NBER WORKING PAPER SERIES

THE EFFECTS OF QUASI-RANDOM MONETARY EXPERIMENTS

Òscar Jordà
Moritz Schularick
Alan M. Taylor

Working Paper 23074
http://www.nber.org/papers/w23074

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
January 2017, Revised November 2017

Previously circulated as "Large and State-Dependent Effects of Quasi-Random Monetary Experiments." Comments and suggestions from James Cloyne, Julian di Giovanni, Gernot Müller, Ricardo Reis, and Jón Steinsson have helped improve the paper. We are grateful to Helen Irvin for outstanding research assistance. We thank James Cloyne and Patrick Hürtgen for sharing their data with us. Seminar and conference participants at the Federal Reserve Board of Governors, the Federal Reserve Bank of Chicago, the Federal Reserve Bank of San Francisco, the Federal Reserve Bank of Cleveland, the Fourth CEPR Economic History Symposium, De Nederlandsche Bank, the Bank of England, Ohio State University, and the NBER Monetary Economics program meeting provided useful feedback. All errors are our. Generous support from the Institute for New Economic Thinking, the Bundesministerium für Bildung und Forschung (BMBF), and the Volkswagen Foundation supported different parts of the data collection and analysis effort. We are grateful for their support. The views expressed in this paper are the sole responsibility of the authors and to not necessarily reflect the views of the Federal Reserve Bank of San Francisco, the Federal Reserve System, or the National Bureau of Economic Research.

At least one co-author has disclosed a financial relationship of potential relevance for this research. Further information is available online at http://www.nber.org/papers/w23074.ack

The effects of quasi-random monetary experiments
Òscar Jordà, Moritz Schularick, and Alan M. Taylor
NBER Working Paper No. 23074
January 2017, Revised November 2017
JEL No. E01,E30,E32,E44,E47,E51,F33,F42,F44

ABSTRACT

The trilemma of international finance entails that fluctuations in interest rates—for countries with fixed exchange rates that allow unfettered cross-border capital flows—are mostly due to international arbitrage. Consequently, we can locate a valid source of exogenous variation to identify monetary policy effects with instrumental variable methods. Paired with conventional instruments based on central bank staff forecasts, and using historical data since 1870, we estimate local average treatment effects (LATE) of monetary policy interventions for different subpopulations. Using a novel control function approach we determine the robustness of our findings to possible spillovers via alternative trade-based channels. Our results reveal and rectify attenuation bias in previous estimates, are consistent with theory, and provide a good approximation to the ATE. The effects that we report are quantitatively important and state-dependent.

Òscar Jordà
Economic Research, MS 1130
Federal Reserve Bank of San Francisco
101 Market St.
San Francisco, CA 94105
and University of California, Davis
oscar.jorda@sf.frb.org

Moritz Schularick
University of Bonn
53113 Bonn, Germany
moritz.schularick@uni-bonn.de

Alan M. Taylor
Department of Economics and
Graduate School of Management
University of California
One Shields Ave
Davis, CA 95616-8578
and CEPR
and also NBER
amtaylor@ucdavis.edu

1. Introduction

Few questions in economics have been studied and debated more intensely than the effects of monetary policy on output and inflation. In most of our theories money is taken to be neutral in the long-run, but the short-run effects can be large and non-neutral, depending on various market frictions. If this was ever a consensus view, the debate has now re-opened. The aftermath of the global financial crisis has seen central banks struggle to hit their targets, raising the question as to whether monetary policy has any short-run potency at all. At the same time, a recent but growing literature on labor market hysteresis, argues that monetary policy can have very persistent effects (see, e.g., Galí 2016), challenging standard long-run neutrality assumptions. As a result, determining whether monetary policy is a useful cyclical stabilizer is still an empirical matter. However, empirical measures of the effect of interest rates on macroeconomic outcomes are fraught. Macroeconomic aggregates and interest rates are jointly determined since monetary policy reflects the central bank's policy choices given the economic outlook. In the parlance of the policy-evaluation literature (see, e.g., Rubin, 2005), any measure of the treatment effect of policy is contaminated by confounders simultaneously correlated with the treatment assignment mechanism and the outcome.

Over time, several best-practice methods have been devised, but almost all are exclusively based on the Post-WW2 U.S. experience. Identification primarily rests on the assumption that any variation in interest rates that cannot be explained by observable information must reflect random variation in the policy assignment. First, the traditional approach to identification started by using a rich set of controls directly in the regression—see, for example, Christiano, Eichenbaum, and Evans (1999) and the long tradition of structural VARs cited therein and since. Identification, further, relies on additional exclusion (short-run, long-run, sign) restrictions whose validity is often untestable.

Second, an influential "narrative" identification approach was proposed, where textual analysis of central bank minutes could yield information on policy shifts that were plausibly exogenous with respect to contemporaneous economic conditions—see Romer and Romer (1989), and the subsequent literature. Identification still depends on appropriate control—what may surprise the reader, may not have surprised the narrator.

Aware of these issues, a third approach emerged. By conditioning on central bank forecasts available at the time of the policy decision, researchers devised another measure of plausibly random surprises to policy—see Romer and Romer (2004), and the subsequent literature such as Cloyne and Hürtgen (2016) or Coibion et al. (2017). Again, measurement of the policy surprise hinges on the assumption that the policymaker does not rely on additional variables for which the staff may not have prepared a forecast.

Lastly, researchers have drawn on the capacity of financial markets to aggregate vast amounts of information used in "high frequency" identification. Shifts in funds rate futures (or other asset prices) in tight windows around policy announcements have been used, among others, by Kuttner (2001); Faust, Wright, and Swanson (2004); Gertler and Karadi (2015); and Nakamura and Steinsson (forthcoming). While advantageous from the point of view of control, what surprised markets may not have surprised policymakers. As debate still rages, each method confronts its critics, and a recent survey stresses the limits to our knowledge (Ramey 2016).

Therefore, this paper proposes a new and different approach based on a natural experiment. How a country manages its exchange rate and how freely capital flows across its borders has direct implications for its domestic interest rates, and hence for monetary policy. This is especially true for safe assets with low liquidity and risk premiums, such as government securities and interbank credit. Consider such assets which, though denominated in different currencies, are otherwise perfect substitutes. Assume their currencies are credibly pegged for the duration of the investment and that investors can freely transfer funds. Then, without exchange rate risk, arbitrage would equalize the rates of return across all markets. Economic forces thus limit a country's policy choices with respect to the triad of capital mobility, exchange rates, and interest rates. The *trilemma* faced by policymakers is that they can have control over two out of the three policies, but not all three simultaneously. Articulating how and when the trilemma functions as a source of natural experiments in domestic monetary policy is one of the contributions of this paper.

The general empirical validity of the trilemma, in recent times and in distant historical epochs, has been recognized for more than a decade: exogenous base country interest rate movements spill into local interest rates for open pegs[1] (Obstfeld, Shambaugh, and Taylor 2004, 2005; Shambaugh 2004). Some important corollaries directly follow for empirical macroeconomics. A key contribution by di Giovanni, McCrary, and von Wachter (2009) exploits the resulting identified local monetary policy shocks to estimate impulse response functions for other macroeconomic outcome variables of interest using standard VAR methods with instrumental variables.[2]

In this paper we take these ideas further. Variation in base rates can be used as a natural experiment, but only by appropriately sorting the different channels of transmission. We do this by extending the lessons from the policy evaluation literature and identification with instrumental variables (IV) to a time series setting using local projections (Jordà 2005). The framework we develop is quite general. For example, we entertain variation in the *dose given treatment* (interest rates can be moved up or down, sometimes by bigger amounts than at other times) and we allow for non-monotonic and state-dependent treatment effects (increasing rates by twice the usual amount may not be twice as effective). There are numerous extensions that we explore in follow up work. Here we focus on illustrating the core features of the approach.

Instrumental variable applications using local projections (LP-IV) have recently appeared in a variety of settings (see, e.g. Jordà, Schularick, and Taylor 2015, Ramey and Zubairy 2017; and Stock and Watson 2017). Here we develop these econometric methods further to link impulse responses with the notion of the local average treatment effect (LATE). Finally, we expand the sample from the commonly studied period of post-WW2 in the U.S. to all of advanced economy macroeconomic

[1] Open pegs are countries that fix their exchange rate but allow relatively free movement of capital.

[2] In related work, di Giovanni and Shambaugh (2008) used the trilemma to investigate post-WW2 output volatility in fixed and floating regimes. Ilzetzki, Mendoza, and Végh (2013) partition countries by exchange rate regime to study the impact of a fiscal policy shock. In previous work (Jordà, Schularick, and Taylor 2015), we studied the link between financial conditions, mortgage credit, and house prices.

history since 1870. Our results are thus based on a large cross-sectional sample spanning over a century and are therefore important in that they generalize extant findings based on U.S. post-WW2 data alone.

Our main findings are easily summarized. First, using the subpopulation of open pegs, we find evidence of considerable attenuation bias in policy responses when we estimate the responses to monetary policy using traditional OLS selection-on-observables identification versus identification with instrumental variables. Second, we investigate the robustness of our new IV estimates of LATE for open pegs as compared to the LATE found by investigating the combined behavior of post-WW2 data from the U.S. and U.K. using the established instrumental variables approaches of Romer and Romer (2004) and Cloyne and Hürtgen (2016). In the latter, instruments are based on staff forecasts available at the time policy was set. Third, we assess the robustness of our findings to potential spillover confounding using control functions (see Wooldridge 2010; and Conley, Hansen and Rossi 2012) by taking advantage of the heterogeneity across the subpopulations in our data. Although it might appear obvious which way this bias goes, careful calculation shows that its direction is ambiguous. Guided by economic reasoning, we place plausible bounds on this bias and show that this source of confounding, if anything, tends to reinforce our main findings.

To sum up, we find that our LATE estimates based on different instruments and samples nevertheless paint a similar picture for what the ATE might be. Relative to selection-on-observables identification, the attenuation bias found across subpopulations is remarkably similar, and it is large. Moreover, LATE estimates turn out to be very similar to each other across a variety of subsamples examined. Altogether the evidence strongly suggests that our new IV estimates of LATE are a very good and reliable approximation to the unknowable ATE.

Moreover, although the policy evaluation literature typically focuses on binary treatments, our analysis can easily accommodate varying treatment levels or doses. Central bank interventions consist of sometimes raising interest rates, sometimes lowering them, and often by different amounts. Dose effectiveness can also be state-dependent. Recent studies on the fiscal multiplier (e.g., Auerbach and Gorodnichenko 2013ab; Jordà and Taylor 2016; and Ramey and Zubairy 2017) have found fiscal policy to have different effects depending on the state of the economy. Framed against a similar desire to understand monetary stabilization, our analysis sheds light on the recent results on monetary policy asymmetry reported in Angrist, Jordà, and Kuersteiner (2016) and Tenreyro and Thwaites (2016).

Taking a step further, our empirical investigations also consider recent theoretical developments that examine the efficacy of monetary policy as a function of (household and/or firm) leverage. For example, Auclert (2017) shows that the effectiveness of monetary policy depends on household balance sheet exposure through the redistributive channels that interest rates can have. In Kaplan, Moll, and Violante (2017), the mechanism operates via the heterogeneity in marginal propensities to consume of households facing uninsurable income shocks in incomplete markets. Narrowing the focus, Iacioviello (2005) and Cloyne, Ferreira, Surico (2015) argue that households' reactions to monetary policy shocks varies depending on variation in levels of mortgage indebtedness. On the firm side, a venerable literature (Bernanke and Gertler, 1995; Kashyap, Lamont, and Stein 1994; and Kashyap and Stein 1995) argued that credit constrained firms are more responsive to monetary policy. More recently, Ottonello and Winberry (2017) focus on firm (credit) heterogeneity to investigate asymmetry in the investment channel of monetary policy. This is just an example of other extensions permitted in our framework, some of which we examine in more detail.

In particular, we find—like others before us—that one source of state dependence comes from how the economy responds to monetary policy in the boom versus the slump. Specifically, we find that stimulating a weak economy is much harder than reining in a strong one. Another source of state dependence comes from the level of inflation, however. Advanced economies have recently struggled with a low-growth, low-inflation environment, referred to by commentators as "lowflation." The historical data show that policy turns out to be rather ineffective in lowflation environments, thus revealing another hitherto unexamined dimension in which monetary policy is asymmetric.

Turning to credit and using data that has only recently become available (Jordà, Schularick, and Taylor 2017), we find that relatively more mortgage borrowing translates to interest rates restraining economic activity by more and prices by less than with relatively less mortgage borrowing. Thus, the evidence suggests that, for example, in a mortgage boom where house prices are appreciating rapidly and the economy is growing above potential, higher interest rates can slow down economic activity quite rapidly. However, output losses would pile up quickly if the objective of the central bank were to bring down inflation in such an environment as the inflation-output tradeoff becomes particularly costly. Interestingly, we find that non-mortgage (unsecured) lending does not generate such an asymmetry, once more casting some doubt on channels operating through firm leverage.

This paper makes a number of contributions to monetary economics and to empirical macroeconomics broadly speaking. On average, monetary policy has stronger and longer-lasting effects than previously understood, but consistent with results from a recent literature that uses external instruments for identification. However, monetary policy effects can vary greatly with the state of the economy, and that state may depend on a rich set of characteristics, including output, inflation, and leverage. Both of these findings matter when drafting theoretical models of monetary economies. In terms of empirical methods, we are careful to spell out, in a dynamic local projection framework, the crucial instrumental variable identification assumptions and how they might be valid for some subpopulations but not for others. Moreover, we introduce methods to provide well-reasoned bounds on potential biases coming from untestable failures of the instrumental variables exclusion restriction. All of these new developments should provide for a tighter link with long-standing policy evaluation methods in applied microeconomics and bring empirical macroeconomics closer to a unified protocol for data analysis.

2. The trilemma of international economics: a quasi-natural experiment

Throughout the history of modern finance, countries have managed international goods and capital flows using a variety of policies. The interplay of these policies had, at times, important consequences for domestic monetary conditions. We exploit situations when external conditions bleed into domestic policy as a way to identify exogenous movements in monetary conditions. Given the important role that this mechanism plays, we begin by presenting the main statistical properties of our instrumental variable before discussing our overall empirical approach. The specific construction of this instrumental variable is one of the important contributions of our paper.

We investigate economies broken down into the three subpopulations naturally defined by the trilemma. We set aside *base* countries from all others since their currency serves as the focal value for pegging economies or *pegs*. We call the remaining economies, which allow their currency to be determined freely in the market, *floats*. Hence, we define a binary regime indicator $q_{i,t} \in \{0,1\}$ where if a country is currently in a peg and it was in a peg the previous year then $q_{i,t} = 1$, otherwise $q_{i,t} = 0$. This definition is conservative and ensures the peg is well-established before including a particular country-time pair $\{i,t\}$ into the subpopulation of pegs. Below we show that such a conservative definition imposes a higher burden of proof for our analysis thus making our results more credible.

Let $\Delta r_{i,t}$ denote the change in short-term nominal interest rates in the home country i at time t, and let $\Delta r_{b(i,t),t}$ denote the short-term nominal interest in country i's base country at time t, which can differ across i and over time, hence the notation $b(i,t)$. In our international historical dataset (discussed in Section 3 in detail), these interest rates are taken from the short end of the sovereign yield curve, specifically three-month government bond rates, the closest measure to a policy rate that we were able to obtain consistently for our long and wide panel of historical data since 1870.[3]

Denote with $\Delta\widehat{r}_{b(i,t),t}$ movements in base country $b(i,t)$ rates explained by observable controls for that base country, denoted $\boldsymbol{x}_{b(i,t),t}$. Similarly, denote with $\boldsymbol{x}_{i,t}$ a broad set of domestic macroeconomic controls in country i at time t. Such controls include current and lagged values of macroeconomic aggregates, and lagged values of the policy variable. Thus, define with $z_{i,t} \equiv \Delta r_{b(i,t),t} - \Delta\widehat{r}_{b(i,t),t}$ unpredictable movements in base country interest rates. The idea is to narrow the focus to movements in base country rates that would not have been predicted using observable information by country i.[4]

Next, define the variable $k_{i,t} \in [0,1]$ which indicates whether country i is open to international markets or not. We base this capital mobility indicator on the index (from 0 to 100) in Quinn, Schindler, and Toyoda (2011). We use a continuous version of their index rescaled to the unit

[3]Swanson and Williams (2014) and Gertler and Karadi (2015) are two recent examples of papers that use medium-term government rates to measure monetary policy effects as well.

[4]A lower standard of proof would be to use $\Delta r_{b(i,t),t}$ directly. Experiments with such an instrument produced similar results to those reported below and are available upon request.

interval, with 0 meaning fully closed and 1 fully open. International full capital mobility has been mainly interrupted by the two world wars. Resumption of mobility was nearly immediate after WW1. Not so after WW2, in large part due to the tight constraints of the Bretton Woods regime. Nowadays, capital mobility is commonplace. It is fair to say, therefore, that restrictions on capital mobility have not been used as a high frequency policy tool by pegging economies.

Economies generally allow capital to flow relatively freely. The average value of k for the pegs in the full sample is 0.87 (with a standard deviation of 0.21) versus 0.70 (0.31) for floats. In the post-WW2 era, these averages are virtually indistinguishable from one another, with values of 0.76 (0.24) for pegs and 0.74 (0.30) for floats. It cannot be said that pegs used more restrictions on capital to regain control over monetary policy after pegging.

Using these definitions we can now construct the function $\lambda_{i,t} = \lambda(z_{i,t}, k_{i,t})$. This function spells out, in general terms, how fluctuations in $z_{i,t}$ affect domestic rates. In principle the association between home and base rates may be nonlinear. Notice that $\lambda_{i,t}$ also depends on $k_{i,t}$, which modulates the transmission of fluctuations in $z_{i,t}$ to $\Delta r_{i,t}$ depending on how open the economy is to capital flows. This is a direct consequence of the trilemma. Finally, we will be more detailed on the role that the function $\lambda_{i,t}$ plays in the context of investigating local average treatment effects later on. Thus, while it is unusual to specify an instrumental variable using an unknown function of random variables, it allows us to be more accurate in stating the main assumptions.

Similarly, define the function $\delta_{i,t} = \delta(\Delta r_{i,t})$. The purpose of this function is to allow for potentially nonlinear effects of interest rates on macroeconomic aggregates in the potential outcomes framework, as we discuss momentarily. In particular, when treatment is binary, there are only two possible states, units that receive the treatment versus units that do not. Instead the function $\delta_{i,t}$ allows us to consider the effect of the *dose*, given treatment. Another advantage of our more general setting is to allow for state-dependence, as we show later in the paper.

For example, to gain intuition, consider a simple special case with $\lambda_{i,t} = z_{i,t}$, with $k_{i,t} \equiv 1$, and $\delta_{i,t} = \Delta r_{i,t}$. This is the familiar textbook case of pure uncovered interest parity under an frictionless hard peg regime (i.e., with no band), where the pass through from exogenous base country rates to home rates is one-for-one and perfectly linear. As we see in a moment, reality is not quite so simple.

Given these definitions, we further assume that:

Assumption 1. Monotonicity
Let δ and λ denote the random variables defined earlier, where the subindices are omitted for simplicity. Then we assume that in population:

$$\frac{\partial E(\delta|\lambda)}{\partial \lambda} \geq 0.$$

That is, according to this assumption, if the base country raises its interest rate, on average the home country will too—as the trilemma suggests it should be. This assumption is similar to the assumption of monotonicity in, for example, Angrist and Imbens (1994) although in their application both δ and λ are binary variables.

In addition, and again based on the trilemma, we will also assume that:

$$\frac{\partial \lambda(z,k)}{\partial k} \geq 0,$$

where again we omit subindices. In fact, we will assume that $\lambda \to 0$ as $k \to 0$. Here, we allow for k to modulate λ, so that as capital moves more (less) freely, the effect of base rates on home rates rises (vanishes).

Sometimes exchange rates are managed over a small band around the peg. This was the case for several European economies in the lead up to the euro. This poses no difficulty for our instrument construction, however. We do not require fluctuations in home country rates to be perfectly explained by base country rates. All that is needed is for these two rates to be associated. Fluctuations inside a corridor limit exchange rate variation such that wide interest rate differentials cannot persist in practice.

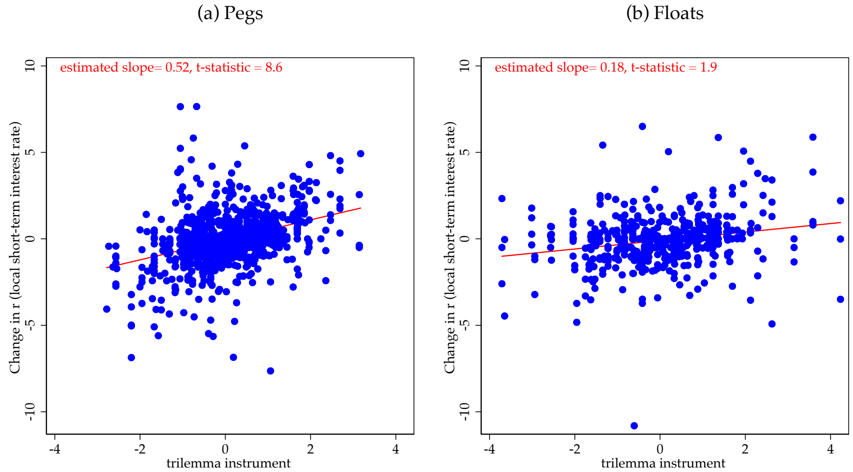
Over the span of nearly 150 years, countries have come in and out of fixed exchange regimes. However, once in such a regime, countries tend to stick to the arrangement. In contrast to Obstfeld and Rogoff (1995), who found the average duration of fixed exchange rates since 1973 to be about 5 years, we find in our longer-run sample that average to be about 21 years. Part of the reason is that our sample includes advanced economies only, whereas theirs includes emerging market economies as well. Another is that our sample includes longer-lived peg episodes in the gold standard and Bretton Woods eras.

To get a sense of the relationship between $\delta_{i,t}$ and $\lambda_{i,t}$, Figure 1 shows scatterplots and fitted values from a regression of the change in the policy function, defined for now as $\delta_{i,t} = \Delta r_{i,t}$, on the trilemma instrument functional, defined for now as $\lambda_{i,t} = z_{i,t}\, k_{i,t}$ for simplicity (including controls and fixed effects). We compare the subpopulation of pegs ($q_{i,t} = 1$) with the subpopulation of floats ($q_{i,t} = 0$). These two regressions correspond to the results reported in column 4 of Table 1. To better evaluate the strength of the instrument, Table 1 reports first-stage regression results of the endogenous variable, $\delta_{i,t} = \Delta r_{i,t}$ on the instrument $\lambda_{i,t}$, without controls in columns 1–3, and then more formally with controls, in columns 4–6. We do this for the subpopulation of pegs in the first block of rows, and then for the subpopulations of floats is the second block of rows.

Table 1 reaffirms the evidence displayed in Figure 1 formally, and clearly shows that $\lambda_{i,t}$ is not a weak instrument. Columns 4–6 refer to the formal first-stage regression with controls, country fixed effects and robust (clustered) standard errors (the regression also allows the coefficients of the controls to differ for the period 1973 to 1980 to account for the two oil crises as we also discuss below). The t-statistic on $\lambda_{i,t}$ is well above 3 for the full and post-WW2 samples. Moreover, notice that the slope estimates in columns 4–6 are similar to those in columns 1–3, which suggests that the instrument contains information that is quite different from that contained in the controls.[5]

[5]The control list will be discussed in more detail below but it basically includes 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.

Figure 1: *Pegs vs. floats: changes in home short-rates in the home country vs. the trilemma instrument*



Notes: Full sample: 1870–2013 excluding 1914–1919 and 1939–1947. Country fixed effects included. These regressions also include 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. Panel (a): all peg economies with $q_{i,t} = 1$. Panel (b): all floating economies with $q_{i,t} = 0$. See text.

Table 1: *Relationship between change in short-rates for pegs and the trilemma instrument*

	No controls			With controls		
	(1) All years	(2) Pre-WW2	(3) Post-WW2	(4) All years	(5) Pre-WW2	(6) Post-WW2
Pegs: $q = 1$						
instrument	0.58***	0.40***	0.67***	0.52***	0.35*	0.56***
	(0.09)	(0.09)	(0.10)	(0.06)	(0.17)	(0.06)
t-statistic	[6.58]	[4.32]	[6.46]	[8.62]	[2.05]	[8.97]
Observations	1059	438	621	672	148	524
Floats: $q = 0$						
instrument	0.25**	0.14	0.28*	0.16*	-0.09	0.19**
	(0.11)	(0.10)	(0.14)	(0.08)	(0.06)	(0.08)
t-statistic	[2.24]	[1.46]	[1.99]	[1.92]	[-1.66]	[2.34]
Observations	530	233	297	316	57	259

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors in parentheses. Full sample: 1870–2013 excluding 1914–1919 and 1939–1947. Pre-WW2 sample: 1870–1938 (excluding 1914–1919). Post-WW2 sample: 1948–2013. Country fixed effects included in the regressions for columns 4–6. These regressions also include 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.

Next, consider the subpopulation of floats for which $q_{i,t} = 0$. According to the trilemma, the relationship between base and home country interest rates should rather reflect uncovered interest rate parity mechanisms through expected appreciation/depreciation of the exchange rate and therefore be considerably attenuated. This is indeed the case as the second block of rows in Table 1 shows. Coefficient estimates for this subpopulation are less than half the coefficients estimated for the peg subpopulation, with much lower t-statistics (in fact below 2 for the full and pre-WW2 samples). These estimates are thus consistent with a large and growing literature on the trilemma. However, as we shall see these estimates are also helpful to provide bounds for potential spillovers (or violations of the exclusion restriction that may have gone unsuspected in typical settings).

This discussion serves to establish our proposed λ as a potentially strong instrument. However, before discussing the empirical methods further, we briefly discuss another of the novelties in this paper, the data.

3. The data

This paper relies on a large macro-historical dataset that we have assembled with the help of a small army of research assistants and generous support from many colleagues, and kept up to date over the past years. The database covers a broad range of real and financial variables for 17 countries at annual frequency for the modern era from 1870 to 2013. In Jordà, Schularick, and Taylor (2017), we describe the underlying data and its sources in great detail. Virtually all the data used in the paper are publicly available at http://www.macrohistory.net/data/. This website includes a long and very detailed appendix describing all the data sources and relevant transformations.

The data come from a broad range of sources, including series constructed by economic historians and statistical offices, economic and financial history papers, yearbooks of central banks, as well as sources from central and private bank archives. The database contains the near-universe of available modern macroeconomic data for the following advanced economies: Australia, Belgium, Canada, Denmark, Finland, France, Germany, Italy, Japan, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, U.K., and U.S. Unlike the mass of empirical papers investigating the effects of monetary policy, our analysis rests both on a long time dimension and a large group of countries.

On the national accounts side, we have near complete coverage of real GDP data, interrupted only by a few war-time gaps. However, as we generally exclude war years from the empirical analysis this does not impact the results. Other than GDP, we also study the effects of changes in monetary conditions on consumption and investment decisions. For consumption and investment, we typically rely on country-level historical national account reconstructions as well as the data assembled by Barro and Ursúa (2010).

On the financial side, we can work with a wide range of both quantity and price data. The credit data cover total bank lending to the non-financial private sector. We also have broad coverage of asset prices, again relying on numerous sources. The dataset comprises stock price indices as well as house price data from Knoll, Schularick, and Steger (2017), the first major attempt to

Table 2: *Selection of base country short-term interest rate for pegged exchange rates by era*

	Home country interest rates associated with the base country			
Base country interest rate	Pre-WW1	Interwar	Bretton Woods	Post-BW
UK (Gold standard/BW base)	All countries		Sterling bloc: AUS*	
UK/USA/France composite (Gold standard base)		All countries		
USA (BW/Post-BW base)			All other countries	Dollar bloc: AUS, CAN, CHE, JPN, NOR
Germany (EMS/ERM/Eurozone base)				All other countries

*We treat Australia as moving to a U.S. dollar peg in 1967.
Notes: See text, Jordà, Schularick, and Taylor (2015), and Obstfeld, Shambaugh, and Taylor (2004, 2005). Pre-WW1: 1870–1914; Interwar: 1920–1938; Bretton Woods: 1948–1971; Post-BW: 1972–2013.

construct long-run house price indices for the advanced economies. A central variable for our analysis is the short-term interest rate. We typically use a risk-free short-term treasury bill rate with a maturity between 1 and 3 months. In some cases, we relied on central bank discount rates or short-term deposit rates of large financial institutions. For long-term interest rates we have yields on government bonds with a maturity between 5 and 10 years.

3.1. Defining the base country in fixed exchange rate regimes

Part of the data construction effort consists of defining the base countries to which home countries pegged their exchange rates across different eras. This is described in Table 2. The possible base country interest rates used at different times in the history of exchange rate regimes correspond to the four rows in the table. The four major eras correspond to the four columns in the table. The table cells indicate which home countries correspond to each base in each era.

Prior to WW2, peg codings are taken from Obstfeld, Shambaugh, and Taylor (2004, 2005). After WW2 they are gleaned from Ilzetzki, Reinhart, and Rogoff (2008) and updates thereto. One exception with respect to this literature is that we do not code Germany as being pegged from 1999 onwards to emphasize the outsize role that Germany plays within the euro zone as a continuation of its central position with the EMS/ERM fixed exchange rate system. Prior to 1914 we treat the U.K. as the base for everyone, and after 1945 we treat the U.S. as the base for everyone, with the exception of EMS/ERM/Eurozone countries for which Germany is the base after 1973. In the interwar period, the choice of a suitable base country is more challenging and subjective given the instability of the interwar gold standard period; we follow Obstfeld, Shambaugh, and Taylor (2004) in using a hybrid "gold center" short-term interest rate, which is an average of U.S., U.K., and French short term rates depending on which of the three countries was pegged to gold in a particular year; our results are not sensitive to this choice and we replicate our findings using any one of these three countries as the sole interwar base as in Obstfeld, Shambaugh, and Taylor (2004).

Base countries, like the U.S. in the Bretton Woods era, are conventionally understood to pay no attention whatsoever to economic conditions in partner countries when making policy choices. Such behavior finds ample support in the historical record, as discussed in Jordà, Schularick, and Taylor (2015). Thus, to peg is to sacrifice monetary policy autonomy, at least to some degree. It is therefore natural for us to treat the trilemma instrument as an exogenous shifter of local monetary conditions in the home economy. The next section elaborates on the statistical design based on this idea.

4. Methods

The statistical analysis discussed in this section is tailored to the particular features of our data and the economic questions we investigate. Specifically, we want to evaluate the effects of a monetary policy intervention via interest rates on important macroeconomic outcomes. In order to evaluate the effects of this intervention, we want to compare the average observed outcome against the average counterfactual absent the intervention. The notation that we use refers to random variables for which a sample of (panel) data is available. We will omit country and time subindices initially to streamline the presentation of the main derivations. We will bring back the more standard notation when presenting the specific models to be estimated.

Define the random variable $y(h)$ in reference to a macroeconomic outcome of interest observed h periods from today for $h = 0, 1, \ldots, H-1$. For example, we may be interested in the growth rate approximation given by 100 times the log difference between real GDP in $t+h$ relative to t when computing a cumulative impulse response resulting from a monetary policy intervention. We can collect all such random variables into an $H \times 1$ vector $\boldsymbol{y} = (y(0), y(1), \ldots, y(h), \ldots, y(H-1))$.

We can think of the observed random variable $\boldsymbol{y}$ as determined by a mixture of two latent random variables, $\boldsymbol{y}_1$ and $\boldsymbol{y}_0$. Using the potential outcomes notation (see, e.g., Rubin 1974) $\boldsymbol{y}_1$ refers to the potential outcome when there is an intervention and $\boldsymbol{y}_0$ is the potential outcome absent the intervention. This notation usually refers to binary interventions, say $\delta = 0, 1$ using the policy function notation introduced earlier. Later we discuss how to deal with a continuous policy variable—think of it as treatment with a variable dose.

In reporting the average effects of a policy intervention over time, we will adopt the convention of normalizing the intervention to a 1 percentage point (100 bps) increase in interest rates compared to a counterfactual of leaving interest rates unchanged. We note that a quantity such as $E(\boldsymbol{y}_1|\delta = 0)$ is a well defined object. It refers to the counterfactual expected outcome from the treated subpopulation had intervention been withheld.

Each potential outcome is characterized by its own unconditional mean, $E(\boldsymbol{y}_j) = \boldsymbol{\mu}_j$ for $j = 0, 1$. Without loss of generality, we can write $\boldsymbol{y}_j = \boldsymbol{\mu}_j + \boldsymbol{v}_j$, where, by construction, $E(\boldsymbol{v}_j) = \mathbf{0}$. In applied microeconomics, it is often of interest to report the *average treatment effect* or ATE given by $\mathcal{R}_{ATE} \equiv E(\boldsymbol{y}_1 - \boldsymbol{y}_0) = \boldsymbol{\mu}_1 - \boldsymbol{\mu}_0$. In our context, given the dynamic aspects of our data, this ATE has a natural mapping to the usual definition of an impulse response function, as we explain below.

We assume that any heterogeneity gets collected in the term $\boldsymbol{v}_j$ and in particular, any influence on the outcomes coming from controls $\boldsymbol{x}$. The vector of controls will typically include lags of the outcome, lags of the intervention variable, and any other exogenous or predetermined variables. Moreover, we assume that the controls affect the outcome linearly so that $\boldsymbol{v}_j = (\boldsymbol{x} - E(\boldsymbol{x}))\boldsymbol{\gamma} + \boldsymbol{\nu}_j$ for $j = 0, 1$. The controls enter in deviations from population means to ensure that $E(\boldsymbol{v}_j) = \boldsymbol{0}$. As a further simplification, we assume that the controls affect the outcomes in the same manner for the treated and control subpopulations so that $\gamma_0 = \gamma_1 = \gamma$. This assumption is easily relaxed in our setup, but it is a natural starting point and sets our analysis on the same footing as a standard vector autoregression or VAR.

Under our temporary maintained assumption that $\delta = 0, 1$ the observed data are generated by $\boldsymbol{y} = (1 - \delta)\boldsymbol{y}_0 + \delta\boldsymbol{y}_1 = \boldsymbol{y}_0 + \delta(\boldsymbol{y}_1 - \boldsymbol{y}_0)$. In general, interventions are such that $-\infty < \delta < \infty$, however. Think of δ as reflecting the administered dose. Such an extension does not materially change how $\boldsymbol{y}$ is generated if we assume that treatment doses have a linear effect on outcomes. In general this assumption may hold for doses that take values over a small range. However, it is unlikely to hold more broadly. A drug that is beneficial in small doses can kill at higher doses. For this reason, we investigate making the dose δ more flexible later in the analysis. For now, this assumption is also no different than that in a typical VAR. Using our assumptions on $\boldsymbol{y}_j$, we can therefore write:

$$\boldsymbol{y} = \boldsymbol{\mu}_0 + (\boldsymbol{x} - E(\boldsymbol{x}))\boldsymbol{\gamma} + \delta\boldsymbol{\beta} + \boldsymbol{\nu}\,, \tag{1}$$

where $\boldsymbol{\nu} = \boldsymbol{\nu}_0 + \delta(\boldsymbol{\nu}_1 - \boldsymbol{\nu}_0)$.

In practice, we would estimate this equation as the following local projection:

$$y_{i,t+h} = \alpha_{i,h} + \boldsymbol{x}_{i,t}\gamma_h + \delta_{i,t}\beta_h + \boldsymbol{\nu}_{i,t+h}\,; \qquad \text{for } h = 0, \ldots, H-1\,, \tag{2}$$

and for a longitudinal sample given by $i = 1, \ldots, N$; and $t = 1, \ldots, T$. These local projections could be naïvely estimated by standard OLS panel-data methods, and we will refer to that as LP-OLS. Notice that $\alpha_{i,h}$ is a fixed effect. From this regression it is easy to see that $\mathcal{R}_{ATE} = \boldsymbol{\beta} = (\beta_0, \ldots, \beta_{H-1})'$, that is, the impulse response calculated from local projections using expression (2) for $h = 0, \ldots, H-1$. Notice that in a linear model the constant term α_h will capture the means of the controls $E(\boldsymbol{x})$ along with $\boldsymbol{\mu}_0$; and $\delta_{i,t}$ will attain values that depend on whatever policy intervention variable one is interested in considering. Later, it will be the case here that $\delta_{i,t} = \Delta r_{i,t}$, that is, the change in the short-term rate on government securities that we will describe in more detail in the data section.

What about causality? Under OLS, the identification of causal effects for the coefficient β_h in expression (2) relies, roughly speaking, on δ being randomly assigned given the controls. In a linear setup such as ours, it would suffice to assume the following:

Assumption 2. Conditional mean independence
We assume that:

$$E(\boldsymbol{y}_1|\delta, \boldsymbol{x}) = E(\boldsymbol{y}_1|\boldsymbol{x}) \quad \textit{and} \quad E(\boldsymbol{y}_0|\delta, \boldsymbol{x}) = E(\boldsymbol{y}_0|\boldsymbol{x})\,, \tag{3}$$

that is, conditional on the controls, $\boldsymbol{x}$, there is no selection mechanism that explains differences in the conditional means of the potential outcomes.

Implicitly, this is the type of assumption that permeates most of the VAR literature. For example, identification based on zero short-run restrictions using a Cholesky ordering is equivalent to including appropriate contemporaneous values in the control vector $\boldsymbol{x}$ for each outcome $\boldsymbol{y}$ considered. However, as we show below, the conditional mean assumption is not supported in our application and for this reason we have to explore other methods to achieve identification that we now discuss.

4.1. Identification with instrumental variables

Consider the potential outcomes framework from the previous section, but now applied to the domestic intervention variable $\delta = 0, 1$. Define, as in Section 2, the binary random variable $\lambda = 0, 1$, an instrument. This variable takes the value of 1 whenever the base economy intervenes in its own interest rates, and it is 0 otherwise. Later, we will allow λ to be a continuous random variable, just as we did with δ earlier and discussed in Section 2. For the subpopulation of pegs, that is $q = 1$, we can think of the observed δ as coming from a mixture of latent random variables, δ_0 and δ_1, using potential outcomes notation, this time based on whether or not the base economy intervenes on its own interest rate—that is, $\lambda = 0, 1$. Specifically:

$$\delta = \delta_0 + \lambda(\delta_1 - \delta_0)\,. \tag{4}$$

Angrist and Imbens (1994) next assume that λ is independent of $(\boldsymbol{y}_0, \boldsymbol{y}_1, \delta_0, \delta_1)$ and also make a monotonicity assumption similar to our Assumption 1. Based on these assumptions, Angrist and Imbens (1994) define the local average treatment effect or LATE as:

$$\mathcal{R}_{LATE} = E(\boldsymbol{y}_1 - \boldsymbol{y}_0|\delta_1 - \delta_0 = 1)\,, \tag{5}$$

in other words, we can compute the impulse response using base country interventions when they to lead to domestic policy interventions.

Next, assume, in parallel fashion to our discussion of the stochastic processes for $\boldsymbol{y}_1$ and $\boldsymbol{y}_0$, that:

$$\delta_j = m_j + (\boldsymbol{x} - E(\boldsymbol{x}))\boldsymbol{g}_j + \eta_j\,. \tag{6}$$

with $E(\delta_j) = m_j$. As before, assume that $\boldsymbol{g}_0 = \boldsymbol{g}_1 = \boldsymbol{g}$ to simplify matters. Moreover, as in Wooldridge (2010), assume $\eta_0 = \eta_1 = \eta$. That is, assume that the stochastic part of δ_j is common for $j = 0, 1$.

Now, combining expressions (4) and (6) and given a sample of data, we can write:

$$\delta_{i,t} = a_i + \boldsymbol{x}_{i,t}\, \boldsymbol{g} + \lambda_{i,t}\, b + \eta\,, \tag{7}$$

a direct parallel to expression (1). At this stage, we consider the more general setting in which $-\infty < \lambda < \infty$ in which case the parameter b is normalized to be a 1 percentage point (100 bps) change of effective base country rates (effective in the sense that rate changes are scaled appropriately by the degree of capital openness as discussed earlier). This expression is basically the first-stage regression in instrumental variable estimation. The advantage of using the δ and λ notation is now becoming apparent. Any transformation of the intervention variable designed to account for nonlinearity and state-dependence modifies the first stage regression appropriately (so that we do not run afoul of Jensen's inequality, that is, the expected value of a function is not the same as the function of the expected values). Moreover, we now state assumptions for estimation with instrumental variable methods that may be valid for λ but may not be valid for z directly. Specifically:

Assumption 3. Relevance and exogeneity

$$\begin{aligned} L(\delta|\boldsymbol{x},\lambda;q=1) &\neq L(\delta|\boldsymbol{x};q=1) && && \textit{relevance} \\ L(\boldsymbol{y}_j|\boldsymbol{x};\delta;\lambda;q=1) &= L(\boldsymbol{y}_j|\boldsymbol{x};\delta;q=1) && \textit{for } j=0,1 && \textit{exogeneity} \end{aligned} \tag{8}$$

where, for example, $L(\delta|\boldsymbol{x};\lambda)$ refers to the linear projection of δ on $\boldsymbol{x}$ and λ as in expression (7).

This assumption is milder than making assumptions on conditional expectations since identification is based on the usual covariance between instrument and intervention. Note that we explicitly condition on $q=1$ to emphasize that in our trilemma-based application these assumptions only hold for the subpopulation of pegs. We note that one of the robustness checks that we introduce later allows for failure of the exclusion restriction in expression (8).

We now have the ingredients required to estimate the causal effects of a policy intervention for the subpopulation of pegs. Using Assumptions 1 and 3, and given a sample of panel data, we can now estimate the following set of local projections using instrumental variables (LP-IV):

$$y_{i,t+h} = \alpha_{i,h} + \boldsymbol{x}_{i,t}\gamma_h + \widehat{\delta}_{i,t}\beta_h + \nu_{i,t+h}\,; \qquad \text{for } h=0,\ldots,H-1\,, \tag{9}$$

which can be compared to the LP-OLS form at (2), and where the estimates from expression (7) are given by:

$$\widehat{\delta}_{i,t} = \hat{a}_i + \boldsymbol{x}_{i,t}\hat{g} + \lambda_{i,t}\hat{b}\,. \tag{10}$$

As before, the impulse response $\mathcal{R}_{LATE} = E(\boldsymbol{y}_1 - \boldsymbol{y}_0|\lambda;q=1) = \boldsymbol{\beta} = (\beta_0,\ldots,\beta_{H-1})'$ can be estimated from the sequence of equations in expression (9). Several remarks are worth making. First, the control vector $\boldsymbol{x}$ will include contemporaneous values of all the variables except the outcome variable (since we begin at $h = 0$). This is to provide insurance against variation in the policy intervention that could have been explained by information observed concurrently with the policy treatment. Second, the starting point for our analysis is to set $\delta_{i,t} = \Delta r_{i,t}$ and $\lambda_{i,t} = z_{i,t}\, k_{i,t}$ but later we consider state-dependence. Third, expressions (9) and (10) include fixed effects. Fourth, standard errors are estimated using a clustered-robust covariance matrix estimator. This option allows for a completely unrestricted specification of the covariance matrix of the residuals in the time series dimension taking advantage of the cross-section. This conveniently takes care of serial correlation in the residuals induce by the local projection setup.

So far we have a way of computing $\mathcal{R}_{LATE}$ for pegs. However, it turns out that there are instruments available to examine the subpopulation of bases. Using a completely different instrumental variable based on Romer and Romer (2004) for the U.S. and extended to the U.K. by Cloyne and Hürtgen (2016), we also compute the $\mathcal{R}_{LATE}$ for the U.S. and the U.K. in the post-WW2 period using the same methods. The instrument reflects staff forecasts from the staff at each of these central banks and is meant to capture the information available to policymakers when they made decisions on policy. This is the best estimate of the LATE for the subpopulation of bases. Despite using this alternate instrument, which is motivated quite differently from the new trilemma instrument we propose, and which is available only for two base economies, the derived LP-IV estimates of $\mathcal{R}_{LATE}$ for the U.S. and the U.K. in the post-WW2 era will turn out to be encouragingly similar to the estimates of $\mathcal{R}_{LATE}$ obtained for the subpopulation of pegs using the trilemma instrument.

5. Monetary policy interventions: understanding the subpopulations

Throughout its history, a country can fall into any of the following bins defined earlier: *pegs*, *floats*, and *bases*. For example, during Bretton Woods, Germany was in a *peg* to the dollar. With the end of Bretton Woods, and later the introduction of the European Monetary System, we consider Germany to become a *base* for many European economies. And there are other periods where we classify Germany as a *float*, as was the case for much of the interwar period. Other than bases, all other countries are either floats or pegs, depending on the period.

When presenting results, we will always measure and display the outcome variable in deviations relative to its initial value in year 0, with units shown in percent of the initial year value (computed as log change times 100), except in the case of interest rates where the response will be measured in units of percentage points. The policy intervention variable will be defined as the one-year change in the short-term interest rate in year 0, and normalized in all cases to a 1 percentage point, or 100 basis points (bps) increase.

The vector of explanatory variables includes a rich set of macroeconomic controls consisting of the first-difference of the contemporaneous values of all variables (excluding the response or outcome variable), and up to 2 lags of the first-difference of all variables, including the response variable. The list of macroeconomic controls is: log real GDP per capita; log real real consumption per capita; log real 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.[6]

In almost all respects, we found that this estimation setup produced stable outcomes. However, in line with the well-known "price puzzle" literature (e.g., Eichenbaum 1992; Sims 1992; Hanson 2004), we found that there was substantial instability in the coefficients of the control variables, and that this finding was driven by the postwar high-inflation period of the 1970s. The traditional resolution of this puzzle has been to include commodity prices as a way to control for oil shocks. Given the constraints of our data, we addressed this issue by allowing the controls to take on a potentially different coefficient for the subsample period of years from 1973 to 1980 inclusive, thus bracketing the volatile period of the two oil crises.

5.1. LP-OLS: subpopulation $\mathcal{R}_{ATE}$ under conditional mean independence

We begin our analysis by following the older selection-on-observables tradition. We naïvely estimate via LP-OLS the effect of an interest rate intervention on output (measured by log real GDP per capita) and prices (measured by log CPI)—two variables that commonly feature in many central bank mandates. These results are provided in Table 3 and are based on a panel regression that allows the relevant coefficient estimates to vary for each of the three subpopulations that we consider: pegs, floats and bases.

These estimates are a natural benchmark: *if* regression control is sufficient to achieve identification, then we could quite easily obtain estimates of $\mathcal{R}_{ATE}$ simply by averaging standard panel-based estimates across subpopulations. Hence the table evaluates whether estimates across subpopulations are statistically different from one another. In addition, we also provide a joint test that, over the 5 horizons considered, the effect of interest rates on output and prices is zero. The analysis is conducted over the full and the post-WW2 samples.

Consider the output responses first, reported in columns 1–3. Full sample results indicate some minor differences across subpopulations. The p-values of the null that the coefficients are equal is reported in column 4. The differences are economically minor, however. The post-WW2 results in column 4 suggest that if anything, the differences are even less important over this sample. Generally speaking, the coefficient estimates have the expected signs. An increase in interest rates causes output to decline. Note that in all cases the effect is statistically different from zero as reported in the rows labeled $H_0 : sATE = 0$, by which we denote "subpopulation ATE."

The price responses reported in columns 5–7 fit intuition less neatly. The overall effect of an interest rate increase on prices in the full sample is essentially null for pegs and floats (columns 5 and 6 respectively), but negative for bases with a -0.85 significant response in year $h = 4$. The

[6]The data are described in more detail in Jordà, Schularick, and Taylor (2017), and its online appendix.

Table 3: *LP-OLS. Real GDP per capita and CPI price responses to interest rates*

Responses at years 0 to 4 (100× log change from year 0 baseline).

(a) Full sample	*Output response*			*P=F=B*	*Price response*			*P=F=B*
Year	Pegs (1)	Floats (2)	Bases (3)	p-value (4)	Pegs (5)	Floats (6)	Bases (7)	p-value (8)
$h=0$	0.09***	-0.11	0.22***	0.01	0.15**	0.35***	-0.05	0.01
	(0.03)	(0.10)	(0.06)		(0.07)	(0.10)	(0.08)	
$h=1$	-0.17	-0.51***	-0.20**	0.14	0.17	0.57**	-0.11	0.03
	(0.11)	(0.15)	(0.08)		(0.15)	(0.24)	(0.14)	
$h=2$	-0.25*	-0.80***	-0.53***	0.06	0.15	0.40	-0.28	0.19
	(0.14)	(0.22)	(0.12)		(0.20)	(0.41)	(0.18)	
$h=3$	-0.34**	-0.86***	-0.43***	0.17	0.07	0.07	-0.58**	0.22
	(0.17)	(0.23)	(0.09)		(0.28)	(0.54)	(0.23)	
$h=4$	-0.33	-0.84***	-0.34**	0.20	0.00	-0.06	-0.85**	0.22
	(0.20)	(0.31)	(0.13)		(0.39)	(0.64)	(0.33)	
$H_0 : sATE = 0$	0.00	0.01	0.01		0.26	0.01	0.93	
Observations		1253				1285		

(b) Post-WW2	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$h=0$	0.04*	-0.01	0.13**	0.23	0.14**	0.16**	0.03	0.19
	(0.02)	(0.08)	(0.05)		(0.06)	(0.07)	(0.06)	
$h=1$	-0.18	-0.31**	-0.29***	0.62	0.22*	0.36**	0.11	0.30
	(0.11)	(0.15)	(0.10)		(0.11)	(0.18)	(0.12)	
$h=2$	-0.29**	-0.55***	-0.66***	0.08	0.13	0.23	-0.04	0.59
	(0.15)	(0.20)	(0.13)		(0.17)	(0.30)	(0.15)	
$h=3$	-0.37**	-0.52***	-0.66***	0.17	-0.09	-0.18	-0.41**	0.55
	(0.16)	(0.19)	(0.10)		(0.24)	(0.40)	(0.19)	
$h=4$	-0.35*	-0.52**	-0.72***	0.21	-0.23	-0.40	-0.69***	0.45
	(0.19)	(0.24)	(0.12)		(0.32)	(0.49)	(0.26)	
$H_0 : sATE = 0$	0.00	0.01	0.00		0.00	0.00	0.16	
Observations		897				929		

Notes: *** $p<0.01$, ** $p<0.05$, * $p<0.1$. P denotes pegs, F floats, B bases. Cluster robust standard errors in parentheses. Full sample: 1870–2013 excluding WW1: 1914–1919 and WW2: 1939–1947. PostWW2 sample: 1948–2013. The column *P=F=B* displays the p-value of the null that for a given horizon h, estimates of the corresponding elasticity are equal across subpopulations. $H_0 : sATE = 0$ refers to the null that the coefficients for $h = 0, \ldots, 4$ are jointly zero for a given subpopulation. See text.

picture changes somewhat for the post-WW2 subsample. Responses are essentially zero for $h = 0, 1$, and 2. Negative signs appear for $h = 3$, and 4, but the responses are generally not very different from zero in the statistical sense (except for bases, again).

What are the main takeaways from the naïve LP-OLS estimates Table 3? On first glance there is little evidence that anything is amiss. Output and price responses across subpopulations are similar, have the expected signs, and are statistically significant (although for prices only after year $h = 3$). On average across subpopulations, post-WW2 results indicate that a one percent increase in interest rates would reduced output and price levels about 0.5 percentage points over 5 years, roughly a 0.1 percent in annual rate of decline. The price responses offer a less reassuring picture, in large part because the responses are generally insignificant and often have the "wrong" sign early on. The next step is to examine the estimates of the local average treatment effect for the pegs and for bases. Any departure from the parameter estimates reported here would be indicative of a violation of the conditional mean independence assumption.

5.2. LP-IV $\mathcal{R}_{LATE}$: two instruments, two subpopulations

We now compare LP-OLS estimates with LP-IV estimates based on our trilemma instrument for the subpopulation of pegs by matching the samples. This will generate small differences between the LP-OLS estimates in Table 4 and those in columns 1 and 5 in Table 3. We calculate the $\mathcal{R}_{LATE}$ of an interest rate intervention and evaluate any attenuation bias from violations of conditional mean independence using a Hausman test. Table 4 summarizes the main results.

The table is organized as follows. The output and price responses in columns 1 and 4 are LP-OLS estimates over the same sample as the LP-IV estimates reported in columns 2 and 5. Column 3 reports the p-value of the Hausman test of the null that the estimate in column 1 is equal to that in column 2; likewise, column 6 reports the Hausman test for columns 4 and 5. We check (again) whether the trilemma instrument is weak with Kleibergen-Paap tests. Finally, we test the null that all ATE coefficients are jointly zero, reporting the p-value of the test in the row labeled $H_0 : sATE = 0$.

The first task is to compare the LP-OLS responses reported in columns 1 and 4 here, with those reported in Table 3 in columns 1 and 5. Recall that in Table 3 we estimate the model using all observations, but allow coefficients to vary by subpopulation. The differences are relatively minor, owing to slight differences in the sample used given the availability of the instrument.

The important result of Table 4 is the size of the attenuation bias in the case of LP-OLS compared to LP-IV. The differences are economically sizable and statistically significant as indicated by the Hausman tests of columns 3 and 6. Conditional mean independence clearly fails. Using LP-OLS estimates (column 1) and the full sample, output would be estimated to be about 0.18% lower four years after an increase in interest rates of 1%. In contrast, the LP-IV effect is measured to be nearly a 2.9% decline, or about a 0.5% annualized rate of lower growth. A similar pattern is observable for the price response. Full sample LP-OLS estimates are largely insignificant and often have the wrong sign. LP-IV estimates are sizable, significant, and have the right sign.

Table 4: *LP-OLS vs. LP-IV. Attenuation bias of real GDP per capita and CPI price responses to interest rates. Trilemma instrument. Matched samples*

Responses at years 0 to 4 (100× log change from year 0 baseline).

(a) Full sample	*Output response*		*OLS=IV*	*Price response*		*OLS=IV*
Year	LP-OLS (1)	LP-IV (2)	p-value (3)	LP-OLS (4)	LP-IV (5)	p-value (6)
$h = 0$	0.11***	-0.21*	0.01	0.10*	-0.21	0.11
	(0.03)	(0.11)		(0.05)	(0.19)	
$h = 1$	-0.18*	-0.99***	0.00	0.17	-0.68**	0.01
	(0.10)	(0.23)		(0.11)	(0.33)	
$h = 2$	-0.22	-1.88***	0.00	0.02	-1.57***	0.00
	(0.16)	(0.33)		(0.18)	(0.43)	
$h = 3$	-0.26	-2.20***	0.00	-0.22	-2.84***	0.00
	(0.21)	(0.42)		(0.30)	(0.70)	
$h = 4$	-0.18	-2.87***	0.00	-0.39	-3.83***	0.00
	(0.25)	(0.54)		(0.43)	(0.92)	
KP weak IV		87.83			70.57	
$H_0 : sATE = 0$	0.00	0.00		0.01	0.00	
Observations	667	667		667	667	

(b) Post-WW2	(1)	(2)	(3)	(4)	(5)	(6)
$h = 0$	0.06***	-0.02	0.24	0.05	0.17	0.39
	(0.02)	(0.07)		(0.05)	(0.14)	
$h = 1$	-0.13	-0.75***	0.01	0.11	0.08	0.89
	(0.09)	(0.25)		(0.08)	(0.26)	
$h = 2$	-0.22*	-1.58***	0.00	-0.03	-0.46	0.18
	(0.13)	(0.35)		(0.13)	(0.32)	
$h = 3$	-0.24	-1.70***	0.00	-0.27	-1.21***	0.03
	(0.16)	(0.37)		(0.22)	(0.42)	
$h = 4$	-0.17	-2.20***	0.00	-0.47	-1.78***	0.02
	(0.20)	(0.50)		(0.34)	(0.54)	
KP weak IV		91.04			77.14	
$H_0 : sATE = 0$	0.00	0.00		0.01	0.01	
Observations	522	522		522	522	

Notes: *** p <0.01, ** p <0.05, * p <0.1. Cluster robust standard errors in parentheses. Full sample: 1870–2013 excluding WW1: 1914–1919 and WW2: 1939–1947. PostWW2 sample: 1948–2013. *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. $H_0 : sATE = 0$ refers to the p-value of the test of the null hypothesis that the coefficients for $h = 0, \ldots, 4$ are jointly zero for a given subpopulation. *OLS=IV* shows the p-value for the Hausman test of the null that OLS estimates equal IV estimates. See text.

Table 5: *LP-OLS vs. LP-IV. Attenuation bias of real GDP per capita and CPI price responses to interest rates. U.S. and U.K. using RRCH instrument.*

Responses at years 0 to 4 (100× log change from year 0 baseline).

RRCH IV	*Output response*		*OLS=IV*	*Price response*		*OLS=IV*
Year	LP-OLS (1)	LP-IV (2)	p-value (3)	LP-OLS (4)	LP-IV (5)	p-value (6)
$h = 0$	0.35***	0.05	0.74	0.19**	-0.17	0.15
	(0.10)	(0.40)		(0.08)	(0.31)	
$h = 1$	-0.04	-0.59	0.41	0.58***	0.32	0.61
	(0.20)	(0.74)		(0.19)	(0.47)	
$h = 2$	-0.54	-1.71	0.19	0.63**	0.24	0.66
	(0.34)	(1.20)		(0.30)	(0.80)	
$h = 3$	-0.50	-2.19	0.18	0.34	-1.00	0.31
	(0.45)	(1.44)		(0.38)	(1.44)	
$h = 4$	-0.37	-1.88	0.23	-0.03	-3.05	0.16
	(0.46)	(1.41)		(0.47)	(2.42)	
KP weak IV		4.25			5.73	
$H_0 : sATE = 0$	0.00	0.35		0.00	0.06	
Observations	71	71		71	71	

Notes: *** p <0.01, ** p <0.05, * p <0.1. Cluster robust standard errors in parentheses. RRCH refers to the Romer and Romer (2004) and Cloyne and Hürtgen (2016) IV. U.S. sample: 1969–2007. U.K. sample: 1976–2007. *KP weak IV* refers to the Kleibergen-Paap test for weak instruments. $H_0 : sATE = 0$ refers to the p-value of the test of the null hypothesis that the coefficients for $h = 0, \ldots, 4$ are jointly zero for a given subpopulation. *OLS=IV* shows the p-value for the Hausman test of the null that OLS estimates equal IV estimates. See text.

Comparing the full sample results with the post-WW2 results we find differences in the output response to be relatively minor. The price response, however, becomes somewhat delayed after WW2. The LP-IV response suggests that on impact and the year after, the price response is essentially zero although by year 4, prices are expected to be about 1.8% lower than they were four years earlier. Tests for weak instruments suggest the trilemma instrument is relevant and tests of the null that the $\mathcal{R}_{LATE}$ estimated with LP-IV is statistically different from zero. Interest rates have a strong effect on output and prices for the subpopulation of pegs.

Next, Table 5 compares these results with estimates based on a different instrumental variable and subpopulation. We turn to the Romer and Romer (2004) instrument for the U.S., as updated and extended to the U.K. by Cloyne and Hürtgen (2016). We shall henceforth refer to this instrument as *RRCH*. Both the U.S. and the U.K. can be thought of as belonging to the subpopulation of bases and thus provide the best approximation of the $\mathcal{R}_{LATE}$ results for this group, where we should note that a similar IV for Germany as a base is not (yet) available.

The results in Table 5 are organized in a manner similar to Table 4. The table reports the output and price responses estimated by LP-OLS and LP-IV using the RRCH instrument. We note that the RRCH instrument is available only from 1969 to 2007 for the U.S. and 1976 to 2007 for the U.K. Because of the abbreviated sample, we limit the control set to save on degrees of freedom. We allow up to 3 lags of interest rates, output and inflation, but omit all other controls. This parsimonious specification still allows coefficients to vary over the oil crisis period of 1973–1980, as before. In terms of a formal model, one can think of this specification as the empirical counterpart to a three-variable New Keynesian VAR specification.

We can see that the LP-OLS and LP-IV estimates of the output response appear similar from the statistical perspective of the Hausman test reported in column 3. However, the sample is rather limited. Only 71 observations taxed by 11 regressors (a different constant for U.S. and U.K. observations and nine regressors). Economically speaking, there is a fair amount of attenuation bias. By year $h = 4$ the LP-OLS response is -0.37 compared to -1.88 for LP-IV. That said, both methods generally deliver the correct sign and the responses have a similar shape. The differences are more apparent for the price response. The LP-OLS response of prices to an interest rate intervention (column 4) is economically and statistically small, with coefficients that have the wrong sign (except for the last). The LP-IV response has a similarly muted response initially but it becomes increasingly negative. By year 4, prices are expected to be about 3.1% lower than they would otherwise would be, a response that is similar to that based on the trilemma IV as reported in Table 4.

5.3. The exclusion restriction: a robustness check

It is well-understood that without overidentification—that is, a situation when there are more instruments than endogenous variables—the exclusion restriction in expression (8) cannot be tested. Economically speaking, a violation could occur if base rates affect home outcomes through channels other than movements in home rates. Such channels are sometimes referred to as a spillover effects. Our problem offers an opportunity to assess such spillover effects. As shown in Table 1, the trilemma channel operates for pegs but not for floats. Information from this subpopulation will turn out to be particularly helpful. Consider a simple example to present the basic idea. Appendix A contains formal derivations. As before, let y be a univariate outcome variable, δ the intervention, and λ the instrument. We abstract from the constant term, controls, state dependence, and any other complication for now. The standard IV setup consists of the first and second stage regressions given by:

$$\begin{aligned} \delta &= \lambda\, b + \eta\,, \\ y &= \hat{\delta}\,\beta + \lambda\,\phi + \nu\,. \end{aligned} \tag{11}$$

Typically we assume $E(\delta\,\nu) \neq 0$, but $E(\lambda\,\nu) = 0$. The exclusion restriction refers to the assumption that $\phi = 0$. If this restriction were not to hold, it is easy to see that

$$\hat{\beta}_{IV} \xrightarrow{p} \beta + \frac{\phi}{b}\,.$$

This last expression is both simple and intuitive: the bias induced by the failure of the exclusion restriction depends on both the size of the failure (ϕ) *and* the strength of the instrument (b). Weaker instruments will tend to make the bias worse. This point was made in, for example, Conley, Hansen, and Rossi (2011).

The float subpopulation ($q = 0$) contains useful information that we now exploit. Continue to assume that $E(\lambda\, \nu) = 0$. We think this is justified since large base economies with monetary policy autonomy are unlikely to consider the macroeconomic outlook of smaller floats when setting rates. Hence, consider estimating (11) using OLS when $q = 0$. Estimates of the intervention effect β and the spillover effect ϕ will be biased as long as $E(\delta\, \nu) \neq 0$ and $b \neq 0$. However, it is easy to show (under standard regularity conditions) that the equivalent OLS estimates of expression (11) are such that

$$\left.\begin{array}{lcl} \hat{\beta}_{OLS} & \xrightarrow{p} & \beta - \theta \\ \hat{\phi}_{OLS} & \xrightarrow{p} & \phi + b\,\theta \end{array}\right\} \qquad \text{with a bias term} \qquad \theta = \frac{E(\delta\,\nu)E(\lambda^2)}{E(\lambda\,\delta)^2 - E(\lambda^2)E(\delta^2)}\,. \tag{12}$$

Expression (12) is intuitive. If $E(\delta\, \nu) > 0$, then $\theta > 0$ and the effect of domestic interest rates on outcomes, $\hat{\beta}_{OLS}$, will be *attenuated* by the bias term θ. Similarly, the spillover effect, $\hat{\phi}_{OLS}$, will be *amplified* by an amount $b\,\theta$. This amplification will be larger the stronger the correlation between δ and λ, as measured by the pseudo first-stage coefficient b.

The difference between OLS and IV estimates reported in Table 4 suggests that there is considerable attenuation bias in $\hat{\beta}$. The implication is that simple OLS will tend to make the spillover effect seem *larger* than it really is, and the interest rate response *smaller* than it really is. Of course, if $E(\delta\, \nu) < 0$, then $\theta < 0$, and the sign of the biases would be reversed. A priori the direction of the bias is ambiguous, as we cautioned earlier.

Without loss of generality, suppose that $\beta = \aleph\phi$, that is, the true domestic interest rate effect on outcomes is a scaled version of the spillover effect from the foreign interest rate. In this case,

$$\hat{\phi}(\aleph) = \frac{(\hat{\phi}_{OLS} + \hat{b}\hat{\beta}_{OLS})}{1 + \aleph\hat{b}} \;\xrightarrow{p}\; \beta(\aleph)\,. \tag{13}$$

Taking $\aleph$ as given, we can use a control function approach to correct our LP-IV estimates of $\mathcal{R}_{LATE}$ for biases due to potential spillover effects. Expression (11) can then be rewritten as

$$(y - \lambda\hat{\phi}(\aleph)) = \delta\beta + \nu + \lambda(\hat{\phi}(\aleph) - \phi(\aleph))\,.$$

Moreover, the usual moment conditions imply that

$$E(\lambda(y - \lambda\hat{\phi}(\aleph))) = E(\lambda\,\delta)\beta + E(\lambda\,\nu) + E(\lambda^2(\hat{\phi}(\aleph) - \phi(\aleph)))\,,$$

$$\text{with } E(\lambda\,\nu) = 0\,, \;\text{ and }\; (\hat{\phi}(\aleph) - \phi(\aleph))\frac{1}{N_p}\sum_{j}^{N_p}\lambda_j^2 \;\xrightarrow{p}\; 0\,,$$

as long as

$$\frac{1}{N_p}\sum_j^{N_p}\lambda_j^2 \xrightarrow{p} Q_\lambda < \infty\,, \text{ and } N_f \longrightarrow \infty \text{ as } N_p \longrightarrow \infty\,,$$

with N_f and N_p denoting the sizes of the subpopulations of floats and pegs respectively.

From this, we can now present a extension of our IV estimator *corrected for potential spillover effects*, where this new variant is constructed by subtracting the spillover term from the outcome variable in the standard IV coefficient estimator, whereby

$$\hat{\beta}(\aleph) \equiv \frac{\frac{1}{N_p}\sum\lambda_j(y_j - \lambda_j\hat{\phi}(\aleph))}{\frac{1}{N_p}\sum\lambda_j\,\delta_j} \xrightarrow{p} \beta(\aleph)\,.$$

We assume that the sample sizes of both float and peg subpopulations tend to infinity. In practice, $\aleph$ is unknown. We proceed below by using economic arguments to provide an interval of plausible values $\aleph \in [\underline{\aleph}, \overline{\aleph}]$ over which we compute $\hat{\beta}(\aleph)$. This interval provides a sense of the sensitivity of our benchmark LP-IV estimates of $\mathcal{R}_{LATE}$ to potential spillover contamination.

Table 6 reports OLS estimates of expression (11), based on the float subpopulation (and by including the usual control set). We employ the same trilemma instrument, defined exactly as before, but now utilize it in the *float* subpopulation to operationalize our control function approach. The left-hand side variables are log real GDP per capita and log CPI price level. Table 6 also reports the coefficient associated with the pseudo-first stage regression of δ on λ. This provides an estimate of the parameter b in expression (12).

Table 6 makes clear the intuition behind expression (12). The interest rate responses of real GDP per capita reported in column 1 are economically small. They are statistically insignificant for the full sample estimates reported in panel (a) of the table, and only significant in years 3 and 4 in the Post-WW2 sample reported in panel (b). In contrast, the response to the instrument (think of it as a shock to the base country interest rate) is almost three times larger and significant. Price responses follow a different pattern, with responses to the own interest rate shock of the wrong sign, but responses to the base country interest rate (measured by the instrument) of the correct sign. This is a feature we will return to in the results reported below.

Finally, we note that the regression of the domestic interest rate on the instrument and the control set is generally non-zero, but about half to one third the magnitude of the coefficient estimated in the first stage regression for the peg subpopulation and reported in Table 1. Compare 0.20 for the floats with 0.40 for the pegs (using full sample estimates in the case for output). These results are consistent with those reported in Obstfeld, Shambaugh, and Taylor (2005).

Lastly, to make any progress we need auxiliary assumptions on $\aleph$, which cannot be determined from the data. We assert that it is natural to assume that $\aleph \geq 1$. That is, we assume that home rates affect outcomes at least as strongly as rates in the foreign base country. In order to provide bounds, we use a range of values of $\aleph$ between 1 and 4. In other words, the home rate effect is assumed to be 1 to 4 times larger than the base rate effect.

Table 6: *LP-OLS. Real GDP per capita and CPI price responses to domestic and base-country interest rates. Full and post-WW2 samples for subpopulation of exchange rate float economies*

Responses at years 0 to 4 (100× log change from year 0 baseline).

(a) Full sample	*Output response to*		$\delta = \lambda$	*Price response to*		$\delta = \lambda$
	δ	λ	p-value	δ	λ	p-value
	(1)	(2)	(3)	(4)	(5)	(6)
$h = 0$	-0.05	0.09	0.43	0.53**	-0.15	0.09
	(0.15)	(0.09)		(0.22)	(0.24)	
$h = 1$	-0.24	-0.13	0.80	1.17**	-0.30	0.11
	(0.26)	(0.22)		(0.49)	(0.50)	
$h = 2$	-0.23	-0.57**	0.48	1.45**	-0.62	0.10
	(0.31)	(0.22)		(0.66)	(0.71)	
$h = 3$	-0.21	-1.08***	0.10	1.35	-1.13	0.10
	(0.29)	(0.32)		(0.80)	(0.88)	
$h = 4$	-0.06	-1.40***	0.04	1.48	-1.70	0.06
	(0.30)	(0.45)		(0.89)	(1.00)	
δ on λ	0.20***			0.21***		
first stage estimate	(0.06)			(0.07)		
Observations	269			269		

(b) Post-WW2	(1)	(2)	(3)	(4)	(5)	(6)
$h = 0$	-0.06	-0.03	0.78	0.37	-0.36*	0.05
	(0.10)	(0.11)		(0.22)	(0.19)	
$h = 1$	-0.22	-0.28	0.86	0.58	-0.64	0.10
	(0.16)	(0.26)		(0.44)	(0.43)	
$h = 2$	-0.21	-0.84**	0.29	0.82	-0.86	0.13
	(0.26)	(0.36)		(0.62)	(0.62)	
$h = 3$	-0.34*	-1.51***	0.01	0.78	-1.31	0.13
	(0.18)	(0.33)		(0.71)	(0.78)	
$h = 4$	-0.41**	-1.90***	0.00	0.99	-1.87*	0.07
	(0.18)	(0.46)		(0.78)	(0.94)	
δ on λ	0.18*			0.16**		
first stage estimate	(0.09)			(0.07)		
Observations	210			210		

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Cluster robust standard errors in parentheses. δ is the short-term interest rate of the domestic economy considered and indexed by i; λ is defined as $\lambda_{i,t} = z_{i,t} \times k_{i,t}$ and refers to the base country short-term interest rate times the openness index. We use the U.K. as the base before 1939; we use the U.S. for the Bretton-Woods era (1946–1971) for all countries, and for Australia, Canada, Japan, and the U.K. after Bretton-Woods (1972 onward); and we use Germany for all remaining economies (all in Europe) in the post Bretton-Woods era (1972 onward). $\delta = \lambda$ refers to the test of the null that the coefficient for the float rate is equal to that for the instrument and report the p-value of the test. δ on λ refers to the coefficient of the regression of the float rate on the instrument (called b in expression (11)). See text.

Using these assumptions and our spillover correction, Figure 2 shows our adjusted estimates of $\mathcal{R}_{LATE}$, for a 100 bps impulse, as before. The figure displays responses to a 1% increase in home rates for real GDP per capita and the price level using alternative estimators. The red dashed lines show the responses reported in Table 4, columns 1 and 4, based on LP-OLS and the rich control set described above. As we noted, the responses for GDP generally have the correct sign but are economically and statistically small. This is in line with the attenuation bias described in expression (12). Next, the solid blue lines with associated point-wise error bands in grey show the LP-IV estimates reported in Table 4, columns 2 and 5. As we noted, the responses are considerably larger, both statistically (the null $H_0 : sATE = 0$ is rejected at the 1% level) and economically. Finally, the light green shaded region with a dashed border displays the range of responses that would result from our spillover correction using $\aleph \in [1, 4]$.

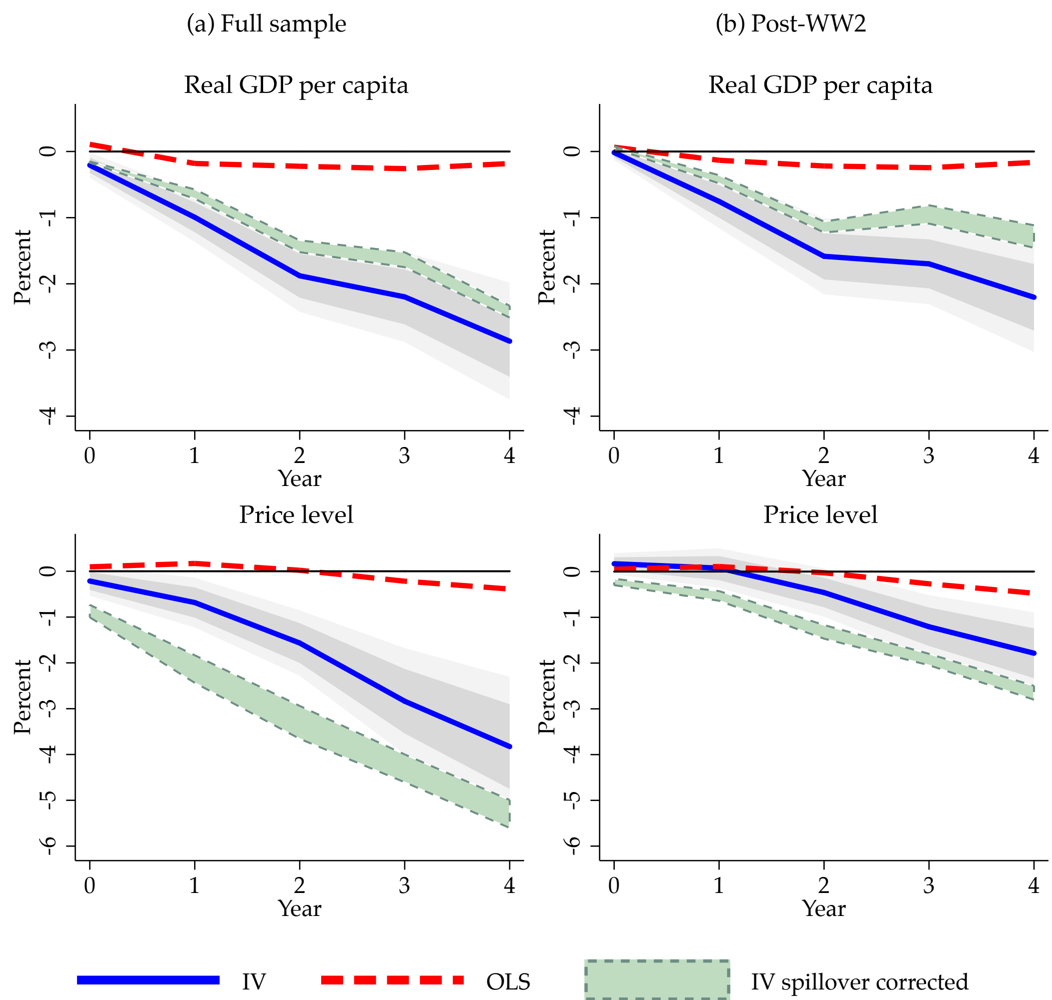
Several results deserve comment. First, the correction for spillover effects tends to attenuate the output responses as compared to our preferred LP-IV estimates. The correction suggests that the output response is about 0.5 to 1 percentage points less negative by year 4 than that reported using LP-IV alone. In the post-WW2 era this means that the cumulative change in output is probably closer to -1.5% (about -0.3% per annum) than to -2.5% (about -0.5% per annum). Interestingly, the response of prices is amplified. The reason is easy to see in Table 6. The LP-OLS estimate of the effect of domestic interest rates on prices is positive rather than negative. Meanwhile, the effect of base country interest rates is strongly negative. Thus, the correction makes the LP-IV price response more sensitive to interest rates, especially on impact. This feature has often been an achilles heel of the post-WW2 U.S. VAR literature, where findings often have a price response with the wrong sign or a price response that remains largely muted for a prolonged period of time even as the response of output shows a more immediate response. These results are reassuring. Our LP-IV estimates are reasonably robust to potential spillover effects under economically meaningful scenarios. Even if the spillover effects were large, the true estimate would be still be much closer to the LP-IV estimates that we report than to the LP-OLS estimates commonly found in the literature.

5.4. Some takeaways

The main findings so far can be summarized as follows. First, we report strong evidence of attenuation bias when comparing LP-OLS to LP-IV subpopulation. This evidence suggests that conditional mean independence, despite the extensive set of controls that we include, is insufficient to achieve identification of the causal effect of interest rates on output and prices.

Second, we find that the LP-IV responses calculated for pegs using our trilemma instrument and those calculated for the U.S. and U.K. bases using the RRCH instrument are reasonably close. Therefore going forward, we use both instruments in tandem to maximize statistical power. Although we can only speculate about the causal effects of interest rates for floats, the similarities of LP-OLS estimates reported in Table 3, and the results in Table 5 suggest that interest rate responses are similar for this subpopulation as well.

Figure 2: *Real GDP per capital and CPI price responses to a 1 percent increase in interest rates. LP-OLS, LP-IV, and spillover corrected LP-IV*



Notes: Full sample: 1870–2013 excluding WW1: 1914–1919 and WW2: 1939-1947. LP-OLS estimates displayed as a dashed red line, LP-IV estimates displayed as a solid blue line and 1 S.D. and 90% confidence bands, LP-IV spillover corrected estimates displayed as a light green shaded area with dashed border, using $\aleph \in [1, 4]$. See text.

Is our estimated $\mathcal{R}_{LATE}$ a good approximation to the unobserved $\mathcal{R}_{ATE}$ in all bins under different monetary and exchange-rate regimes? We argue that it is. Of course, pegs approach systematic monetary policy very differently than floats. In the latter, systematic policy sets interest rates to target explicit inflation and economic activity goals. In the former, a different systematic policy sets interest rates that mirror the base country to maintain the peg, while seeking other stability goals implicitly. In principle, pegs seek the same price and output stability objectives as bases. They do it by taking advantage of the credibility earned by the central bank in the base country.

Consider a highly stylized setting, with many details omitted for clarity. Suppose we are in a symmetric world with one base economy and one other economy that can either peg or float. If business cycles in the two economies were perfectly synchronized, and if the central bank in the base economy behaved optimally given the same policy goals as the other economy, the decision to peg or float would be indeterminate from the point of view of systematic monetary policy and irrelevant for economic outcomes.

The world is much more complicated, of course. But for us the relevant question is this: what would allow us to extrapolate the $\mathcal{R}_{LATE}$ measured for pegs to the $\mathcal{R}_{LATE}$ we are unable to estimate for floats? Don't they have very different approaches to systematic monetary policy? There is a two pronged response to this question. First, we have succinctly articulated a case for why differences between pegs and floats may be smaller than feared at first glance—indeed, our stylized example is a case where the systematic policies are equivalent. Second and perhaps more importantly, the causal effect identified for pegs is based *only* on non-systematic movements of base country interest rates. These are as close to an exogenous source of variation as we can get. Importantly, they are, by construction, silent about the policy choices a country makes.

In this sense, the key to the validity of our approach is no different than that found throughout the vast literature on empirical monetary policy. Unless the fundamental mechanics of the macroeconomy differ between pegs and floats, there is no reason to expect this causal effect to be different in the two subpopulations.

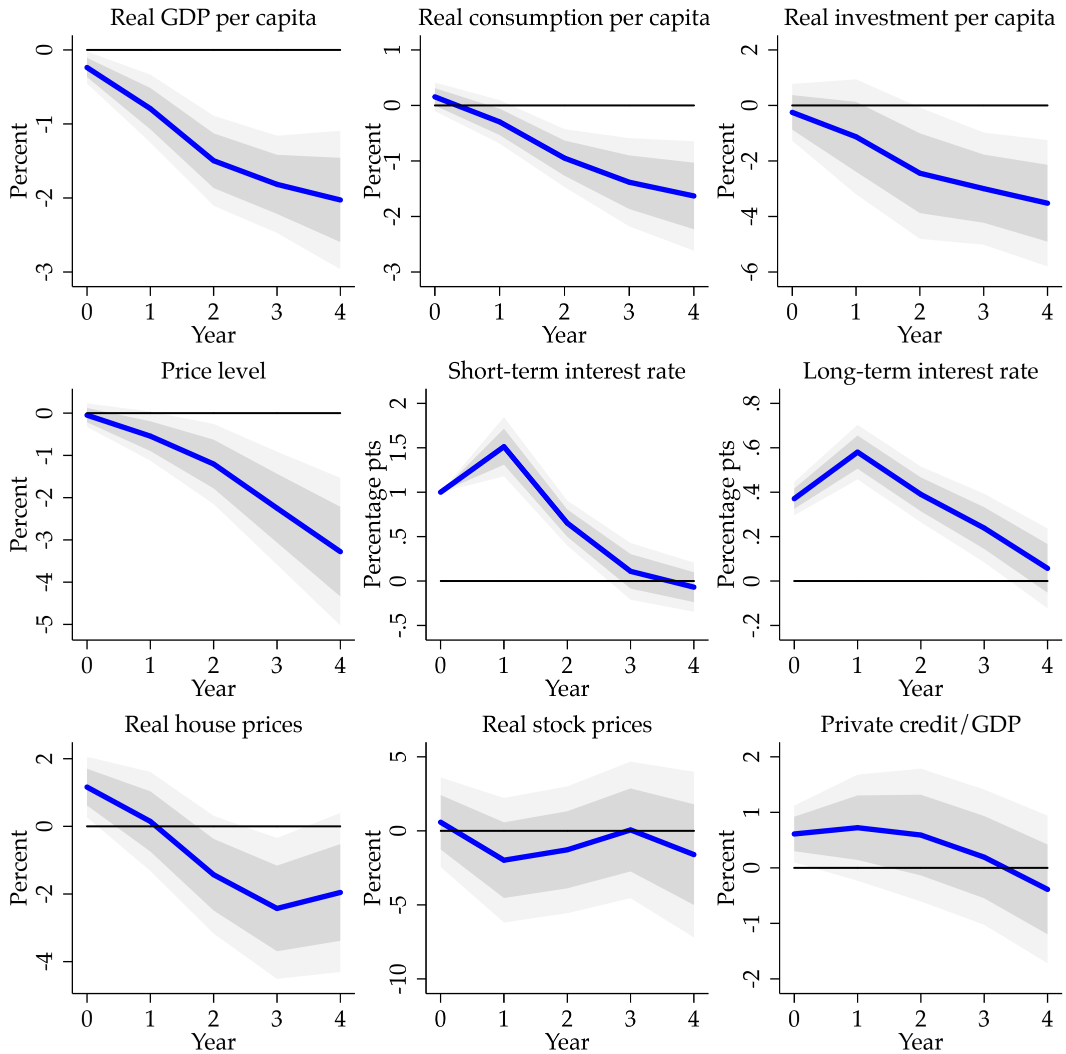
6. The causal effects of interest rates on the macroeconomy

In this section we briefly present a comprehensive study of the causal response of a wider array of macroeconomic outcomes to a short-term interest rate increase of +100 bps. Figure 3 summarizes the responses of our full set of variables for the full sample (the equivalent figure using the Post-WW2 sample only is virtually identical and provided in Appendix B for completeness as Figure B.1).

Starting at the top left chart of Figure 3, we see that a +100 bps increase in the short rate *causally* leads to a 2% decline in real GDP per capita (or about −0.5% per annum), a 1.5% decline in real consumption per capita, a 3.5% decline in real investment per capita, and a 3% decline in the price level (in the second row, first column), where all effects are relative to the no-change policy counterfactual and the measurements are cumulative over the horizon of 4 years.

Moving along the second row of charts in Figure 3, we look first at the own response of short-term interest rates to a +100 bps rate rise in year 0 (row 2, column 2). This path reflects the intrinsic persistence of changes in interest rates. In this case, short-term interest rates increase by +150 bps in year 1, drop back to +75 bps in year 2, and then decline to effectively zero in both years 3 and 4. The next chart (row 2, column 3) shows the response of long-term interest rates, which are, as is well-known, more subdued in amplitude than short rates; the long-term interest rate moves about half as much. A +100 bps rise in the short rate *causally* leads to the long rate rising +40 bps in year 1, rising to +60 bps in year 2, and then falling back towards zero by year 4.

Figure 3: *Full baseline results. Full sample*



Notes: Full sample: 1870–2013 excluding WW1: 1914–1919 and WW2: 1939-1947. LP-IV estimates displayed with a solid blue line and 1 S.D. and 90% confidence bands. See text.

Proceeding to the last row of charts in Figure 3 (columns 1 and 2), we can examine the responses of two key asset prices: a +100 bps rise in the short rate *causally* leads to a cumulative 2% decline in real house prices and a cumulative 1.5% decline in real stock prices over 4 years, again as compared to the no-change policy counterfactual. Although these real responses may appear quite small, bear in mind that the nominal responses are far larger in the negative direction, as reported in Figure B.2 in Appendix B. That is, nominal asset prices drop strongly and quickly in response to an interest rate hike. Overall, these asset price responses are consistent with a significant wealth-effect channel for monetary policy, working alongside the more often noted income-effect channel visible in the path of real GDP.

Finally, in the last row and column of Figure 3, we display the causal response of aggregate credit (bank lending to the nonfinancial sector relative to GDP). This chart shows that a +100 bps rise in the short rate has a relatively muted effect on the ratio of bank loans to GDP cumulated over four years, although this effect is in the end consistent with models where contractionary monetary policy leads to less demand and/or supply of credit. If anything, the effect is slightly more pronounced when using the Post-WW2 sample although the responses are very similar. Bauer and Granziera (2016) have reported similar patterns based on post-1970 OECD data. Loan contracts are difficult to undo in the short-run relative to the decline of GDP. Therefore, although nominal private credit declines on impact (not shown), the private credit to GDP ratio may take a bit longer to decline.

The takeaway from these impulse responses is clear. An exogenous shock to interest rates has sizable effects on real variables (larger than those measured using conventional VARs), but along the lines predicted by most monetary models with rigidities or frictions. Term structure responses conform very well with standard results in the literature. Nominal variables decline strongly. Perhaps the only variable that appears to be somewhat unresponsive to interest rates is the credit to GDP variable. Loans decline in response to a shock to interest rates, but their rate of decline matches closely the rate of decline in real economic activity.

Figures B.3 and B.4 in Appendix B report the full set of impulse responses to year 10 after impact for the full and post-WW2 samples respectively. These figures are provided to investigate the neutrality of money in the long-run. We are able to confirm that this is the case, although perhaps at a much slower rate than one would presume. Investment, the more volatile of the components of output, is nearly back to its starting value by year 10 only.

The responses of all nine variables displayed in Figures 3, and B.1, B.3, and B.4 in Appendix B, are consistent with the intuition gained from the long monetary tradition. The differences that we report with that literature are mostly about somewhat stronger effects than previously reported. The next step in our investigation is to assess the stability of these results when we allow them to depend on the state of the economy.

7. State dependence

In this section we address the possible state dependence of the impulse-response functions by stratifying our analysis of output and inflation responses. In particular, we use indicators for whether the economy is in a boom or slump, in low or high inflation, or in a credit surge or slowdown, to address big issues from the research literature and current policy debates.

Our first stratification in Figure 4 uses the output gap ("boom/slump"). The boom/slump stratification delivers a state-dependent analysis that echoes the analysis in Auerbach and Gorodnichecko (2013b) and Jordà and Taylor (2016) for fiscal policy, namely, that stabilization policy can have different effects in the boom (when the output gap is positive) versus the slump (when the output gap is negative). We measure the gap using a simple low-pass HP filter.

Our second stratification in Figure 5 uses inflation ("high/low"). This links with current debates about the effectiveness of monetary policy in an environment of unusually low inflation. Despite widespread belief in the "divine coincidence" (Blanchard and Galí 2005), the returns to ultra-loose monetary policy are debatable in this setting. Thus, we investigate state dependence based on high/low inflation environments (defined by an annual 2% CPI inflation rate cutoff).

Our third and final stratification in Figure 6 looks at the growth of credit relative to GDP in the 3 years leading up to the policy intervention, broken down by mortgage versus non-mortgage loans (adjusted for country-specific variation). The data on credit come from Jordà, Schularick, and Taylor (2017), and are now publicly available.[7] This stratification directly speaks to a nascent literature that investigates mechanisms by which household/firm heterogeneity in leverage can affect the effectiveness of monetary policy as discussed in the introduction.

7.1. State dependence in boom and slump episodes

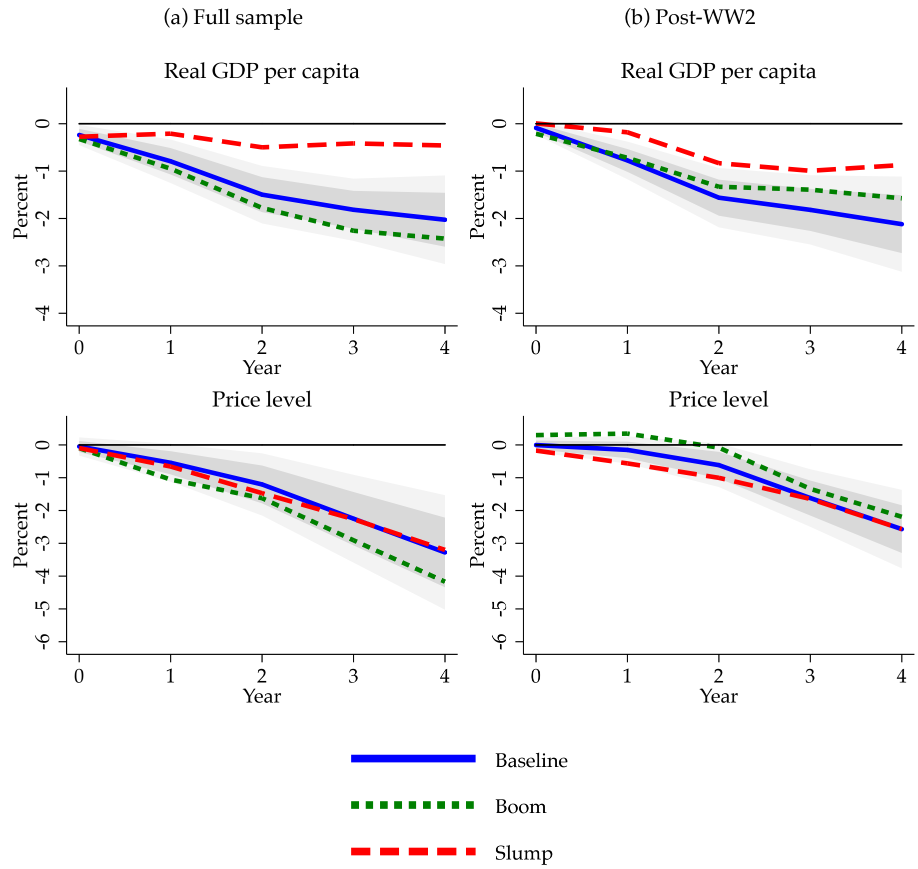
In the first set of nonlinear experiments shown in Figure 4, we set the state variable equal to 1 in booms and 0 in slumps. Booms (slumps) are years when log real GDP per capita is above (respectively, below) its long-run country-specific trend component, as measured by an HP filtered series with a very low-pass setting (λ_{HP}=100, annual).

These results are strongly indicative of an asymmetric macroeconomic response to interest rates. The experiment is always normalized to be a +100 bps increase in the short-term rate. This is done to facilitate impulse-response comparisons across states even though, of course, tight policy in a slump is unlikely. The response of real GDP per capita to monetary policy appears to be quite strong in booms (about -2.5% by year 4, full sample), but considerably weaker in slumps (about -0.5% by year 4, full sample). This difference of nearly 2% (closer to 1% Post-WW2) is broadly consistent with Angrist, Jordà, and Kuersteiner (forthcoming), Tenreyro and Thwaites (2016), and Barnichon and Matthes (2016). Evidence of asymmetry is less clear for the inflation response, however. Differences using the full sample are very minor and not very different from the average response. The Post-WW2 responses suggest that if the monetary authority tried to stimulate the economy (opposite to what is displayed but consistent with the outlook in many economies today), it would take 2 years before the effects would be felt in prices, although after 4 years the effects would be similar in size to those estimated for the boom state.

To sum up, a central bank rate tightening of $+100$ bps in a boom would have causal effects on output that are strongly contractionary on output going forward. In contrast, a central bank rate loosening of -100 bps in a slump would have causal effects on output that are, proportionately, only weakly expansionary. However, the effects on prices would be roughly symmetric across the two states. Thus, the "sacrifice ratios" for monetary policy are markedly different, and so are the incentives to pursue pre-emptive monetary policy, in one state versus the other.

[7] http://www.macrohistory.net/data/.

Figure 4: *State dependence: monetary policy has a weaker effect on output in slumps*



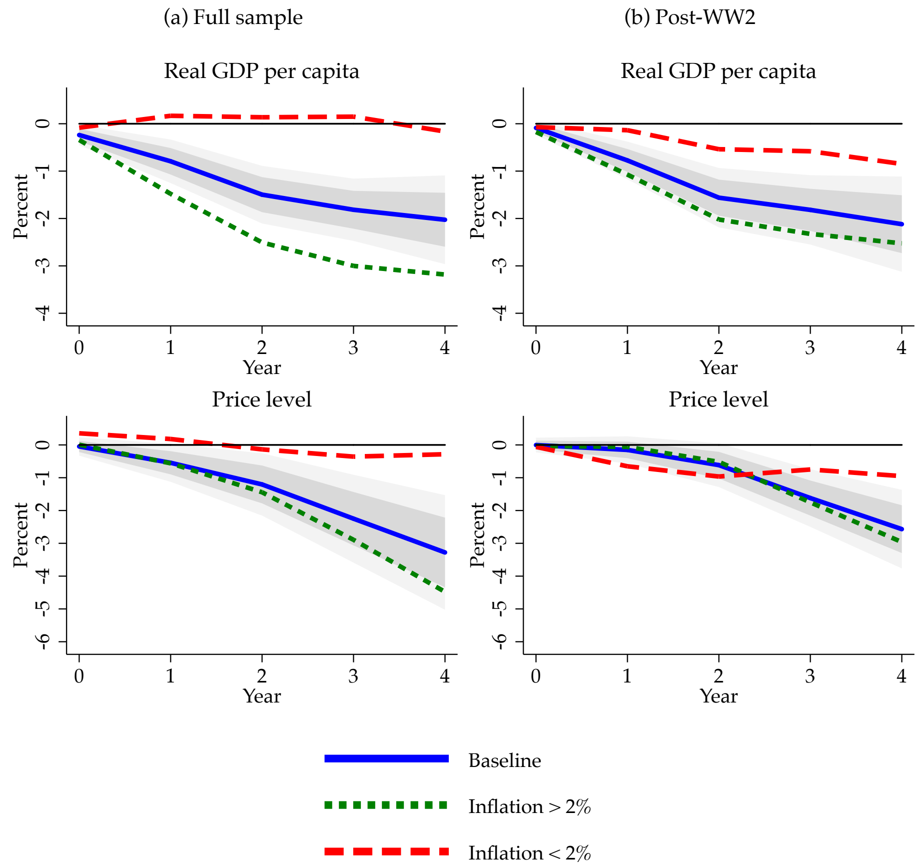
Notes: Full sample: 1870–2013 excluding world wars (1914–1919 and 1939–1947). Post-WW2 sample: 1948–2013. Linear LP-IV estimates displayed with a solid blue line and 1 S.D. and 90% confidence bands. Estimates stratified by the boom displayed with a green dotted line whereas estimates in the slumps are displayed with a red dashed line. See text.

7.2. State dependence in "lowflation" episodes

Our next set of nonlinear experiments, in Figure 5, concerns the hypothesis that monetary policy may have different effects in times of low inflation, a topic of rising interest since the advanced economies entered an era of "lowflation" as the Great Recession wore on after 2008. To examine this nonlinearity we chose to set an indicator state variable equal to 1 when inflation is low (at or below 2%) and 0 otherwise.

These results reinforce some findings in the previous section. Just as in Figure 4, we normalize the responses to a +100 bps increase in interest rates, while fully recognizing rates are unlikely to go up when there is lowflation. However, the normalization greatly facilitates the comparison across states. The response of real GDP to monetary policy appears to be quite strong when inflation is above 2%. In that scenario, the cumulated response at year 4 is slightly above -3%, somewhat

Figure 5: *State dependence: monetary policy has weaker effects when there is lowflation*



Notes: Full sample: 1870–2013 excluding world wars (1914–1919 and 1939–1947). Post-WW2 sample: 1948–2013. Linear LP-IV estimates displayed with a solid blue line and 1 S.D. and 90% confidence bands. Estimates stratified by the lowflation regime displayed with a red dashed line whereas estimates when inflation is above 2% are displayed with a green dotted line. See text.

stronger than the response in booms reported in figure 4, which was −2.5%. However, monetary policy loses most of its traction when the inflation rate dips below the 2% threshold and economies tip into a lowflation state. In the full sample results, we find that there is no effect on output although in the Post-WW2 sample, the effect on output is somewhat more visible (around −1%).

Taken together—the boom/slump and high/low inflation stratification—the historical data are consistent with the scenarios playing out in many advanced economies following the financial crisis. But the financial crisis and theories that try to explain its aftermath often build on what happens when a credit boom goes bust, as we discussed in the introduction (see, e.g., Schularick and Taylor 2012). Thus, the natural next step is to stratify the responses to monetary policy according to credit. This we do in the next subsection by further distinguishing between mortgage and non-mortgage credit growth.

7.3. State dependence in credit boom and slump episodes

Our final set of experiments is presented in Figure 6 using two panels. The results presented in panel A refer to above/below country-specific mean changes in mortgage credit over GDP, whereas panel B focuses instead on changes in non-mortgage credit over GDP.

In previous research (Jordà, Schularick, and Taylor 2016) we have found this distinction to matter and in the context of the theoretical literature, and the limitations of the data, it is the closest we can get to the evaluating household versus firm leverage. Historically, most mortgage credit refers to household rather than commercial real estate borrowing, a trend that has been increasing of late (e.g., the commercial share of mortgages in the U.S. has fallen in the past 40 years from about 40% to 25%). Non-mortgage credit on the other hand, is usually unsecured loans issued to firms rather than households. Hence one can interpret panel A as suggestive that household leverage affects the efficacy of monetary policy in a way that lending to firms does not, as shown in panel B.

In particular, panel A shows that in a period of rapid recent growth in mortgage credit, a 1% shock in interest rates can depress output by about 3 percentage points more than after a period of below average credit growth, 4 years after intervention. Meanwhile, prices respond much more to interest rates during periods of low recent growth in mortgage credit, relative to other times. In response to the same 1% interest rate shock discussed earlier, the decline in prices can be about 3 percentage points higher 4 years after intervention.

Meanwhile, panel B shows that there is essentially no difference in the response to monetary policy whether or not non-mortgage credit grows above or below average in the preceding 3 years. Thus, theories that stress heterogeneity coming from consumer leverage appear to find stronger support in our data than theories that rely on firm leverage instead. Furthermore, there seems to be little difference in results including or excluding pre-WW2 data. The effects that we report are fairly stable across eras despite considerable evolution in the financial sector.

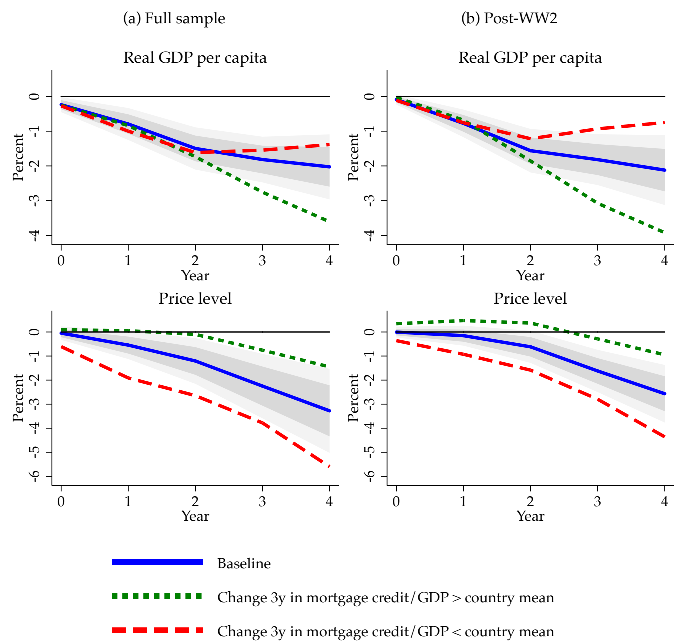
8. Conclusion

The effectiveness of monetary stabilization policy is not only a major policy concern but also an important matter of ongoing controversy among both theoretical and empirical macroeconomists. This paper argues that interest rates can have a considerable effect on macroeconomic outcomes. The source of the attenuation bias that we report suggests that common assumptions, often based on regression control arguments (conditional mean independence), do not provide an adequate basis for identification. Using a quasi-natural experiment in international finance and novel empirical methods, we show why this bias occurs and how to resolve it.

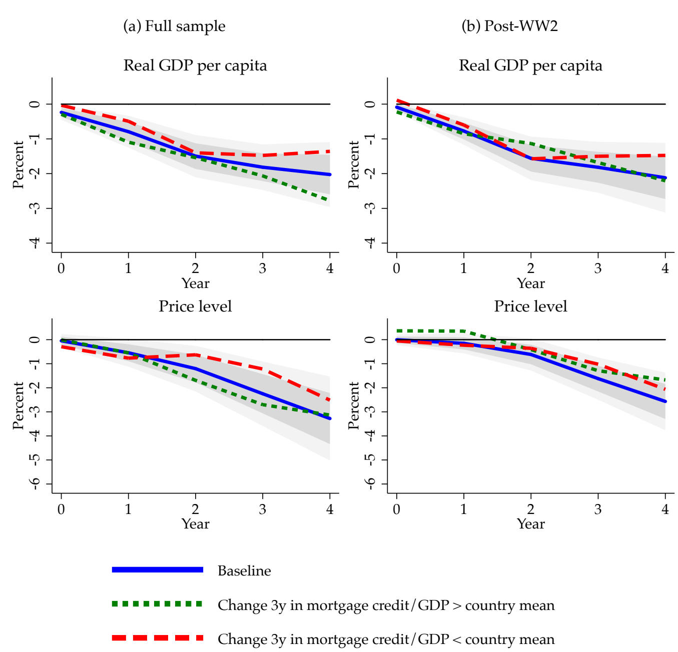
Monetary stabilization policy turns out to be state-dependent, a critical observation in the context of the Great Recession and its aftermath. Policymakers have faced a situation of consistent undershoot relative to their stated objectives and forecasts. A slower than expected growth trajectory has been seen in the U.S. economy, and worse yet in the U.K., Europe, and Japan; and we have

Figure 6: *State dependence: monetary policy and mortgage vs. non-mortgage credit*

A: *Mortgage credit*



B: *Non-mortgage credit*



Notes: Full sample: 1870–2013 excluding world wars (1914–1919 and 1939–1947). Post-WW2 sample: 1948–2013. Linear LP-IV estimates displayed with a solid blue line and 1 S.D. and 90% confidence bands. Estimates stratified with below country-average (3y) credit over GDP growth regime displayed with a red dashed line, and above country-average (3y) credit over GDP growth regime displayed with a green dotted line. See text.

seen persistent sub-2% inflation afflicting all of these economies. Amid worries of deflation risk and secular stagnation, the extensive and unconventional application of monetary policy tools has failed to shunt the macroeconomic locomotive onto a faster track. Debate centers on why the central banks got so derailed, and the subsequent lack of policy traction.

Our quantitative evidence from the largest advanced economy macroeconomic dataset ever assembled shows that all of these travails are, in a sense, nothing new. In conditions where output is weak, inflation is low, or credit is stagnant, the ability to stimulate the economy out of its torpor is that much more difficult. These challenging conditions have been prevalent in a number of western economies for almost a decade, but they were also present in the distant past. Our results therefore have profound implications for how today's monetary models are formulated and applied.

References

Angrist, Joshua D., Guido W. Imbens, and Donald B. Rubin. 1996. Identification of Causal Effects Using Instrumental Variables. *Journal of the American statistical Association* 91(434): 444–55.

Angrist, Joshua D., Òscar Jordà, and Guido Kuersteiner. Semiparametric Estimates of Monetary Policy Effects: String Theory Revisited. *Journal of Business and Economic Statistics.* Forthcoming.

Auclert, Adrien. 2017. Monetary Policy and the Redistribution Channel. NBER Working Paper 23451.

Auerbach, Alan J., and Yuriy Gorodnichenko. 2013a. Measuring the Output Responses to Fiscal Policy. *American Economic Journal: Economic Policy* 4(2): 1–27.

Auerbach, Alan J., and Yuriy Gorodnichenko. 2013b. Fiscal Multipliers in Recession and Expansion. In *Fiscal Policy After the Financial Crisis* edited by Alberto Alesina and Francesco Giavazzi. Chicago: University of Chicago Press, pp. 63–98.

Barnichon, Regis, and Christian Matthes. 2016. Gaussian Mixture Approximations of Impulse Responses and The Non-Linear Effects of Monetary Shocks. CEPR Discussion Paper 11374.

Barro, Robert J., and José F. Ursúa. 2008. Macroeconomic Crises since 1870. *Brookings Papers on Economic Activity* 39(1): 255–335.

Bauer, Gregory H., and Eleonora Granziera. 2016. Monetary Policy, Private Debt, and Financial Stability Risks. *International Journal of Central Banking* conference in San Francisco. November 21–22, 2016.

Bernanke, Ben S. and Mark Gertler. 1995. Inside the Black Box: The Credit Channel of Monetary Transmission. *Journal of Economic Perspectives* 9(4): 27–48.

Blanchard, Olivier, and Jordi Galí. Real Wage Rigidities and the New Keynesian Model. 2007. *Journal of Money, Credit and Banking* 39(S1): 35–65.

Christiano, Lawrence J., Martin S. Eichenbaum, and Charles L. Evans. 1999. Monetary Policy Shocks: What Have We Learned And To What End? *Handbook of Macroeconomics*, vol. 1, edited by John B. Taylor and Michael Woodford. Amsterdam: Elsevier, pp. 65–148.

Cloyne, James, Clodomiro Ferreira, and Paolo Surico. 2015. Monetary Policy when Households have Debt: New Evidence on the Transmission Mechanism. CEPR Discussion Paper 11023.

Cloyne, James, and Patrick Hürtgen. 2016. The Macroeconomic Effects of Monetary Policy: A New Measure for the United Kingdom. *American Economic Journal: Macroeconomics* 8(4): 75–102.

Coibion, Olivier, Yuriy Gorodnichenko, Lorenz Kueng, and John Silvia. 2017. Innocent Bystanders? Monetary Policy and Inequality in the U.S. *Journal of Monetary Economics* 88: 70–88.

Conley, Timothy G., Christian B. Hansen, and Peter E. Rossi. 2012. Plausibly Exogenous. *Review of Economics and Statistics* 94(1): 260–72.

Cox, David R. 1958. *Planning of Experiments*. New York: John Wiley.

di Giovanni, Julian, Justin McCrary, and Till von Wachter. 2009. Following Germany's Lead: Using International Monetary Linkages to Estimate the Effect of Monetary Policy on the Economy. *Review of Economics and Statistics* 91(2): 315–31.

di Giovanni, Julian, and Shambaugh, Jay C. 2008. The Impact of Foreign Interest Rates on the Economy: The Role of the Exchange Rate Regime. *Journal of International Economics* 74(2): 341–61.

Eichenbaum, Martin S. 1992. Comments 'Interpreting the Macroeconomic Time Series Facts: The Effects of Monetary Policy' by Christopher Sims. *European Economic Review* 36(5): 1001–11.

Faust, Jon, Eric T. Swanson, and Jonathan H. Wright. 2004. Identifying VARs Based on High Frequency Futures Data. *Journal of Monetary Economics* 51(6): 1107–31.

Gertler, Mark, and Peter Karadi. 2015. Monetary Policy Surprises, Credit Costs, and Economic Activity. *American Economic Journal: Macroeconomics* 7(1): 44–76.

Gürkaynak, Refet S., Brian Sack, and Eric Swanson. 2005. The Sensitivity of Long-Term Interest Rates to Economic News: Evidence and Implications for Macroeconomic Models. *American Economic Review* 95(1): 425–36.

Hanson, Michael S. 2004. The "Price Puzzle" Reconsidered. *Journal of Monetary Economics* 51(7): 1385–1413.

Hanson, Samuel G., David O. Lucca and Jonathan H. Wright. 2017. Interest Rate Conundrums in the Twenty-First Century. FRBNY Staff Report no. 810.

Iacoviello, Matteo. 2005. House Prices, Borrowing Constraints, and Monetary Policy in the Business Cycle. *American Economic Review* 95(3): 739–764.

Imbens, Guido W. 2014. Instrumental Variables: An Econometrician's Perspective. NBER Working Paper 19983.

Imbens, Guido W., and Joshua Angrist. 1994. Identification and Estimation of Local Average Treatment Effects. *Econometrica* 61(2): 467–76.

Ilzetzki, Ethan, Enrique G. Mendoza, and Carlos A. Végh. 2013. How Big (Small?) are Fiscal Multipliers? *Journal of Monetary Economics* 60(2): 239–54.

Jordà, Òscar. 2005. Estimation and Inference of Impulse Responses by Local Projections. *American Economic Review* 95(1): 161–82.

Jordà, Òscar, Moritz Schularick, and Alan M. Taylor. 2015. Betting the House. *Journal of International Economics* 96(S1): S2–S18.

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 2016*, no. 31, edited by Martin Eichenbaum and Jonathan A. Parker. Chicago: University of Chicago Press. Forthcoming.

Jordà, Òscar, and Alan M. Taylor. 2016. The Time for Austerity: Estimating the Average Treatment Effect of Fiscal Policy. *Economic Journal* 126(590): 219–55.

Kaplan, Greg, Benjamin Moll and Giovanni L. Violante. 2016. Monetary Policy According to HANK. NBER Working Paper 21897

Kashyap, Anil K., Owen A. Lamont, and Jeremy C. Stein. 1994. Credit Conditions and the Cyclical Behavior of Inventories. *Quarterly Journal of Economics* 109(3): 565–593.

Kashyap, Anil K., and Jeremy C. Stein. 1995. The Impact of Monetary Policy on Bank Balance Sheets. *Carnegie-Rochester Conference Series on Public Policy* 42: 151–19.

Knoll, Katharina, Moritz Schularick, and Thomas Steger. No Price Like Home: Global House Prices, 1870–2012. *American Economic Review* 107(2): 331–53.

Kuttner, Kenneth N. 2001. Monetary Policy Surprises and Interest Rates: Evidence from the Fed Funds Futures Market. *Journal of Monetary Economics* 47(3): 523–44.

Nakamura, Emi, and Jón Steinsson. 2013. High Frequency Identification of Monetary Non-Neutrality. *Quarterly Journal of Economics*. Forthcoming.

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

Obstfeld, Maurice, Jay C. Shambaugh, and Alan M. Taylor. 2004. Monetary Sovereignty, Exchange Rates, and Capital Controls: The Trilemma in the Interwar Period. *IMF Staff Papers* 51(S): 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): 423–38.

Ottonello, Pablo, and Thomas Winberry. 2017. Financial Heterogeneity and the Investment Channel of Monetary Policy. University of Michigan. Unpublished.

Owyang, Michael T., Valerie A. Ramey, and Sarah Zubairy. 2013. Are Government Spending Multipliers Greater during Periods of Slack? Evidence from Twentieth-Century Historical Data. *American Economic Review* 103(3): 129–34.

Quinn, Dennis P., 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*, Volume 2, edited by J. B. Taylor and H. Uhlig. Amsterdam: Elsevier, pp. 71–162.

Ramey, Valerie A., and Sarah Zubairy. 2014. Government Spending Multipliers in Good Times and in Bad: Evidence from US Historical Data. NBER Working Paper 20719.

Romer, Christina D. and David H. Romer. 1989 Does Monetary Policy Matter? A New Test in the Spirit of Friedman and Schwartz. *NBER Macroeconomics Annual* 4: 121–70.

Romer, Christina D. and David H. Romer. 2004. A New Measure of Monetary Shocks: Derivation and Implications. *American Economic Review* 94(4): 1055–84.

Rubin, Donald B. 1974. Estimating Causal Effects of Treatments in Randomized and Nonradomized Studies. *Journal of Educational Psychology* 66(5): 688–701.

Rubin, Donald B. 1978. Bayesian Inference for Causal Effects: The Role of Randomization. *Annals of Statistics* 6(1): 34–58.

Rubin, Donald B. 2005. Causal Inference Using Potential Outcomes: Design, Modeling, Decisions. *Journal of the American Statistical Association* 100(469): 322–331.

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

Sims, Christopher A. 1992. Interpreting the Macroeconomic Time Series Facts: The Effects of Monetary Policy. *European Economic Review* 36(5): 975–1000.

Swanson, Eric T., and John C. Williams. 2014. Measuring the Effect of the Zero Lower Bound on Medium- and Longer-Term Interest Rates. *American Economic Review* 104(10): 3154–85.

Tenreyro, Silvana, and Gregory Thwaites. 2016. Pushing on a String: US Monetary Policy is Less Powerful in Recessions. *American Economic Journal: Macroeconomics*. 8(4): 43–74.

Wooldridge, Jeffrey M. 2015. Control Function Methods in Applied Econometrics. *Journal of Human Resources* 50(2): 420–45.

Appendices

A. Evaluating the exclusion restriction: general case

Detailed derivations of the results in Section 5.3 are provided for a panel of data. Denote the generic dimensions of the panel for each subpopulation considered as N for the cross section and T for the time series dimension. Different subpopulations will have different dimensions but, in the interest of clarity, we refrain from using subscripts for now.

In our application, we use the auxiliary subpopulation of floats to correct the IV estimates based on the peg subpopulation for spillover bias. The consistency results are based on $N, T \to \infty$ for both of these subpopulations. Because we do not need inferential procedures (ours is a robustness check to examine the potential biases from spillover effects) we do not derive the asymptotic distribution of the bias corrected estimator. In other settings, auxiliary information may be available that would allow one to obtain point-identified bias corrections. The derivations provided here could be readily extended to derive the asymptotic distribution.

To simplify the notation, consider the non-state dependent version of expression (9), extended to include potential spillovers:

$$\begin{aligned} \delta_{i,t} &= a_i + \lambda_{i,t} b + \boldsymbol{x}_{i,t}\boldsymbol{g} + \eta_{i,t}\,, \\ y_{i,t+h} &= \alpha_i^h + \hat{\delta}_{i,t}\beta_h + \lambda_{i,t}\phi + \boldsymbol{x}_{i,t}\gamma_h + \nu_{i,t+h}\,; \qquad h = 0, 1, ..., H\,. \end{aligned} \tag{14}$$

We make standard regularity assumptions about the error processes and assume, in particular, that the following moment conditions hold $E(\lambda_{i,t}\eta_{i,t}) = E(\lambda_{i,t}\nu_{i,t+h}) = 0$. That is, $\lambda_{i,t}$ is an exogenously determined valid instrument.

Next, define the projection matrix $M_{i,t} = 1 - \boldsymbol{w}_{i,t}(\boldsymbol{w}_{i,t}'\boldsymbol{w}_{i,t})^{-1}\boldsymbol{w}_{i,t}'$, where $\boldsymbol{w}$ includes the fixed effects and the $\boldsymbol{x}$, but excludes λ and δ. Pre-multiplying expression (14) by $M_{i,t}$ allows us to focus on the coefficients of interest, b, β^h, and ϕ. For example, using this projection matrix, note that the OLS estimator of b is simply

$$\hat{b} = \left(\frac{1}{NT}\sum_{t=1}^{T}\sum_{i=1}^{N}\lambda_{i,t}M_{i,t}\lambda_{i,t}\right)^{-1}\left(\frac{1}{NT}\sum_{t=1}^{T}\sum_{i=1}^{N}\lambda_{i,t}M_{i,t}\delta_{i,t}\right).$$

An additional piece of notation will help make the derivations that follow more transparent. Consider the moment condition $E(\lambda_{i,t}M_{i,t}\eta_{i,t}) = 0$. We define the operator

$$S(\lambda, \eta) = \left(\frac{1}{NT}\sum_{t=1}^{T}\sum_{i=1}^{N}\lambda_{i,t}M_{i,t}\eta_{i,t}\right) \xrightarrow{p} E(\lambda_{i,t}M_{i,t}\eta_{i,t})$$

as a way to refer to the equivalent sample moment condition implied by the population statement. Using this operator and the usual OLS moment conditions, it is easy to see that

$$\hat{\beta}_h = \frac{S(\lambda,\delta)\,S(\lambda,y(h)) - S(\lambda^2)\,S(\delta,y(h))}{S(\lambda,\delta)^2 - S(\lambda^2)\,S(\delta^2)}\,, \tag{15}$$

$$\hat{\phi} = \frac{S(\lambda,\delta)\,S(\delta,y(h)) - S(\delta^2)\,S(\lambda,y(h))}{S(\lambda,\delta)^2 - S(\lambda^2)\,S(\delta^2)}\,. \tag{16}$$

These moment conditions rely on the usual OLS assumptions $E(\delta_{i,t} M_{i,t} \nu_{i,t+h}) = 0$, and $E(\lambda_{i,t} M_{i,t} \nu_{i,t+h}) = 0$. However, here we explore what happens to estimators that assume these conditions but where in reality the first moment condition is violated due to the endogeneity of $\delta_{i,t}$. In that case, OLS estimates of $\hat{\beta}_h$ and $\hat{\phi}$ based on (15) and (16) will be biased.

In particular, we can define the bias term

$$\hat{\theta} = \frac{S(\lambda^2)\, S(\delta, \nu(h))}{S(\lambda, \delta)^2 - S(\lambda^2)\, S(\delta^2)} \xrightarrow{p} \theta\,.$$

It is then straightforward to show using expressions (15) and (16) that

$$\begin{aligned} \hat{\beta}_h &\xrightarrow{p} \beta_h - \theta\,, \\ \hat{\phi} &\xrightarrow{p} \phi + b\theta\,, \end{aligned}$$

where we made use of the standard OLS result

$$\hat{b} = \frac{S(\lambda, \delta)}{S(\lambda^2)} \xrightarrow{p} b\,. \tag{17}$$

Notice that under the assumptions of the model, this estimator is consistent for b. Next, without loss of generality, let $\beta_h = \aleph\phi$, for $\aleph$ an unrestricted parameter. Then, it is easy to show that

$$\hat{\phi}(\aleph) = \frac{\hat{b}\hat{\beta}_h + \hat{\phi}}{1 + \aleph\hat{b}} \xrightarrow{p} \phi(\aleph)\,. \tag{18}$$

In practice the true value of $\aleph$ is unknown. The approach that we follow here is to calculate expression (18) for a range of values of $\aleph \in [\underline{\aleph}, \overline{\aleph}]$. Because $\aleph$ is unknown, we do not attempt to characterize the asymptotic distribution of our spillover corrected estimator.

The next step consists of incorporating $\hat{\phi}(\aleph)$, which is based on the float subpopulation, to correct the IV estimator for the subpopulation of pegs. The control function approach consists in using the usual IV estimator on the following auxiliary expression:

$$\begin{aligned} y^*_{i,t+h} &= \delta_{i,t}\beta_h + \boldsymbol{x}_{i,t}\gamma_h + u_{i,t+h}\,, \\ y^*_{i,t+h} &\equiv (y_{i,t+h} - \lambda_{i,t}\hat{\phi}(\aleph))\,, \\ u_{i,t+h} &\equiv \nu_{i,t+h} + \lambda_{i,t}(\hat{\phi}(\aleph)) - \phi(\aleph)))\,. \end{aligned}$$

Consistency of this modified IV estimator only requires the usual consistency conditions of IV estimators applied to the sample of pegs. Notice that if $\hat{\phi}(\aleph) \rightarrow \phi(\aleph)$, then

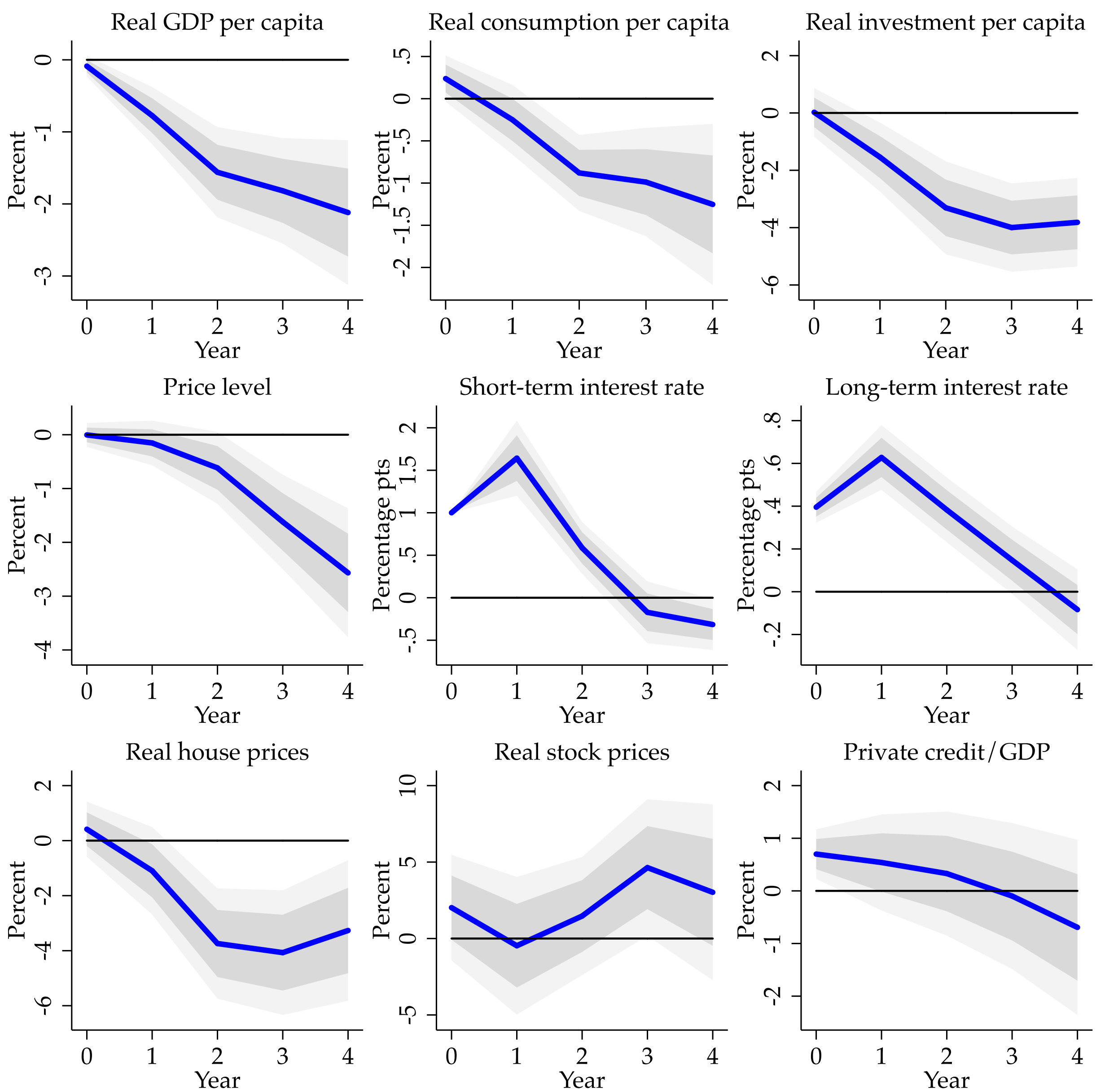
$$\left(\frac{1}{NT} \sum_{t=1}^{T} \sum_{i=1}^{N} \lambda_{i,t} M_{i,t} \lambda_{i,t} \right) (\hat{\phi}(\aleph)) - \phi(\aleph)) \xrightarrow{p} 0 \quad \text{if} \quad E(\lambda\, M\, \lambda) < \infty\,.$$

This is the condition that, along with the typical IV conditions, ensures consistency for $\hat{\beta}_h$ as $N, T \rightarrow \infty$ in both the peg and the float subpopulations.

B. Robustness checks

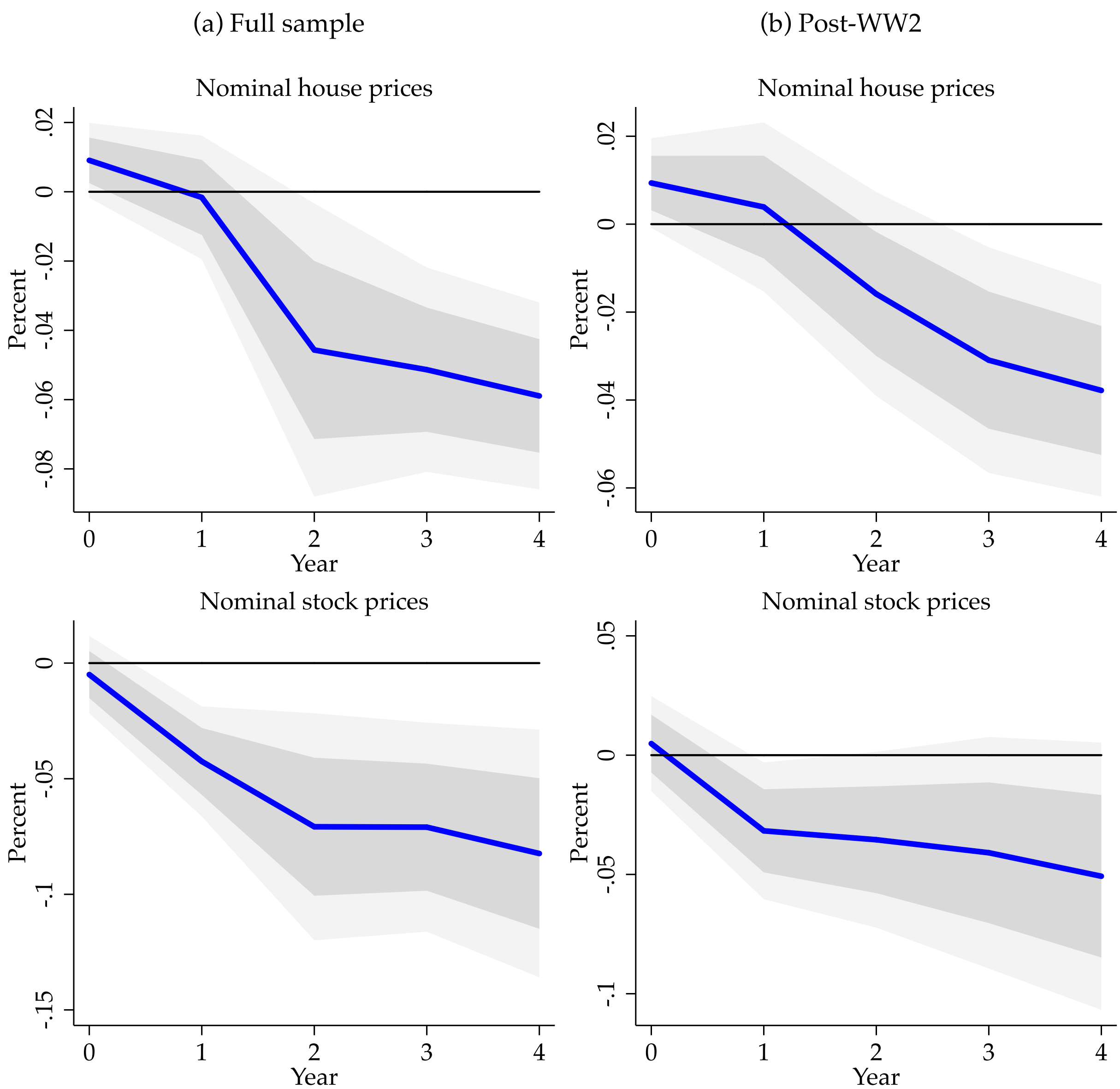
Figures B.1 through B.4 in this appendix contain a series of robustness check which are mentioned in the main text.

Figure B.1: *Full baseline results. Post-WW2 sample*



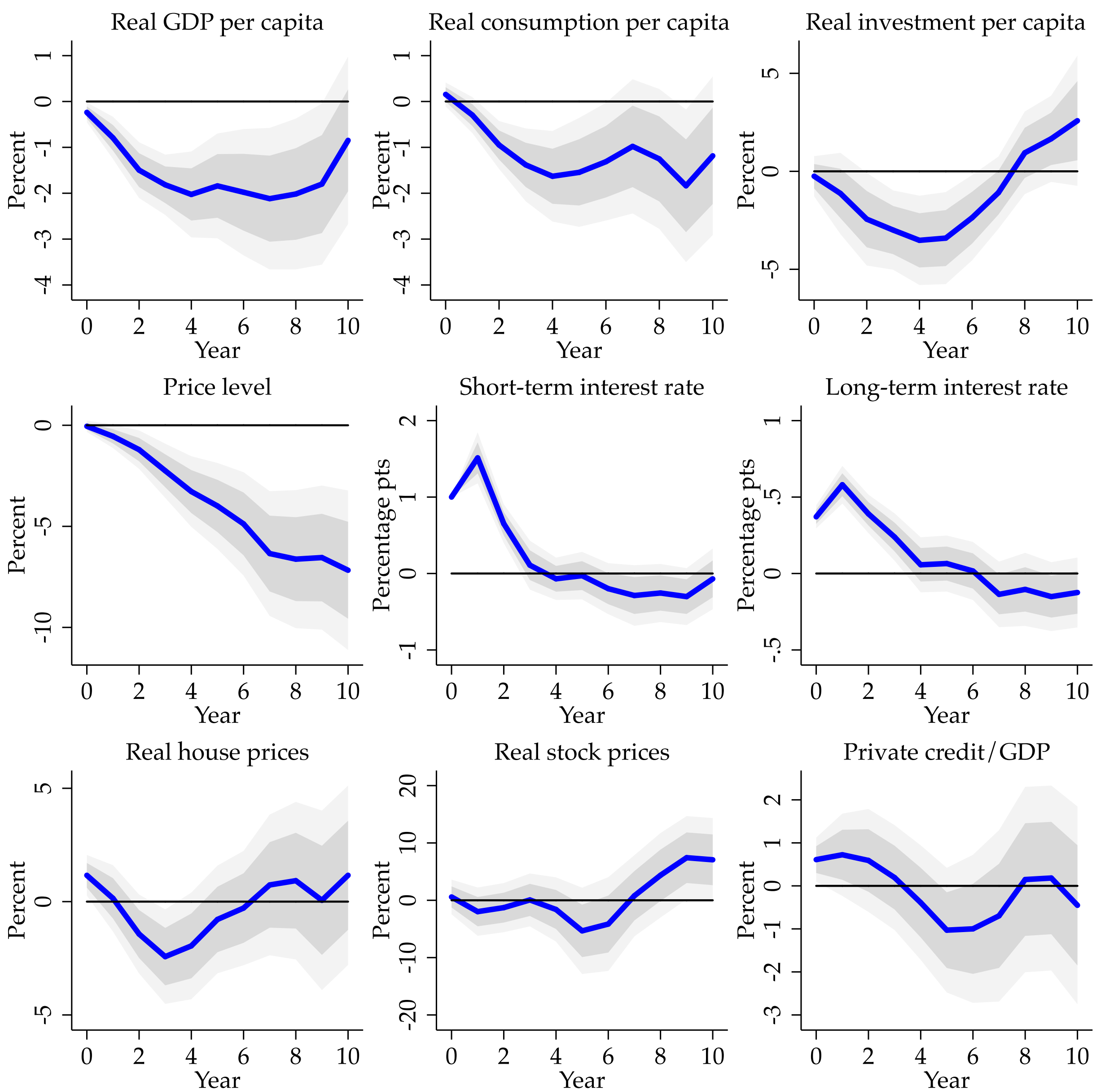
Notes: Post-WW2 sample: 1948–2013. LP-IV estimates displayed with a solid blue line and 1 S.D. and 90% confidence bands. See text.

Figure B.2: *Nominal asset price responses to interest rates. Full and post-WW2 samples*



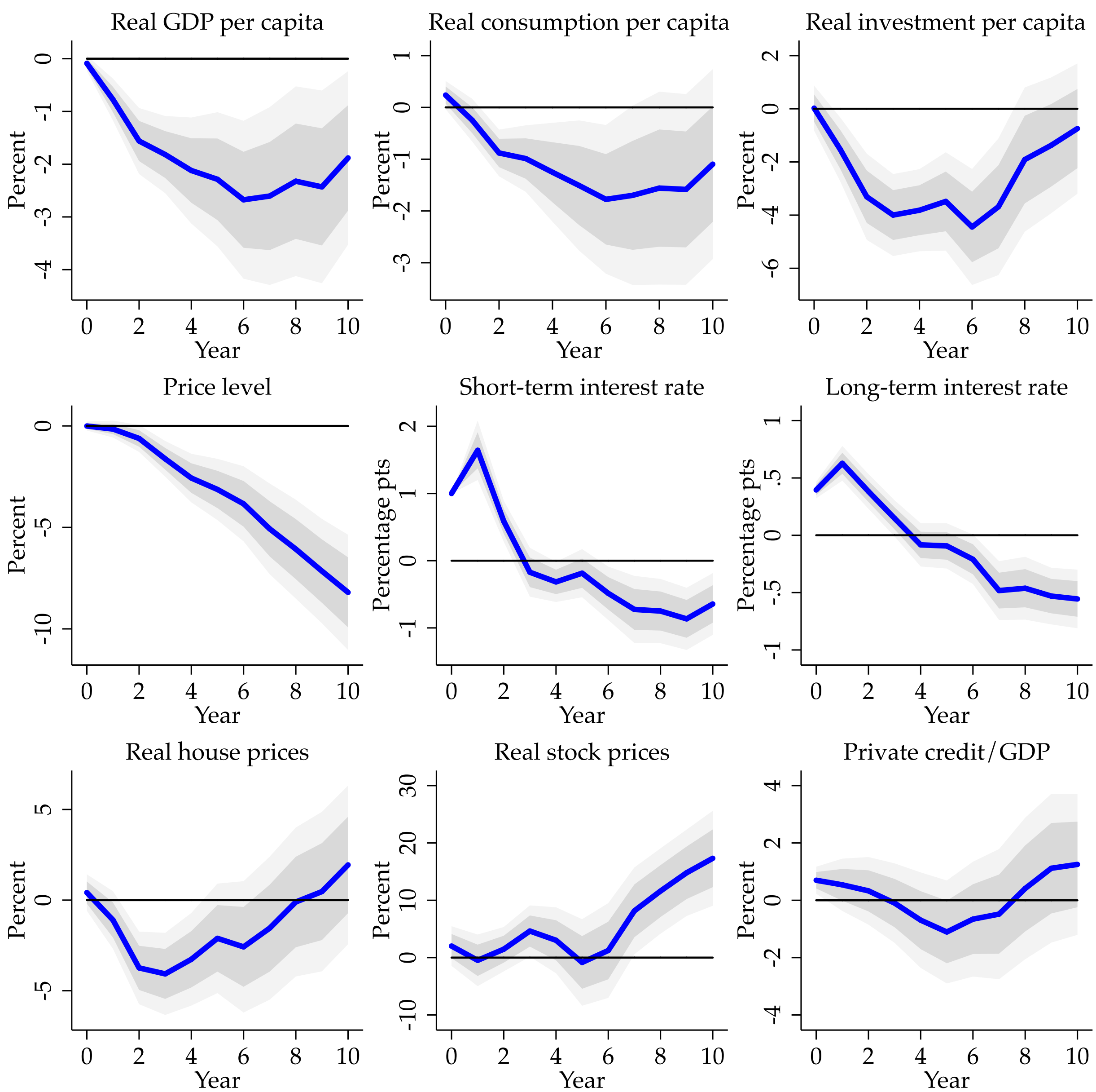
Notes: Full sample: 1870–2013. Post-WW2 sample: 1948–2013. LP-IV estimates displayed with a solid blue line and 1 S.D. and 90% confidence bands. See text.

Figure B.3: *Full baseline results. Full sample. 10-year horizon*



Notes: Full sample: 1870–2013. LP-IV estimates displayed with a solid blue line and 1 S.D. and 90% confidence bands. See text.

Figure B.4: *Full baseline results. Post-WW2 sample. 10-year horizon*



Notes: Post-WW2 sample: 1948–2013. LP-IV estimates displayed with a solid blue line and 1 S.D. and 90% confidence bands. See text.