
$h = 12$	0.21	-4.02***	0.00	0.37	-3.72***	0.00
	(0.35)	(0.96)		(0.39)	(0.94)	
KP weak IV		44.40			54.07	
H_0: $LATE = 0$	0.00	0.00		0.01	0.00	
Observations	838	765		658	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.

F. DSGE MODEL

We present the different components of the model in the next few sections, starting with the production side of the economy.

F.1. Monopolistically competitive producers

Assume there is a continuum of differentiated intermediated good producers that sell the intermediate good Y_{kt}. These goods can be aggregated into a Dixit-Stiglitz final composite good Y_t as follows:

$$Y_t = \left[\int_0^1 Y_{kt}^{\frac{1}{1+\lambda_p}} dk \right]^{1+\lambda_p},$$

where $\lambda_p > 0$ is the price markup. The iso-elastic demand for intermediate good k is given by:

$$Y_{kt} = \left(\frac{P_{kt}}{P_t} \right)^{-\frac{1+\lambda_p}{\lambda_p}} Y_t \,.$$

The zero profit condition for competitive final good producers implies that the aggregate price index is

$$P_t = \left[\int_0^1 P_{kt}^{-\frac{1}{\lambda_p}} dk \right]^{-\lambda_p}.$$

Each intermediate good k is produced by a price-setting monopolist using labor L_{kt} and physical capital K_{kt}[23]:

$$Y_{kt} = (Z_t L_{kt})^{1-\alpha} \; K_{kt}^{\alpha} \,,$$

where Z_t is the aggregate TFP. The variable Z_t denotes a non-stationary TFP series that we describe in the next subsection.[24] 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.

$$W_t = (1-\alpha)\; mc_{kt} Z_t^{1-\alpha} \left(\frac{K_{kt}}{L_{kt}} \right)^{\alpha},$$

$$R_t^k = \alpha \; mc_{kt} Z_t^{1-\alpha} \left(\frac{K_{kt}}{L_{kt}} \right)^{\alpha-1}.$$

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

$$\frac{K_{kt}}{L_{kt}} = \frac{K_t}{L_t} = \frac{\alpha}{1-\alpha} \frac{W_t}{R_t^k},$$

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

$$P_t mc_{kt} = P_t mc_t = \frac{1}{Z_t^{1-\alpha}} \left(\frac{R_t^k}{\alpha} \right)^{\alpha} \left(\frac{W_t}{1-\alpha} \right)^{1-\alpha}.$$

[23] We can append fixed cost in the production function to eliminate steady state profits. This is usually done to justify no entry and exit in the steady state in a DSGE model (Christiano, Eichenbaum, and Evans, 2005; Justiniano, Primiceri, and Tambalotti, 2013; Smets and Wouters, 2007). In a micro-founded model of growth, positive rents are needed to incentivize investment in growth (Romer, 1990).

[24] Relative to the conventional literature, we will allow for an endogenous relationship between TFP growth and output gap.

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. A resetting firm chooses $\tilde{P}_{kt}$ to maximize:

$$\mathbb{E}_t \sum_{s=t}^{\infty} \theta_p^{s-t} Q_{t,s} \left[\frac{\tilde{P}_{kt}}{P_s} - \ mc_s \right] Y_s(k) \,,$$

subject to demand for its product

$$Y_{kt} = \left(\frac{P_{kt}}{P_t} \right)^{-\frac{1+\lambda_p}{\lambda_p}} Y_t \,,$$

where the stochastic discount factor in period s relative to period t is given by:

$$Q_{t,s} = \beta^{s-t} \frac{\Lambda_s}{\Lambda_t} \frac{P_t}{P_s} \,,$$

and Λ_t is the marginal utility of consumption defined later. The first order condition is :

$$\mathbb{E}_t \sum_{s=t}^{\infty} \theta_p^{s-t} Q_{t,s} \left[\frac{\tilde{P}_t}{\Pi_{t,s}} - (1+\lambda_p)\ mc_s \right] \left(\frac{\tilde{P}_t}{\Pi_{t,s}} \right)^{-\frac{1+\lambda_p}{\lambda_p}} Y_s = 0 \,.$$

By the law of large numbers, the law of motion of the aggregate price index P_t is given by:

$$P_t^{\frac{1}{\lambda_p}} = (1-\theta_p)(\tilde{P}_t)^{\frac{1}{\lambda_p}} + \theta_p P_{t-1}^{\frac{1}{\lambda_p}} \,.$$

F.2. Households

Rest of the model components are standard. We briefly summarize them here and leave the detailed derivations to the appendix.

Households

Each household supplies differentiated labor indexed by j. Household j chooses consumption C_t, risk-free nominal bonds B_t, investment I_t and capital utilization u_t to maximize the utility function, with external habits over consumption:

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

where h is the degree of habit formation, $\nu > 0$ is the inverse Frisch elasticity of labor supply, $\omega > 0$ 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

$$P_t C_t + P_t I_t + B_{t+1} = B_t(1+i_t) + B_{j,t}^S + (1+\tau^w) W_t L_{j,t} + \Gamma_t + T_t + R_t^K u_t K_t^u - P_t a(u_t) K_t^u \,, \qquad (11)$$

where I_t is investment, $B_{j,t}^S$ is the net cash-flow from household j's portfolio of state-contingent securities. Labor income $W_t L_{j,t}$ is subsidized at a fixed rate τ_w. Households own an equal share of all firms, and thus receive Γ_t dividends from profits. Finally, each household receives a lump-sum government transfer T_t. 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_t^K is :

$$K_t = u_t K_t^u \,,$$

The (nominal) cost of capital utilization is $P_t a(u_t)$ 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:

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

where $g_{ss} \equiv \bar{\mu}$ is the steady state growth rate of Z_t. Utility maximization delivers the first order condition linking the inter-temporal consumption smoothing to the marginal utility of holding the risk-free bond

$$1 = \beta \mathbb{E}_t \left[\frac{\Lambda_{t+1}}{\Lambda_t}\left(1+i_t\right)\frac{P_t}{P_{t+1}}\right] . \tag{12}$$

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

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

where Λ_t is the marginal utility of consumption given by:

$$\Lambda_t = \frac{1}{C_t - hC_{t-1}} - \frac{h\beta}{C_{t+1} - hC_t} \,.$$

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:

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

where the (relative) price of installed capital q_t is given by

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

Choice of capital utilization rate yields:

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

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:

$$L_t = \left[\int_0^1 L_{j,t}^{\frac{1}{1+\lambda_w}} dj\right]^{1+\lambda_w},$$

where $\lambda_w > 0$ is the nominal wage markup. 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:

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

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.[25] The first order condition for this problem is given by:

$$\mathbb{E}_t \sum_{s=0}^{\infty} (\beta\theta_w)^s \Lambda_{t+s} \left[(1+\tau_t^W)\tilde{W}_t - (1+\lambda_w)\omega \frac{L_{j,t+s}^{\nu}}{\Lambda_{t+s}}\right] L_{j,t+s} = 0\,. \tag{13}$$

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:

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

where the nominal wage inflation and price inflation are related to each other :

$$\pi_t^w = \frac{W_t}{W_{t-1}} = \frac{w_t}{w_{t-1}} \frac{1}{\pi_t} \frac{1}{1+g_t},$$

where $\pi_t \equiv \frac{P_t}{P_{t-1}}$ is the inflation rate, $w_t \equiv \frac{W_t}{P_t Z_t}$ is the productivity adjusted real wage and g_t is the growth rate of Z_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.

$$\frac{1+i_t}{1+i_{ss}} = \left(\frac{1+i_{t-1}}{1+i_{ss}}\right)^{\rho_R} \left[\left(\frac{\pi_t}{\pi_{ss}}\right)^{\phi_\pi} \left(\frac{Y_t}{Y_t^{f,t}}\right)^{\phi_y}\right]^{1-\rho_R} \epsilon_t^{mp}, \tag{14}$$

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

[25] We assume imperfect wage indexation in our nominal wage rigidity assumption. We ignore specifying it here for ease of exposition. See appendix for details.

$\epsilon_t^{mp} \sim N(0, \sigma_r)$ is the monetary policy shock.

We assume government balances budget every period:

$$P_t T_t = \tau^p \int_0^1 p_{it} x_{it} di + \tau^w W_t L_t + P_t G_t \,,$$

where G_t is the government spending, which is determined exogenously as as a fraction of GDP

$$G_t = \left(1 - \frac{1}{\lambda_t^g}\right) Y_t \,,$$

and where the government spending shock follows the process:

$$\log \lambda_t^g = (1 - \rho_g)\lambda^g + \rho_g \log \lambda_{t-1}^g + \epsilon_t^g; \quad \epsilon_t^g \sim N(0, \sigma_g) \,,$$

F.4. Market clearing

Market clearing requires

$$Y_t = C_t + I_t + a(u_t) K_t^u + G_t \,.$$

F.5. Stationary allocation

We normalize the following variables :

$$\begin{aligned}
y_t &= Y_t / Z_t \,, \\
c_t &= C_t / Z_t \,, \\
k_t &= K_t / Z_t \,, \\
k_t^u &= K_t^u / Z_{t-1} \,, \\
\mathbb{I}_t &= I_t / Z_t \,, \quad \text{capital investment} \,, \\
\mathbb{G}_t &= G_t / Z_t \,, \quad \text{government spending} \,, \\
w_t &= W_t / (Z_t P_t) \,, \\
r_t^k &= R_t^k / P_t \,, \\
\lambda_t &= \Lambda_t Z_t \,, \\
\tilde{\Gamma}_t &\equiv \frac{\Gamma_t}{P_t Z_t} \,.
\end{aligned}$$

Definition 1 (Normalized equilibrium)**.** 19 endogenous variables $\{\lambda_t, g_{t+1}, i_t, \pi_t, \pi_t^w, X_{1t}^p, X_{2t}^p, X_{it}^w, X_{2t}^w, c_t, y_t, z_t, k_{t+1}^u, r_t^K, \mathbb{I}_t, q_t, u_t, k_t, w_t, L_t\}$, 3 shocks $\{\lambda_t^g, \epsilon_t^{mp}, \mu_t\}$ given the natural rate allocation variables $\{y_t^f, g_{t+1}^f\}$.

Consumption Euler equation

$$1 = \beta \mathbb{E}_t \left[\frac{\lambda_{t+1}}{\lambda_t (1 + g_{t+1})} \frac{1 + i_t}{\pi_{t+1}} \right] \,, \tag{15}$$

$$\lambda_t = \frac{1}{c_t - \frac{hc_{t-1}}{1+g_t}} - \frac{h\beta}{c_{t+1}(1+g_{t+1}) - hc_t}, \tag{16}$$

Price-setting

$$\frac{X^p_{1t}}{X^p_{2t}} = \left(\frac{1 - \theta_p \pi_t^{\frac{1}{\lambda_p}}}{1-\theta_p} \right)^{-\lambda_p}, \tag{17}$$

$$X^p_{1t} = \lambda_t y_t mc_t + \theta_p \beta \pi_{t+1}^{\frac{1+\lambda_p}{\lambda_p}} X^p_{1t+1}, \tag{18}$$

$$X^p_{2t} = \frac{1}{1+\lambda_p} \lambda_t y_t + \theta_p \beta \pi_{t+1}^{\frac{1}{\lambda_p}} X^p_{2t+1}, \tag{19}$$

Wage-setting

$$\frac{X^w_{1t}}{X^w_{2t}} = \left(\frac{1 - \theta_w (\pi^w_t)^{\frac{1}{\lambda_w}}}{1-\theta_w} \right)^{-\lambda_w + (1+\lambda_w)\nu}, \tag{20}$$

$$X^w_{1t} = \omega {L_t}^{1+\nu} + \theta_w \beta \left((\pi^w_{ss})^{1-\iota_w} (\pi_{W,t})^{\iota_w} \right)^{-\frac{(1+\lambda_{w,t+1})(1+\nu)}{\lambda_{w,t+1}}} \pi_{W,t+1}^{\frac{(1+\lambda_w)(1+\nu)}{\lambda_w}} X^w_{1t+1}, \tag{21}$$

$$X^w_{2t} = \frac{1+\tau^w}{1+\lambda_w} \lambda_t w_t L_t + \theta_w \beta \left((\pi^w_{ss})^{1-\iota_w} (\pi_{W,t})^{\iota_w} \right)^{\frac{-1}{\lambda_{w,t+1}}} \pi_{W,t+1}^{\frac{1}{\lambda_w}} X^w_{2t+1}, \tag{22}$$

$$\pi_{W,t} = \frac{w_t}{w_{t-1}} \pi_t (1+g_t), \tag{23}$$

Capital investment

$$k^u_{t+1} = \left[1 - S\left(\frac{\mathbb{I}_t}{\mathbb{I}_{t-1}} \frac{1+g_t}{1+g_{ss}} \right) \right] \mathbb{I}_t + (1-\delta_k) \frac{k^u_t}{1+g_t}, \tag{24}$$

$$q_t = \beta \mathbb{E}_t \left[\frac{\lambda_{t+1}}{\lambda_t (1+g_{t+1})} \left(r^K_{t+1} u_{t+1} - a(u_{t+1}) + q_{t+1}(1-\delta_k) \right) \right], \tag{25}$$

$$\begin{aligned} q_t \left[1 - S\left(\frac{\mathbb{I}_t}{\mathbb{I}_{t-1}} \frac{1+g_t}{1+g_{ss}} \right) - S'\left(\frac{\mathbb{I}_t}{\mathbb{I}_{t-1}} \frac{1+g_t}{1+g_{ss}} \right) \frac{\mathbb{I}_t}{\mathbb{I}_{t-1}} \frac{1+g_t}{1+g_{ss}} \right] \\ + \beta \frac{\lambda_{t+1}}{\lambda_t} q_{t+1} \frac{1+g_{t+1}}{1+g_{ss}} \left(\frac{\mathbb{I}_{t+1}}{\mathbb{I}_t} \right)^2 S'\left(\frac{\mathbb{I}_{t+1}}{\mathbb{I}_t} \frac{1+g_{t+1}}{1+g_{ss}} \right) = 1 \end{aligned} \tag{26}$$

Capital utilization rate

$$k_t = u_t \frac{k^u_t}{1+g_t}, \tag{27}$$

$$r^K_t = a'(u_t), \tag{28}$$

Production technologies

$$y_t = k_t^{\alpha} L_t^{1-\alpha}\,, \tag{29}$$

$$r_t^k = \alpha \frac{y_t}{k_t}\,, \tag{30}$$

$$w_t = (1-\alpha)\frac{y_t}{L_t}\,, \tag{31}$$

Endogenous growth

$$g_t = \log Z_t - \log Z_{t-1} = \mu_t + \eta \log(y_{t-1}/y_{t-1}^f)\,, \tag{32}$$

Government

$$\frac{1+i_t}{1+i_{ss}} = \left(\frac{1+i_{t-1}}{1+i_{ss}}\right)^{\rho_R} \left[\left(\frac{\pi_t}{\pi_{ss}}\right)^{\phi_\pi} \left(\frac{y_t}{y_t^f}\right)^{\phi_y}\right]^{1-\rho_R} \epsilon_t^{mp}\,, \tag{33}$$

Market clearing

$$y_t = c_t + \mathbb{I}_t + a(u_t)\frac{k_t^u}{1+g_t} + \left(1 - \frac{1}{\lambda_t^g}\right) y_t\,, \tag{34}$$

Shocks

$$\log \lambda_t^g = (1-\rho_g)\lambda^g + \rho_g \log \lambda_{t-1}^g + \epsilon_t^g; \quad \epsilon_t^g \sim N(0, \sigma_g)\,, \tag{35}$$

$$\epsilon_t^{mp} \sim N(0, \sigma_r)\,, \tag{36}$$

F.6. Steady state

$$1 = \beta \frac{1}{(1+g)} \frac{1+i}{\pi}\,,$$

$$\lambda = \frac{(1+g) - h\beta}{c(1+g-h)}\,,$$

$$mc = \frac{1}{1+\lambda_p}\,,$$

$$\omega L^{1+\nu} = \frac{1+\tau^w}{1+\lambda_w} \lambda w L\,,$$

$$\pi^w = \pi(1+g)\,,$$

$$q = 1\,,$$

$$u = 1\,,$$

$$(1 - \frac{1-\delta_k}{1+g})k^u = \mathbb{I}\,,$$

$$1 = \beta \left[\frac{1}{1+g} \left(r^K + (1-\delta_k) \right) \right] ,$$

$$k = \frac{k^u}{1+g} ,$$

$$r^K = a'(1) ,$$

$$y = k^\alpha L^{1-\alpha} ,$$

$$r^k = \alpha \frac{y}{k} ,$$

$$w = (1-\alpha) \frac{y}{L} ,$$

$$g = \mu ,$$

$$y = c + \mathbb{I} + \left(1 - \frac{1}{\lambda g} \right) y ,$$

F.7. Approximate equilibrium

We log-linearize the variables around the steady state as follows: for any variable x,

$$\hat{x}_t = \log \left(\frac{x_t}{x} \right) ,$$

where x is the steady state, except for the following variables

$$\hat{g}_{t+1} \equiv \log \left(\frac{1+g_{t+1}}{1+g} \right) ,$$

$$\hat{i}_t = \log \left(\frac{1+i_t}{1+i_{ss}} \right) ,$$

Definition 2 (Approximate equilibrium)**.** An approximate competitive equilibrium in this economy with endogenous growth is defined as a sequence of 16 endogenous variables $\{\hat{\lambda}_t, \hat{g}_{t+1}, \hat{i}_t, \hat{\pi}_t, \hat{c}_t, \hat{\pi}_t^w, \hat{mc}_t, \hat{y}_t, \hat{L}_t, \hat{w}_t, \hat{k}_{t+1}^u, \hat{\mathbb{I}}_t, \hat{q}_t, \hat{r}_t^K, \hat{k}_t, \hat{u}_t\}$, which satisfy the following equations, for a given sequence of one exogenous shock $\{\hat{\epsilon}_t^{mp}\}$ and given sequence of flexible price allocations.

Consumption Euler equation

$$(\mathbb{E}_t \hat{\lambda}_{t+1} - \hat{\lambda}_t - \hat{g}_{t+1}) + \hat{i}_t - \mathbb{E}_t \hat{\pi}_{t+1} = 0 , \tag{37}$$

$$\begin{aligned} \hat{\lambda}_t = & \frac{1+g}{1+g-h\beta} \left[\hat{g}_t - \left(\frac{1+g}{1+g-h} (\hat{c}_t + \hat{g}_t) - \frac{h}{(1+g)-h} \hat{c}_{t-1} \right) \right] \\ & + \mathbb{E}_t \frac{h\beta}{(1+g)-h\beta} \left[\left(\frac{1+g}{1+g-h} (\hat{c}_{t+1} + \hat{g}_{t+1}) - \frac{h}{(1+g)-h} \hat{c}_t \right) \right] \end{aligned} \tag{38}$$

Price-setting

$$\hat{\pi}_t = \frac{\beta}{1+\iota_p \beta} \mathbb{E}_t \hat{\pi}_{t+1} + \frac{\iota_p}{1+\iota_p \beta} \hat{\pi}_{t-1} + \kappa_p \hat{mc}_t , \tag{39}$$

where $\kappa_p = \frac{(1-\theta_p\beta)(1-\theta_p)}{\theta_p(1+\iota_p\beta)}$.

Wage-setting

$$\hat{\pi}_t^w = \frac{\beta}{1+\iota_w\beta}\mathbb{E}_t\hat{\pi}_{t+1}^w + \frac{\iota_w}{1+\iota_w\beta}\hat{\pi}_{t-1}^w + \kappa_w[-\hat{\lambda}_t + \nu\hat{L}_t - \hat{w}_t]\,, \tag{40}$$

where $\kappa_w \equiv \frac{(1-\theta_w)(1-\beta\theta_w)}{\theta_w(1+\iota_w\beta)(1+\nu\left(1+\frac{1}{\lambda_w}\right))} > 0$

$$\hat{\pi}_t^w = \hat{w}_t - \hat{w}_{t-1} + \hat{\pi}_t + \hat{g}_t\,, \tag{41}$$

Capital investment

$$\hat{k}_{t+1}^u = \frac{\mathbb{I}}{k^u}\hat{\mathbb{I}}_t + \frac{1-\delta_k}{1+g}\left[\hat{k}_t^u - \hat{g}_{t+1}\right]\,, \tag{42}$$

$$\hat{q}_t = \left[\mathbb{E}_t\hat{\lambda}_{t+1} - \hat{\lambda}_t - \hat{g}_{t+1}\right] + \frac{r^K}{r^K+(1-\delta_k)}\hat{r}_{t+1}^K + \frac{1-\delta_k}{r^K+(1-\delta_k)}\mathbb{E}_t\hat{q}_{t+1}\,, \tag{43}$$

$$\hat{q}_t - S''\left(\hat{\mathbb{I}}_t - \hat{\mathbb{I}}_{t-1} + \hat{g}_t\right) + \beta S''\left(\hat{\mathbb{I}}_{t+1} - \hat{\mathbb{I}}_t + \hat{g}_{t+1}\right) = 0\,, \tag{44}$$

where we parametrize S'' from Smets and Wouters (2007), or Justiniano, Primiceri, and Tambalotti (2013) to be between 3 and 5.

Capital utilization rate

$$\hat{k}_t = \hat{u}_t + \hat{k}_t^u - \hat{g}_t\,, \tag{45}$$

$$\hat{r}_t^K = \frac{a''(1)}{a'(1)}\hat{u}_t\,. \tag{46}$$

Production technologies

$$\hat{y}_t = \alpha\hat{k}_t + (1-\alpha)\hat{L}_t\,, \tag{47}$$

$$\hat{r}_t^k = \hat{y}_t - \hat{k}_t\,, \tag{48}$$

$$\hat{w}_t = \hat{y}_t - \hat{L}_t\,. \tag{49}$$

Endogenous growth

$$\hat{g}_t = \hat{\mu}_t + \eta(\hat{y}_{t-1} - \hat{y}_{t-1}^f)\,, \tag{50}$$

where η is the hysteresis elasticity.

Government

$$\hat{i}_t = \rho_R\hat{i}_{t-1} + (1-\rho_R)\left[\phi_\pi\hat{\pi} + \phi_y(\hat{y}_t - \hat{y}_t^f)\right] + \hat{\epsilon}_t^{mp}\,. \tag{51}$$

Market clearing

$$\frac{1}{\lambda^g}\hat{y}_t = \frac{c}{y}\hat{c}_t + \frac{\mathbb{I}}{y}\hat{\mathbb{I}}_t + \frac{a'(1)k}{y}\hat{u}_t\,. \tag{52}$$

Appendix references

Benigno, Gianluca, Pierpaolo Benigno, and Fabio Ghironi. 2007. Interest rate rules for fixed exchange rate regimes. *Journal of Economic Dynamics and Control* 31(7): 2196–2211.

Benigno, Pierpaolo. 2004. Optimal monetary policy in a currency area. *Journal of international economics* 63(2): 293–320.

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.

Burnside, Craig, and Martin Eichenbaum. 1996. Factor-Hoarding and the Propagation of Business-Cycle Shocks. *American Economic Review* 86(5): 1154–1174.

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.

Corsetti, Giancarlo, Keith Kuester, and Gernot J. Müller. 2013. Floats, pegs and the transmission of fiscal policy. *Central Banking, Analysis, and Economic Policies Book Series* 17: 235–281.

De Paoli, Bianca. 2009. Monetary policy and welfare in a small open economy. *Journal of international Economics* 77(1): 11–22.

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.

Farhi, Emmanuel, and Ivan Werning. 2012. Dealing with the trilemma: Optimal capital controls with fixed exchange rates. Working Paper 18199, National Bureau of Economic Research.

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

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

Gali, Jordi, and Tommaso Monacelli. 2005. Monetary policy and exchange rate volatility in a small open economy. *The Review of Economic Studies* 72(3): 707–734.

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.

Jordà, Òscar, Moritz Schularick, and Alan M. Taylor. 2019. The effects of quasi-random monetary experiments. *Journal of Monetary Economics.* Forthcoming.

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.

Nakamura, Emi, and Jón Steinsson. 2014. Fiscal stimulus in a monetary union: Evidence from US regions. *American Economic Review* 104(3): 753–92.

Romer, Christina D., and David H. Romer. 2004. A New Measure of Monetary Shocks: Derivation and Implications. *American Economic Review* 94(4): 1055–1084.

Romer, Paul M. 1990. Endogenous Technological Change. *Journal of Political Economy* 98(5): 71–102.

Smets, Frank, and Rafael Wouters. 2007. Shocks and frictions in US business cycles: A Bayesian DSGE approach. *American Economic Review* 97(3): 586–606.

Wieland, Johannes F., and Mu-Jeung Yang. 2016. Financial Dampening. NBER Working Papers 22141, National Bureau of Economic Research, Inc.

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