Large and state-dependent effects of quasi-random monetary experiments[⋆]

Òscar Jordà[†] Moritz Schularick[‡] Alan M. Taylor[§]

January 2017

Abstract

Fixing the exchange rate constrains monetary policy. Along with unfettered cross-border capital flows, the trilemma implies that arbitrage, not the central bank, determines how interest rates fluctuate. The annals of international finance thus provide quasi-natural experiments with which to measure how macroeconomic outcomes respond to policy rates. Based on historical data since 1870, we estimate the local average treatment effect (L*ATE*) of monetary policy interventions and examine its implications for the population *ATE* with a trilemma instrument. Using a novel control function approach we evaluate the robustness of our findings to possible spillovers via alternative channels. Our results prove to be robust. We find that the effects of monetary policy are much larger than previously estimated, and that these effects are state-dependent.

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

Keywords: interest rates, monetary experiments, trilemma, fixed exchange rates, instrumental variables, local projections, local average treatment effect

[⋆]We thank James Cloyne and Patrick Hürtgen for sharing their data with us. Comments and suggestions from James Cloyne, Julian di Giovanni, Gernot Müller, Ricardo Reis, and Jón Steinsson have helped improve the paper. Seminar and conference participants at the Federal Reserve Board of Governors, 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, and the NBER Monetary Economics program meeting provided useful feedback. We are grateful to Helen Irvin for outstanding research assistance. Generous grants from the Institute for New Economic Thinking 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 or the Federal Reserve System.

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

[‡]Department of Economics, University of Bonn; and CEPR (moritz.schularick@uni-bonn.de).

[§]Department of Economics and Graduate School of Management, University of California, Davis; NBER; and CEPR (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, with a few notable exceptions, money is neutral in the long-run; but the short-run effects can vary widely and depend critically on various market frictions. As a result, determining whether monetary policy is a useful stabilizer is largely 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 policy effect is contaminated by confounders simultaneously correlated with the assignment mechanism and the outcome.

Over time, a great deal of consensus has emerged almost exclusively built on the Post-WW2 U.S. experience. This consensus grew from a variety of empirical routes that share a common thread: the desire to control for the information used by the central bank in setting interest rates. Whether through a saturated model of observables (see, e.g., Christiano, Eichenbaum, and Evans 1999, and the long tradition of structural VARs); or through indirect measures based on financial data (e.g., Kuttner 2001; Faust, Wright, and Swanson 2004; Nakamura and Steinsson 2013; Gertler and Karadi 2015) the objective is almost always the same: to achieve identification via selection-on-observables arguments. This paper tries a different approach.

Randomized control trials in macroeconomics are, understandably, very rare indeed. However, when countries peg their exchange rate to a *base* country, but allow capital to flow freely across borders, simple arbitrage forces local and international interest rates to cohere. The subsequent loss of monetary control in open pegs, an implication of the trilemma of international finance, is as close to a natural experiment as one can find in macroeconomics.

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 (Obstfeld, Shambaugh, and Taylor 2004; 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.[1]

In this paper we take the idea further: first, we use an instrumental variables local projections (LP-IV) approach which—unlike a VAR—is flexible, parsimonious, and adaptable to nonlinearities (which we find to be important);[2] second, we develop econometric methods to formalize the link

[1] 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.

[2] See Owyang, Ramey, and Zubairy (2013); Jordà, Schularick, and Taylor (2015).

between the average treatment effect of monetary policy in all economies (ATE) and the local average treatment effect (LATE) that can be estimated from open pegs using the trilemma instrument; third, we introduce a novel approach using control functions to place bounds on the bias that may arise if there are confounding spillovers from base interest rates to the local economy via other channels (i.e., violations of the exclusion restriction); and fourth, we expand the sample from post-1970s to all of advanced economy macroeconomic history since 1870.

An important contribution of our paper is to document the attenuation bias generated from the failure of the selection-on-observables assumption implicit in much of the literature. However, the results that we obtain from the direct IV estimator are an estimate of the local average treatment effect (L*ATE*) for the subpopulation of open pegs. What about the other countries?

Comparing the sample of open *pegs* with other subpopulations (e.g. economies that allow the exchange rate to *float*), we find similar outcome responses to interest rate interventions estimated by OLS. If the attenuation bias present in open pegs mirrors that in other subpopulations then, intuitively, the ATE for the full population is likely to be similar to the LATE that we report. Moreover, using the instrument in Romer and Romer (2004) and Cloyne and Hürtgen (2014), we find that the LATE for the U.S. and the U.K., both of which historically have usually been base countries, is similar to the LATE for the open pegs. This finding suggests that the initial LATE estimates are not driven by factors that also determine the decision to peg. For these reasons, we feel reasonably confident that our estimates provide a good approximation to the true, but unmeasurable ATE of interest rates on macroeconomic outcomes.[3]

In principle, unpredictable movements in foreign interest rates could affect pegging economies through channels other than the interest rate. Such spillover effects would be a violation of the exclusion restriction and therefore a potential source of bias. Drawing on the microeconometrics literature on control functions (see, e.g., Wooldridge 2015) we introduce a couple of interesting and useful results. First, although it might appear obvious which way this bias goes, careful calculation shows that its direction is ambiguous. Second, even so, guided by economic reasoning we can place plausible bounds on this bias. These bounds discipline the range of ATE responses that one can reasonably infer from the LATE estimates, and, in fact, the deviations appear to be small compared to our baseline LP-IV results.

Another innovation of our paper is to estimate the state-dependent effect of interest rates on macroeconomic outcomes. The states of the economy that we focus on here refer to the levels of the output gap ("boom/slump") and inflation ("high/lowflation"). Recent studies on the fiscal multiplier (e.g., Auerbach and Gorodnichenko 2013ab; Ramey and Zubairy 2014; and Jordà and

[3]As we will discuss below, we find in auxiliary regressions for our sample (not shown) that the decision to peg is largely unpredictable, and it is also a very persistent choice (in our sample, the average peg is 21 years). In addition, we guard against the risk that very short opportunistic pegs could bias our results by requiring a country to be pegged for two successive years when we construct our trilemma instrument. In other words, we think any endogeneity embedded in the decision to peg or leave a peg appears to be well contained at business cycle frequencies in our application. In addition, the monotonicity assumption requiring instrument and treatment to be positively associated appears to hold in the data.

Taylor 2016) 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 results reported in Angrist, Jordà, and Kuersteiner (forthcoming) and Tenreyro and Thwaites (2016). These papers show that interest rates have asymmetric effects, something we find as well. However, in our context, the source of the asymmetry comes from the difference in the boom-versus-slump state of the economy when policy actions are taken. As for inflation, advanced economies have recently been grappling with a low growth, low inflation environment dubbed by commentators as "lowflation." A burning question in central bank circles is about the effectiveness of monetary policy against this background. To answer such a question, we again investigate the effects of monetary perturbations in lowflation environments. The historical data show that policy is rather ineffective in such cases, revealing another dimension of state-dependent monetary policy. Thus our results have important—albeit pessimistic—implications for the ongoing policy debate and, more generally, for appropriate theoretical models of monetary economies.

2. The trilemma of international macroeconomics: 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 we discuss our overall empirical approach. The specific construction of this instrumental variable is one of the important contributions of our paper.

We suppose that any home country i at time t can either have a flexible (floating) or fixed (pegged) exchange rate with respect to some other *base* country. Let Δr_{it} denote the change in short-term nominal interest rates in the home country, 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 the international historical dataset that we use (discussed in Section 3), these interest rates refer to 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.[4]

Denote $\Delta \hat{r}^*{}_{b(i,t),t}$ movements in base country $b(i,t)$ rates explained by observable controls for that base country, $\boldsymbol{x}_{b(i,t),t}$. Thus, $\Delta r^*_{b(i,t),t} - \Delta \hat{r}^*{}_{b(i,t),t}$ denotes unpredictable movements in base country interest rates. The idea is to take, not just the base country rate as exogenous with respect to the home country, but to control for any predictable movements in base interest rates that might have been expected by country i. Robustness checks based on directly using $\Delta r^*_{b(i,t),t}$ produced similar

[4]See, e.g., Swanson and Williams (2014) for an argument favoring the use of medium-term government rates to measure monetary policy effects.

results to those reported below and are available upon request. Similarly, but for use later, we denote with $\boldsymbol{x}_{it}$ the set of domestic macroeconomic controls used in determining the home interest rate in the absence of a peg.

Going forward, we will be careful to separate country-year pairs into the corresponding subpopulations of *pegs*, *floats*, and *bases*. For now, the variable $PEG_{it} = 0, 1$ simply denotes if country i is a *float* or a *peg* respectively. 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 correlated. 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. Nevertheless, in defining our instrument, we require countries to have been in a peg for at least one year by interacting the variable PEG_{it} with $PEG_{i,t-1}$ as in expression (1) below. This feature eliminates potential distortions of a country that just joined a peg.

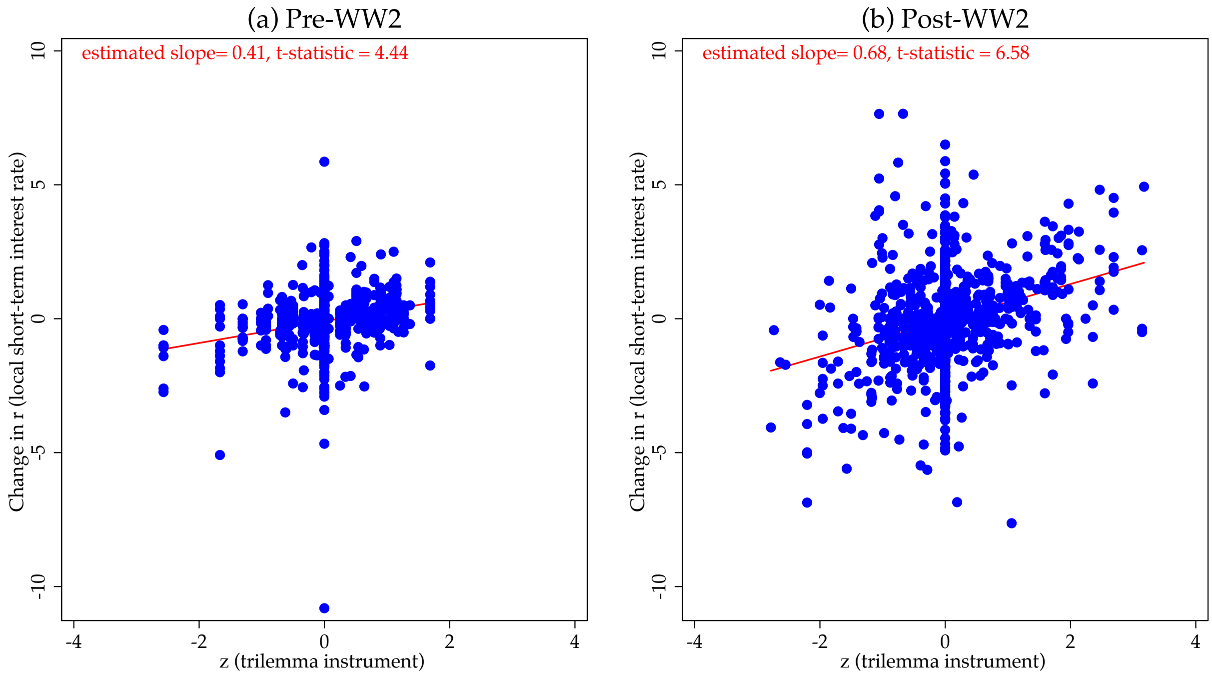
Finally, the variable $KOPEN_{it} \in [0, 1]$ denotes whether the home country is open to international markets or not. This capital mobility indicator is based on the index (from 0 to 100) in Quinn, Schindler, and Toyoda (2011). We use a continuous version of their index rescaled to the unit interval, 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.

Consequently, define the *trilemma instrument* as

$$z_{it} \equiv (\Delta r^*_{b(i,t),t} - \Delta \hat{r}^*{}_{b(i,t),t}) \times PEG_{it} \times PEG_{i,t-1} \times KOPEN_{it}\,. \tag{1}$$

To get a sense of the relationship between the endogenous variable and the proposed instrument, Figure 1 shows scatter plots and fitted values from a regression of the change in local short-term rates, Δr_{it}, on the trilemma instrument z_{it} defined in expression (1), for the pre-WW2 and post-WW2 sample periods. These two regressions correspond to the results reported in columns 2 and 3 of Table 1. To more fully evaluate the strength of the instrument, Table 1 reports first-stage regression results of the endogenous variable, Δr_{it} on the instrument z_{it}, without controls in columns 1–3, in keeping with Figure 1, and then more formally with controls, in columns 4–6.

Figure 1: *Relationship between change in short-rates in the home country and the trilemma instrument*



Notes: Pre-WW2 sample: 1870–1938, excluding WW1, 1914–1919. Post-WW2 sample: 1948–2013.

Table 1: *Relationship between change in short-rates in the home country 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
Constant	-0.08***	-0.10***	-0.06***	-0.32***	-0.34***	-0.35
	(0.01)	(0.02)	(0.01)	(0.10)	(0.08)	(0.21)
z_{it}	0.58***	0.41***	0.68***	0.43***	0.39***	0.43***
	(0.09)	(0.09)	(0.10)	(0.07)	(0.12)	(0.09)
t-statistic	[6.71]	[4.44]	[6.58]	[5.78]	[3.29]	[4.68]
Observations	1961	870	1091	1229	312	917

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.

Table 1 reaffirms the evidence displayed in Figure 1 formally, and clearly shows that z_{it} 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 z_{it} in all cases is well above 3. 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] We will pause further discussion of our instrument to introduce the statistical framework used for causal identification. First, 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 is publicly available at http://www.macrohistory.net/data/. This website includes a 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, UK, and USA. Unlike the vast majority 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 (forthcoming), the first major attempt

[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.

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

	Home country interest rates linked with the base			
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.

to 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. Long-term interest rates are 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, such as 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, which we will denote as z_{it}, 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. Statistical discussion

Our empirical approach integrates elements of the potential outcomes paradigm of the Rubin Causal Model (Rubin 1974) into vector time series analysis using instrumental variable, local projection methods (LP-IV). The analysis relies on a series of assumptions worth spelling informally first.

Random assignment of the instrument reflects the idea that interest rate setting in base economies is independent of economic conditions in peg economies. Relative size disparities between base and peg economies are usually quite large. Moreover, central bank minutes seldom reference conditions in other economies in arguments about rate setting. Thus, to a first approximation, potential outcomes in pegs do not influence base country policy choices. The assumption of stable unit treatment value (SUTVA as it is know in the potential outcomes literature and explained in the appendix) is therefore reasonable to make.[6] The positive association between our instrument and domestic rates discussed in Section 2 fulfills the monotonicity assumption discussed in Imbens and Angrist (1994). That is, in our application interest rate increases in base economies are associated, on average, with rate increases in the associated peg economies.

Based on these assumptions, a roadmap of the analysis is as follows. First, the LP-IV setup lets us estimate the Local Average Treatment Effect (LATE) of interest rates on macroeconomic outcomes for the subpopulation of open pegs. Compared to these LP-IV estimates, typical OLS estimates based on regression control (such as VARs) display considerable attenuation bias. Intuitively, if the central bank lowers interest rates to stimulate the economy, the correlation between low interest rates and low output will tend to attenuate the true effect of the policy intervention—that low interest rates stimulate economic activity. Regression control only goes so far in overcoming this bias.

Second, 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 (2014), we also compute the LATE

[6] We investigated how predictable is the decision to peg using a logit model and controls. We found it difficult to predict when countries decide to enter or leave such arrangements. The results are not reported here but are available upon request.

for the U.S. and the U.K. in the post-WW2 period. This is the best estimate of the LATE for the subpopulation of bases. Despite using an instrument motivated quite differently and that is available for two base economies, the LATE estimates are encouragingly similar to the estimates of the LATE for the subpopulation of pegs.

Third, LP-OLS estimates across the three subpopulations considered (pegs, floats and bases) are very similar to one another. LP-IV estimates for pegs and for the U.S. and the U.K. are also very similar to one another. Similarities of the estimated LATE for pegs, U.S. and U.K., and similarities of the attenuation bias for all three subpopulations lead us to speculate that our LATE estimates are as good an approximation of the Average Treatment Effect (ATE) as the existing data from the universe of macroeconomic data in advanced economies will currently allow.

4.1. Preliminaries and notation

Let $\boldsymbol{w} = (w_1, ..., w_k, ..., w_{K-1}, r)$ be a $K_w \times 1$ vector of random variables containing observations on $K - 1$ macroeconomic outcomes (suitably transformed) plus a final additional variable r, a short-term interest rate variable. We use boldface to denote random vectors, and normal type for random variables. The available sample $\{\boldsymbol{w}_{i,t}\}_{i=1,t=1}^{N,T}$ is observed for $i = 1, ..., N$ with $N = 17$, the sample of countries in our sample. The data are observed annually indexed $t = 1, .., T$ starting in 1870 and ending in 2013.

Define the random variable $y(h) \equiv w_k(h) - w_k(-1)$ for any $k \in \{1, ..., K_w\}$, that is, the difference in the macroeconomic outcome variable w_k (that is, we also include r as a possible outcome) h periods in the future relative to last period. Note that in the sample, an observation of $y(h)$ would simply be $y_{i,t+h} = w^k_{i,t+h} - w^k_{i,t-1}$. Going forward we omit the index k for $y(h)$ for simplicity. On occasion, we expand this notation to $y(r, h)$ to refer to the potential outcomes one would observe, h periods from today, for different values of the treatment (policy) variable r.

The random vector $\boldsymbol{y} = (y(0), ..., y(h), ..., y(H - 1))$ of dimension $1 \times H$ collects the response to r over time. Similar to the notation $y(r, h)$, $\boldsymbol{y}(r)$ will refer to the potential response paths one could observe for a given value of the treatment (policy) variable r. Next, define the vector of observable controls $\boldsymbol{x} = (w_1, ..., w_{k-1}, w_{k+1}, ..., w_{K-1}, r, \boldsymbol{w}(-1)\ , ..., \boldsymbol{w}(-p))$. We assume a sample $\{\boldsymbol{x}_{i,t}\}_{i=1,t=1}^{N,T}$ is available. Note that $\boldsymbol{x}$ is of dimension $1 \times K_x$, with $K_x = K_w(p + 1) - 1$. Importantly, this vector includes contemporaneous observations for r and all the macro-outcomes (except the k^{th}, which appears as a left-hand side variable), as well as up to p lags of all the variables. When investigating the responses to r, including contemporaneous values of all other controls is equivalent to specifying the policy equation last in a VAR based on a Choleski causal order. Our intent here is to ensure that fluctuations in r are orthogonal to any information that was plausibly available to policymakers.

We denote the impulse response $I(\boldsymbol{y}, r, \delta, \boldsymbol{x})$ as the following counterfactual experiment:

$$I(\boldsymbol{y}, r, \delta, \boldsymbol{x}) \equiv \mathbb{E}(\boldsymbol{y}(r + \delta) \mid \boldsymbol{x}) - \mathbb{E}(\boldsymbol{y}(r) \mid \boldsymbol{x}) = \mathbb{E}(\boldsymbol{y}(r + \delta) - \boldsymbol{y}(r) \mid \boldsymbol{x})\,. \tag{2}$$

The notation $\mathbb{E}(.)$ refers to the ideal population expectation over the distribution of potential outcomes r. Note that the random variables, $\boldsymbol{y}(r+\delta)$ and $\boldsymbol{y}(r)$, cannot be simultaneously observed. However, their expectation can be estimated by imposing further assumptions.

A natural starting point would to assume that the conditional mean is linear, so that

$$E[\boldsymbol{y}(r) \mid \boldsymbol{x}] = \boldsymbol{A} + r\boldsymbol{\gamma} + \boldsymbol{x}\boldsymbol{B}\,, \tag{3}$$

where we replace the notation $\mathbb{E}$ to E to denote that we pursue an approximation to the ideal, and where $\boldsymbol{A}$ is a $N \times H$ matrix that collects fixed-effects coefficients for all N countries across all H horizons with $\boldsymbol{A} = (\boldsymbol{\alpha}^0, ..., \boldsymbol{\alpha}^{H-1})$. The effect of the treatment on the outcome is measured by $\boldsymbol{\gamma}$, a $1 \times H$ vector of average experimental responses, with $\boldsymbol{\gamma} = (\gamma_0, ..., \gamma_{H-1})$. Finally, $\boldsymbol{B}$ is a $K_x \times H$ matrix of coefficients such that $\boldsymbol{B} = (\boldsymbol{\beta}_0, ..., \boldsymbol{\beta}_{H-1})$ with $\boldsymbol{\beta}_h$ a $K_x \times 1$ vector for $h = 0, 1, ..., H-1$. Identification relies on $\boldsymbol{x}$ containing all relevant information.[7] When omitted and potentially unobservable factors bias the conditional mean estimate in expression (3), an instrumental variable approach is to be preferred, if available.

Suppose that an instrument, z is available, as in our application, although in the more general case it is natural to think of a vector of instrumental variables $z_1, ..., z_m$. We now define $\boldsymbol{x}^*$ to be the same as $\boldsymbol{x}$ but with the element r omitted, that is, $\boldsymbol{x}^* \equiv (w_1, ..., w_{k-1}, w_{k+1}, ..., w_{K-1}, \boldsymbol{w}(-1)\ , ..., \boldsymbol{w}(-p))$. This is appropriate, since r is the treatment or policy variable for which we have an instrument. We can then define $\boldsymbol{z} = (z_1, ..., z_m, \boldsymbol{x}^*)$.

4.2. Instrumental variables and local average treatment effects (LATE)

Using these definitions we now discuss a set of assumptions that can be thought of as the time series version of some of the conditions in Angrist, Imbens, and Rubin (1996), but for continuous data as in Imbens (2014). Note that the specific timing conventions are motivated by the structure of our data. Moreover, because later on we will analyze state-dependent policy effects, we are mindful to specify these assumptions conditional on the state s used to stratify the analysis. For now the state is left unspecified. In what follows we stick with the single instrument notation (i.e., $m = 1$) to facilitate the exposition without loss of generality. Note that by definition, for country i, $\boldsymbol{x}^*_{it}$ refers to controls observed at time t and its lags.

Assumption 1 (Unconfoundedness or random assignment of the instrument)**.**

$$z \perp \boldsymbol{y}(r), r(z) \mid \boldsymbol{x}^*, s. \tag{4}$$

This assumption is basically the instrument validity assumption of traditional econometric studies. We require the instrument to be independent of potential outcomes and of potential policy choices given controls. Naturally, one expects the instrumental variable to have predictive power for

[7] The appendix carefully explains the selection-on-observables assumption needed for identification.

the policy variable r. However, the assumption states that fluctuations in the instrument are not determined by the potential policy choices that are likely to be made and denoted $r(z)$.

This is one of the features of the potential outcomes notation that is perhaps less common in traditional time series studies. Similarly, the independence of z is with respect to the potential macroeconomic outcomes one could observe given the policy choice $r(z)$, which in turn potentially depends on the value of z. Again, the assumption does not mean that z and $\boldsymbol{y}$ are independent of one another. As explained in Imbens (2014), there are several ways in which this assumption can be stated. For our purposes and under the assumption of linearity, we could relax this assumption and invoke instead conditional mean assumptions. However, we think it is preferable to maintain the more general statement of the assumption for now. Next, we state the exclusion restriction.

Assumption 2 (Exclusion restriction)**.**

$$\boldsymbol{y}(r, z, \boldsymbol{x}^*, s) = \boldsymbol{y}(r, z', \boldsymbol{x}^*, s) \quad \textit{for all } z, z'; \textit{ given } \ \boldsymbol{x}^* \textit{ and } r, \textit{ and state } s. \tag{5}$$

This assumption indicates that any effects of the instrument on the outcome come directly through the policy variable. Since these may be different depending on the state and on the conditioning history up to the point of intervention, we are careful in specifying each element. Later on we provide bounds for our instrumental variable estimates should this assumption fail by using the auxiliary subpopulation of floats to provide estimates based on data.

The next assumption that we state is essentially the usual relevance assumption in instrumental variables analysis, augmented to account for potential differences given the state s.

Assumption 3 (Relevance)**.**

$$z \not\perp r \mid \boldsymbol{x}^*, s. \tag{6}$$

This assumption simply states that the instrument is correlated with the endogenous policy variable. The final assumption we make is that of *monotonicity*.

Assumption 4 (Monotonicity)**.**

$$\textit{If} \quad z > z' \quad \textit{then} \quad E[r \mid z, \boldsymbol{x}^*, s] \ \geq \ E[r \mid z', \boldsymbol{x}^*, s] \quad \textit{for all } \ \boldsymbol{x}^*, s. \tag{7}$$

The monotonicity assumption plays an important role in the identification of the LATE although formal tests of the assumption are usually not available. Several forms of stating this assumption have appeared in the literature, although there are few results that provide guidance in our empirical setting. The stricter forms are based on higher-order stochastic dominance arguments. Here we prefer to make a statement in terms of mean dominance since it has a more natural connection with the results presented in Section 2—when interest rates increase in the base country, interest rates are predicted to increase in the pegging country.

Linearity assumptions further facilitate estimation in practical settings. We make two further assumptions. First, we assume linearity of the conditional mean of r given $\boldsymbol{z}$, that is,

$$E[r \mid \boldsymbol{z}; s] = \boldsymbol{\mu}^s + \boldsymbol{z}\boldsymbol{\theta}^s \equiv \tilde{r}^s\,, \tag{8}$$

where $\boldsymbol{\mu}^s$ are country fixed effects, and $\boldsymbol{\theta}^s$ is an $(m + K_x - 1) \times 1$ vector of coefficients. In principle, we allow the first stage to depend on the state variable s, although in our application we assume the coefficients are constant across states. In addition to economizing on estimated regressors, the assumption is not too restrictive in our context. It simply means that there is a tight connection between the instrument and the policy variable, regardless of the state, as the trilemma would imply.

Second, we assume linearity of the following conditional mean:

$$E[\boldsymbol{y}(\tilde{r}) \mid \boldsymbol{z}; s] = \boldsymbol{A}_s^* + \tilde{r}\boldsymbol{\gamma}_s + \boldsymbol{x}^*\boldsymbol{B}_s^*\,, \tag{9}$$

where $\boldsymbol{A}_s^*$ is a $N \times H$ matrix that for a given state s, collects fixed effects coefficients for all N countries across all H horizons, that is $\boldsymbol{A}_s^* = (\boldsymbol{\alpha}_s^{0,*}, ..., \boldsymbol{\alpha}_s^{H-1,*})$; $\boldsymbol{\gamma}_s$ is the $1 \times H$ vector of average experimental responses $\boldsymbol{\gamma}_s = (\gamma_{0,s}, ..., \gamma_{H-1,s})$; and $\boldsymbol{B}_s^*$ is a $K_x - 1 \times H$ matrix of coefficients such that $\boldsymbol{B}_s^* = (\boldsymbol{\beta}_{0,s}^*, ..., \boldsymbol{\beta}_{H-1,s}^*)$, with $\boldsymbol{\beta}_{h,s}^*$ a $K_x - 1\times$ vector for $h = 0, 1, ..., H-1$.

Hence, the two-step estimator based on expressions (8) and (9) is:

$$\begin{aligned} r_{i,t} &= \mu_i^s + \boldsymbol{z}_{i,t}\boldsymbol{\theta}^s + \epsilon_{i,t} \qquad \Longrightarrow \qquad \hat{r}_{i,t}^s = \hat{\mu}_i^s + \boldsymbol{z}_{it}\hat{\boldsymbol{\theta}}^s\,, \\ y_{i,t+h} &= \alpha_{i,s}^h + \hat{r}_{i,t}^s\gamma_{h,s} + \boldsymbol{x}_{i,t}^*\boldsymbol{\beta}_{h,s}^* + \eta_{i,t+h}\,. \end{aligned} \tag{10}$$

Thus the notation allows for situations in which policy responses may be state dependent, but is otherwise standard.

Given the definition of an impulse response in expression (2) and the instrumental variables machinery now in place, an estimate of the state-dependent version of the impulse response using estimates from (10) is easily seen to be:

$$I(y, h, \delta, s) = \hat{\gamma}_{h,s}\delta\,; \qquad \text{for } \; h = 0, 1, ..., H-1\,;\, s\,. \tag{11}$$

where $I(y, h, \delta, s)$ denotes the response h periods ahead to an impulse of size δ in r, given state s. We omit r, which is the only treatment variable we are considering, to simplify the notation.

5. Subpopulation experiments, population effects

Section 2 showed that the trilemma instrument meets Assumptions 3 (relevance) and 4 (monotonicity). However, we begin by comparing estimates of monetary interventions under the assumption of conditional ignorability (the implicit assumption in the VAR literature stated in greater detail in the appendix) for the three subpopulations defined earlier: country-year pairs for economies that pegged their exchange rate, denoted as *pegs*; country-year pairs of economies that allowed their exchange rate to float, denoted as *floats*; and country-year pairs of economies that served as the base economy to which others pegged their exchange rate to, denoted as *bases*. The objective is to compare estimates of γ in expression (3) across subpopulations.

Next we move on to estimates based on the trilemma instrument for the subpopulation of pegs to calculate the LATE of monetary interventions. As a cross check, we then examine the LATE for bases post-WW2 using a different instrument than the one we propose —the Romer and Romer (2004) and Cloyne and Hürgten (2014) instrument based on the forecast errors by the staff at the Fed and the Bank of England. We will show that LATEs for these subpopulations based on very different instruments yield very similar results.

5.1. Local average treatment effects under conditional ignorability

Recall the three subpopulations defined earlier: *pegs*, *floats*, and *bases*. A country can fall into any of these three bins throughout its history. 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*. 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 treatment variable or impulse 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 control 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.[8]

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 choose to address this issue by allowing the controls to take

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

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.

We begin by estimating 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. 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 $H_0 : LATE = 0$.

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 ($h = 4$) 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 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 Table 3? There is scant 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 based on two different instruments.

5.2. LP-IV LATE estimates: two instruments for two subpopulations

We compare LP-OLS 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. Under the assumptions of Section 4, we calculate the LATE of an interest rate intervention and evaluate any attenuation bias from violations of conditional ignorability using a Hausman test. Table 4 summarizes the main results.

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)	
$H_0 : LATE = 0$	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)	
$H_0 : LATE = 0$	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. $H_0 : LATE = 0$ refers to the null that the coefficients for $h = 0, \ldots, 4$ are jointly zero for a given subpopulation. See text.

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	
$H_0 : LATE = 0$	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	
$H_0 : LATE = 0$	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. $H_0 : LATE = 0$ 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.

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 equivalent to that in column 2. Column 6 reports the Hausman test between columns 4 and 5 instead. We check whether the trilemma instrument is weak with Kleibergen-Paap tests. Finally, we test the null that all LATE coefficients are jointly zero and report the p-value of the test in the row labeled $H_0 : LATE = 0$.

The first check 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 ignorability 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.

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. 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 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 (2014). We denote the 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 LATE results for this group, where we should note that a similar IV for Germany 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. Moreover, one can think of this specification as the empirical counterpart to a three-variable

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	
$H_0 : LATE = 0$	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 (2014) IV. U.S. sample: 1969–2007. U.K. sample: 1976–2007. *KP weak IV* refers to the Kleibergen-Paap test for weak instruments. $H_0 : LATE = 0$ 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.

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. 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 estimates based on the trilemma instrument and broken down by subpopulation. We take this evidence to suggest that conditional ignorability, 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. More broadly, this result casts doubts on the impulse responses reported using standard VARs predicated on common identification assumptions not based on external instruments.

Second, the responses calculated for pegs using our trilemma instrument and those calculated for the U.S. and the U.K. using the RRCH instrument are reasonably close. We conclude from these results that the LATE estimates for pegs and those for bases are somewhat similar. Therefore going forward, we use both instruments in tandem to maximize statistical power while recognizing that the results that we report will be driven primarily by the subpopulation of pegs (given the differences in sample size). 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 provide some reassurance that interest rate responses are similar for this subpopulation as well.

Is our estimated LATE a good approximation to the unobserved 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 LATE measured for pegs to the 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, our approach is no different than that of 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. A novel LP-IV LATE spillover correction

We now show how we can relax the exclusion restriction, which is the assumption that interest rate fluctuations in bases affect outcomes in pegs via the interest-rate channel only. In the absence of other spillover mechanisms, the assumption is natural: interest rates in pegs and bases will tightly cohere when capital is allowed to flow freely, since arbitrage in open financial markets is swift and powerful. In this section we entertain the existence of other spillovers as a possibility, however remote, and develop methods to investigate potential biases using a control function approach and auxiliary information from the subpopulation of floats for external verification.

Assumption 2, the exclusion restriction, states that the instrument affects the outcome only through its correlation with the policy variable. The instrument does not directly affect the outcome otherwise. A priori economic arguments for and against this assumption might be contrived. But what would happen if the exclusion restriction fails? We cannot formally test this assumption. However, we now introduce a novel method to place bounds on the IV estimator depending on direct spillover effects of the instrument on the outcome. We show that the subpopulation of floats provides the auxiliary information that enables us to do this.

Consider a simple example where, as before, y is the outcome variable, r is the intervention, and z is the instrument. We abstract from controls and from state dependence to keep things simple. The appendix provides formal derivations. The standard IV setup is

$$\begin{aligned} y &= r\,\gamma + z\,\delta + \eta\,, \\ r &= z\,\theta + \epsilon\,. \end{aligned} \tag{12}$$

Typically we assume $E(r\,\eta) \neq 0$, but $E(z\,\eta) = 0$. Thus, the second line in expression (12) is the familiar first-stage regression in IV estimation. The exclusion restriction refers to the assumption that $\delta = 0$. If this restriction were not hold, it is easy to see that

$$\hat{\gamma}_{IV} \ \xrightarrow{p}\ \gamma + \frac{\delta}{\theta}\,.$$

In other words, the bias induced by the failure of the exclusion restriction depends on both the size of the failure (δ) *and* the strength of the instrument (θ). Weaker instruments will tend to make the bias worse. This point was made in, for example, Conley, Hansen and Rossi (2011).

The float subpopulation contains useful information that we now exploit. Note that by construction, $z = 0$ for that subpopulation since $PEG_{it} \times PEG_{i,t-1} = 0$ in expression (1). For the purposes of investigating spillover channels for the floats, we redefine the instrument as

$z = KOPEN \times (\Delta\hat{r}^*_b - \Delta r^*_b)$ (omitting indices to keep things simple). That said, we maintain the assumption that $E(z\,\eta) = 0$. We think this justified since large base economies with monetary policy autonomy are unlikely to consider the macroeconomic outlook of smaller floats when setting rates.

Next, consider estimating (12) using OLS. Estimates of the intervention effect γ and the spillover effect δ will be biased as long as $E(r\,\eta) \neq 0$ and $\theta \neq 0$. However, it is easy to show (under standard regularity conditions) that the OLS estimates of expression (12) are such that

$$\left.\begin{array}{rcl} \hat{\gamma}_{OLS} & \xrightarrow{p} & \gamma - \lambda \\ \hat{\delta}_{OLS} & \xrightarrow{p} & \delta + \theta\,\lambda \end{array}\right\} \qquad \text{with a bias term} \qquad \lambda = \frac{E(r\,\eta)E(z^2)}{E(z\,r)^2 - E(z^2)E(r^2)}\,. \tag{13}$$

Expression (13) is intuitive. If $E(r\,\eta) > 0$, then $\lambda > 0$ and the effect of domestic interest rates on outcomes, $\hat{\gamma}_{OLS}$, will be *attenuated* by the bias term λ. Similarly, the spillover effect, $\hat{\delta}_{OLS}$, will be *amplified* by an amount $\theta\,\lambda$. This amplification will be larger the stronger the correlation between r and z, as measured by the pseudo first-stage coefficient θ.

The difference between OLS and IV estimates reported in Section 5 suggests that there is considerable attenuation bias in $\hat{\gamma}$. 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 $E(r\,\eta) < 0$, then $\lambda < 0$, 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 $\gamma = \alpha\delta$, 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,

$$\hat{\delta}(\alpha) = \frac{(\hat{\delta}_{OLS} + \hat{\theta}\hat{\gamma}_{OLS})}{1 + \alpha\hat{\theta}} \xrightarrow{p} \delta(\alpha)\,. \tag{14}$$

Taking α as given, we can use a control function approach to correct our LP-IV LATE estimates for biases due to potential spillover effects. Expression (12) can then be rewritten as

$$(y - z\hat{\delta}(\alpha)) = r\gamma + \eta + z(\hat{\delta}(\alpha) - \delta(\alpha))\,.$$

Moreover, the usual moment conditions imply that

$$\begin{gathered} E(z(y - z\hat{\delta}(\alpha))) = E(z\,r)\gamma + E(z\,\eta) + E(z^2(\hat{\delta}(\alpha) - \delta(\alpha)))\,, \\ \text{with } E(z\,\eta) = 0\,, \text{ and } (\hat{\delta}(\alpha) - \delta(\alpha))\frac{1}{N_p}\sum_j^{N_p} z_j^2 \xrightarrow{p} 0\,, \end{gathered}$$

as long as

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

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

From this, we have our IV estimator *corrected for potential spillover effects*, which is constructed by subtracting the spillover term from the outcome variable in the standard IV coefficient estimator,

$$\hat{\gamma}(\alpha) \equiv \frac{\frac{1}{N_p}\sum z_j(y_j - z_j\hat{\delta}(\alpha))}{\frac{1}{N_p}\sum z_j\, r_j} \xrightarrow{\ p\ } \gamma(\alpha)\,.$$

We assume that the sample sizes of both float and peg subpopulations tend to infinity.

The appendix provides a more formal derivation for a panel setting such as ours and to account for controls. But the intuition is clear. One can use the subpopulation of floats to estimate the spillover effect channel (the δ in our example). Using this estimate, a control function approach to correct the usual IV estimator consists of simply netting out the spillover channel from the outcome variable. The corrected estimate of the treatment effect γ in (13) will then be consistent.

In practice, α is unknown. We proceed below by using economic arguments to provide an interval of plausible values $\alpha \in [\underline{\alpha}, \overline{\alpha}]$ over which we compute $\hat{\gamma}(\alpha)$. This interval provides a sense of the sensitivity of our benchmark IV estimates to potential spillover contamination. We implement these methods in the next section.

6.1. Spillovers: a robustness check

Table 6 reports OLS estimates of expression (12) with the instrument $z = KOPEN \times (\Delta\hat{r}^*_b - \Delta r^*_b)$, based on the float subpopulation (and by including the usual control set). We define this instrument using the U.K. as the base country for all float economies before 1939 (we drop war years associated with WW1 and WW2); the U.S. for all economies in the Bretton Woods period (1946–1971) and thereafter for Australia, Canada, Japan, and the U.K.; and Germany after Bretton Woods (from 1972 onwards) for any remaining countries (all of which are European). 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 the domestic interest rate on z. This provides an estimate of the parameter θ in expression (13).

Table 6 makes clear the intuition behind expression (13). 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).

To make further progress we need assumptions on α, which is as yet an unknown parameter. Given our application, we assert that it is natural to assume that $\alpha \geq 1$. We think it is reasonable to assume that domestic (rather than base-country) interest rates have stronger effects on economies with floating exchange rates. In order to provide bounds, the results adjusted for potential spillover effects that we report below use a range of values of α from 1 to 4. Economically speaking, this means that the local interest rate effect is 1 to 4 times larger than the effect coming from interest rates driven by, say, the actions of a large external central bank.

Using this control function method to correct the IV estimator, Figure 2 shows results based on the subpopulation of pegs but using the subpopulation of floats to adjust the results. The figure displays responses to a 1% increase in local interest rates for real GDP per capita and the price level using alternative estimators. The red dashed lines show the responses reported in Table 4 in columns 1 and 4 based on LP-OLS and the rich control set described in the previous section. As we remarked earlier, 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 (13).

Next, the solid blue line with associated point-wise error bands in grey show the LP-IV estimates reported in columns 2 and 5 of Table 4. As we noted then, the responses are considerably larger, both statistically (the null $H_0 : LATE = 0$ is rejected at the 1% level) and economically. Finally, the light green shaded region with a dashed border displays the range of impulse responses that would result from our spillover adjustment using $\alpha \in [1, 4]$.

Several results deserve comment. First, notice that 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 due to a 1% increase in rates 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 that has 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. Going forward, we make no spillover corrections to our results. The statistics are easier to interpret absent firm evidence on spillovers.

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*

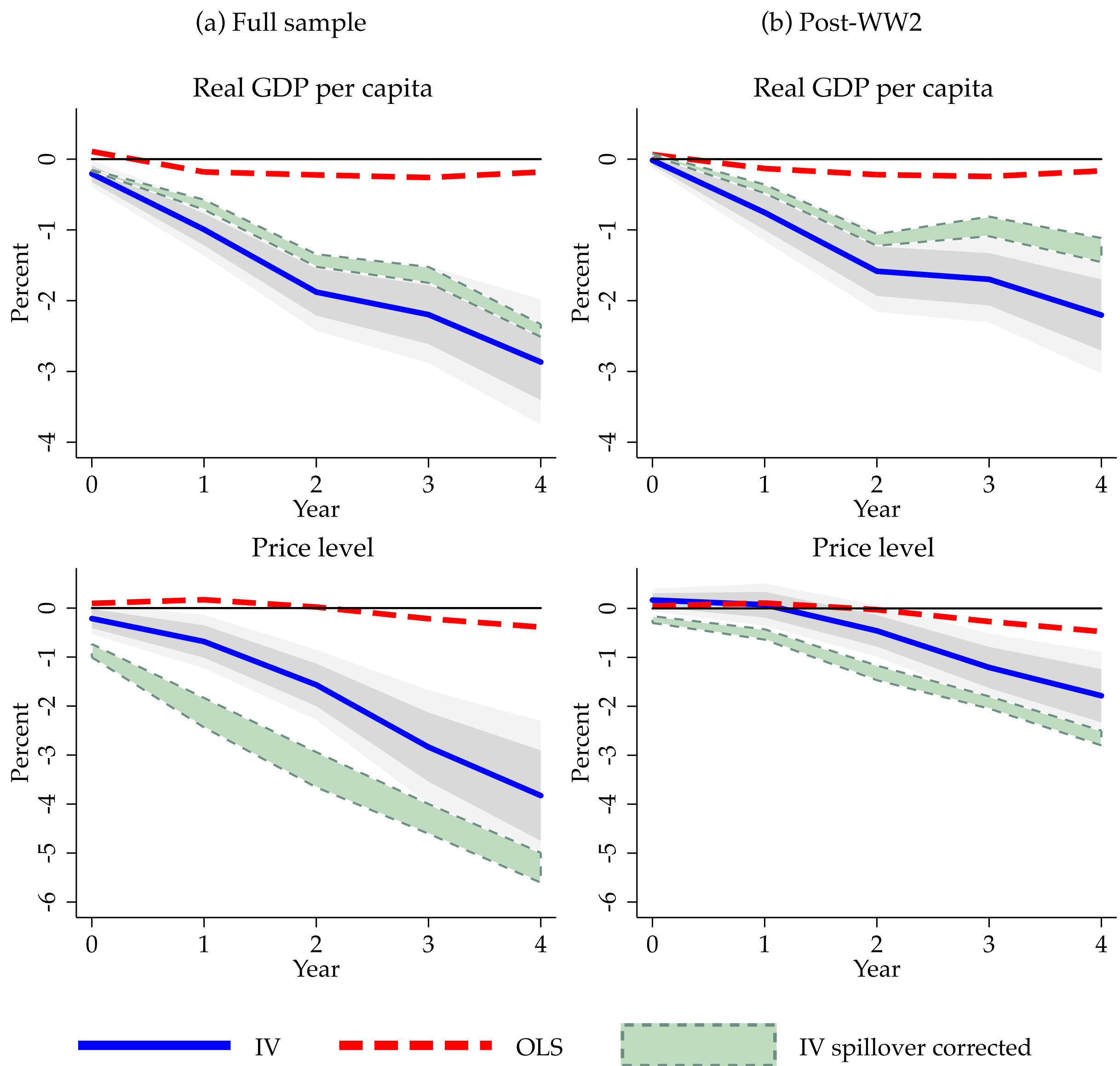
Responses at years 0 to 4 (100× log change from year 0 baseline).

(a) Full sample	*Output response to*		$\Delta r = z$	*Price response to*		$\Delta r = z$
	Δr	z	p-value	Δr	z	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)	
Δr on z	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)	
Δr on z	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. Δr is the rate on short-term (usually 3-months) government bonds of the domestic economy considered and indexed by i; z is defined as $z_{it} = KOPEN_{it} \times (\Delta \hat{r}^*_{b(i),t} - \Delta r^*_{b(i),t})$. $r^*_{b(i),t}$ refers to the base country rate on short-term (usually 3-months) government bonds. 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). $\Delta r = z$ 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. Δr on z refers to the coefficient of the regression of the float rate on the instrument (called θ in the text in expression (12)). See text.

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 $\alpha \in [1, 4]$. See text.

7. The causal effects of interest rates on the macroeconomy

In this section we briefly present a comprehensive study of the causal response 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 the appendix for completeness as Figure A.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.

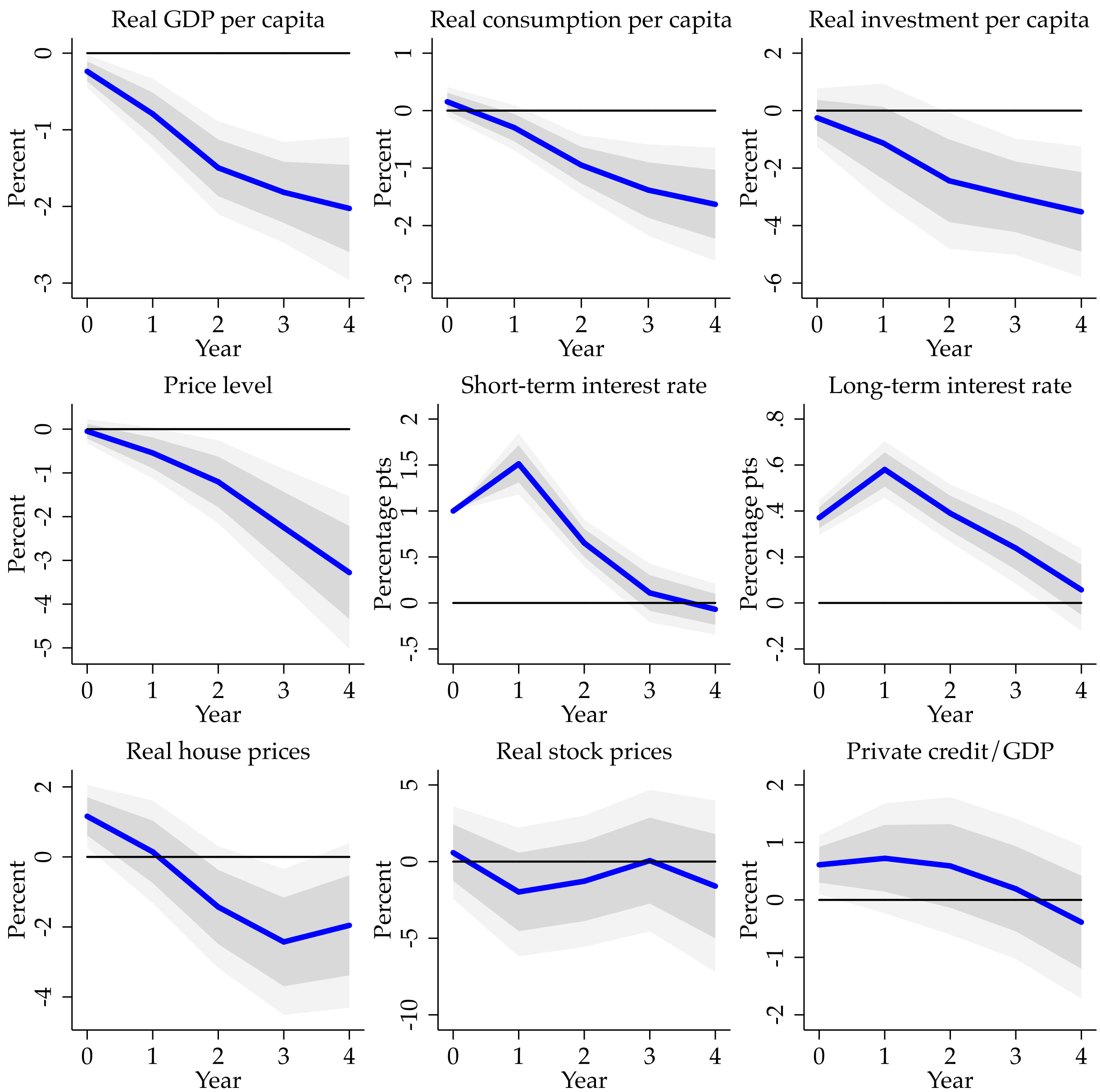
Moving along the second row of charts in Figure 3, we look first at the own response of short-term interest rates to a +100 bps rate rise in year 0 (row 2, column 2). This path reflects the intrinsic persistence of changes in interest rates. In this case, short-term interest rates increase by +150 bps 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.

Proceeding to the last row of charts in Figure 3 (columns 1 and 2) show 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 responses may appear small, note that the nominal responses reported in Figure A.2 in the appendix follow the more familiar pattern. Nominal asset prices drop quickly in response to an interest rate hike. Overall, the responses are consistent with a significant wealth-effect channel for monetary policy, 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

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.

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 A.3 and A.4 in the appendix 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.

All the nine variables displayed in Figures 3, and A.1, A.3, and A.4 in the appendix, are consistent with the intuition gained from the long monetary tradition. The differences that we report with that literature are mostly about much 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.

8. State dependence

In this section we address the possible state dependence of the impulse-response functions. To do this, we shift from a linear to a nonlinear model, a move that is easily accomplished in an LP framework. We stratify our analysis of output and inflation responses, using an indicator of whether the economy is in a boom or a slump, or in low versus high inflation, to address major issues that have emerged in the research literature as well as in current policy debates.

A key concern is that policy responses may be asymmetric. Angrist, Jordà, and Kuersteiner (forthcoming) argue that monetary policy loosening may have a very different impact on macroeconomic outcomes than a tightening. Barnichon and Matthes (2016), using identification via Gaussian basis functions, also find that expansionary monetary policy has virtually no effect, but contractionary policy is powerful. Earlier work by Hamilton and Jordà (2002) investigated a similar asymmetry, based on the nature of the surprise component of changes in the federal funds rate target. In their setting, the asymmetry was motivated less by whether the central bank was in an overall tightening or loosening mood, but whether the public perceived the central bank to be more or less hawkish than current economic conditions would predict.

Adopting an alternative identification method, based on the Romer and Romer (2004) monetary policy instrument used in our analysis, Tenreyro and Thwaites (2016) reach somewhat similar conclusions. They estimate monetary policy responses according to the state of the business cycle with a stratification based on recessions and expansions. Monetary policy has a big effect in expansions, but a smaller effect in recessions. They also find tightening shocks may be more powerful than loosening shocks, but they argue that this is not the entire story since (perhaps surprisingly) in their data contractionary shocks are no more likely in expansions than in recessions.

Similarly, Paul (2015) finds evidence of state dependence when dividing the cycle up into booms and slumps using the "high frequency" identification method of Gertler and Karadi (2015).[9]

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).

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).

8.1. State dependence in boom and slump episodes

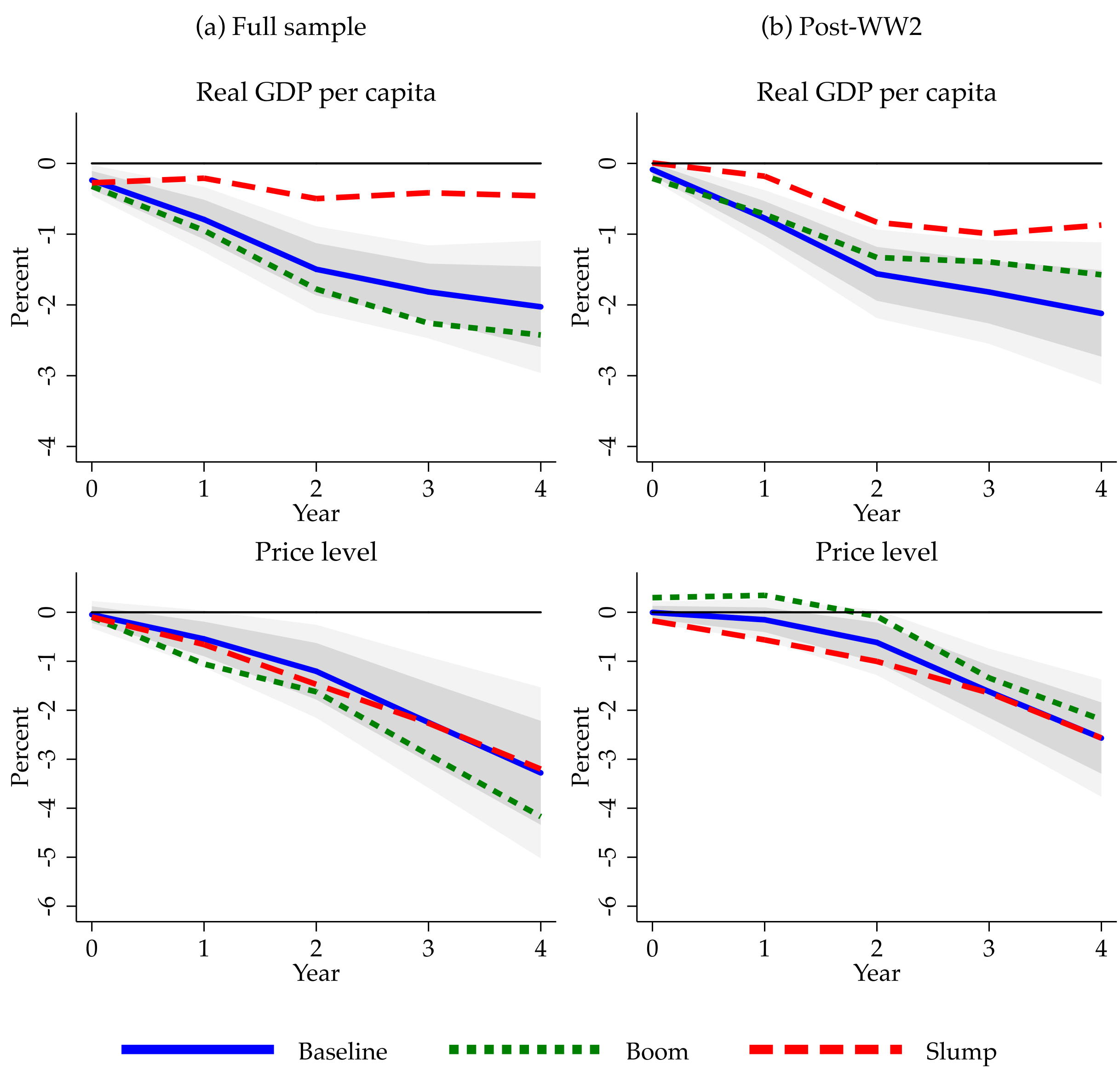
The first set of nonlinear policy experiments are shown in Figure 4 and they investigate the hypothesis that interest rates may have different effects on outcomes in booms than in slumps. We examine this nonlinearity with an indicator state variable equal to 1 in booms and 0 in slumps. Booms (slumps) are periods when the actual level of log real GDP per capita is above (respectively, below) the level of its long-run country-specific trend component as measured by an HP filtered series with a very low-pass setting (λ=100, with annual data).

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. 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

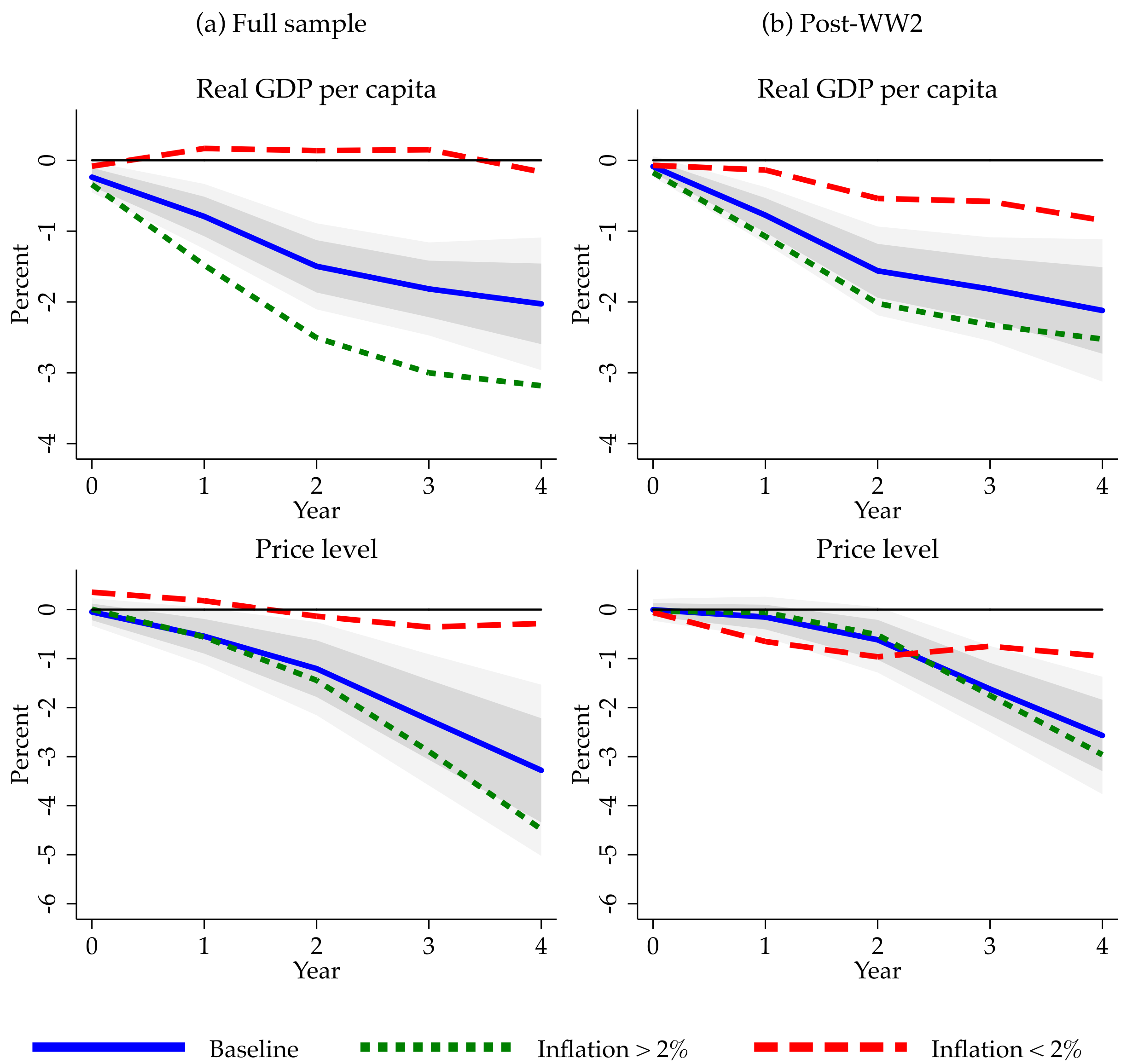
[9] And even before this new wave of studies, which exploit the latest LP methods and identification techniques, there was an older tradition of studies which explored asymmetric monetary policy responses (see, inter alia, Cover 1992; Weise 1999; Ravn and Sola 2004; Lo and Piger 2005).

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.

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.

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.

8.2. State dependence in "lowflation" episodes

Our final 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 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 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%).

To sum up, a central bank rate tightening of $+100$ bps in times when in inflation is above 2% would have causal effects on output that are strongly contractionary. In contrast, a central bank rate loosening of -100 bps in times when in inflation is below 2% would have causal effects on output that are, proportionately, only weakly expansionary. Again, the effects on prices would be roughly symmetric across the two states.

8.3. Summary: implications for a depressed-lowflation trap

Many economists and policymakers now worry that policy could be less effective in some states of the world, such as now, and in a manner that is consistent with several recently emerging theories. This view is strongly supported by our findings. The quantitative message from the long sweep of macroeconomic history in advanced economies is clear, and the two sets of results in this section reinforce each other. The evidence suggests that we will face challenges to policy implementation if we find ourselves in states of the world where output is below potential and/or inflation is very low. In these slow growth or stagnation conditions, the ability to stimulate the economy out of its torpor is that much more difficult; neutral real rates are low, and nominal interest rates are likely to be dragged towards the zero lower bound (Williams 2016). Taken together, these conditions—less bang for a given policy move, and less room to make a move—could make it considerably more difficult for monetary policy to effect change in its principal target variables, output and inflation.

9. Conclusion

The effectiveness of monetary stabilization policy is not only a major policy concern but also an important matter of some controversy among both theoretical and empirical macroeconomists. This paper argues that interest rates have a bigger effect on macroeconomic outcomes than has been previously estimated. The source of the attenuation bias that we report suggests that common assumptions, often based on control arguments, do not provide adequate 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 persistent sub-2% inflation has afflicted 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 this may be nothing new. We find economic responses to monetary policy to be especially weak when the economy is running below potential. Moreover, we also find weak responses in lowflation environments. These challenging conditions have been prevalent in a number of western economies for almost a decade. 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.

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.

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 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, and Patrick Hürtgen. The Macroeconomic Effects of Monetary Policy: A New Measure for the United Kingdom. Bank of England Working Paper 493.

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

Cover, James Peery. 1992. Asymmetric Effects of Positive and Negative Money-Supply Shocks. *Quarterly Journal of Economics* 107(4): 1261–82.

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): 315–331.

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. 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.

Hamilton, James, and Òscar Jordà. 2002. A Model for the Federal Funds Rate Target. *Journal of Political Economy* 5(110): 1135–67.

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

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. 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.

Klein, Michael W., and Jay C. Shambaugh. 2013. Rounding the Corners of the Policy Trilemma: Sources of Monetary Policy Autonomy. NBER Working Paper 19461.

Knoll, Katharina, Moritz Schularick, and Thomas Steger. No Price Like Home: Global House Prices, 1870–2012. *American Economic Review*. Forthcoming.

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.

Lo, Ming Chien, and Jeremy Piger. 2005. Is the Response of Output to Monetary Policy Asymmetric? Evidence from a Regime-Switching Coefficients Model. *Journal of Money, Credit and Banking* 37(5): 865–86.

Nakamura, Emi, and Jón Steinsson. 2013. High Frequency Identification of Monetary Non-Neutrality. NBER Working Paper 19260.

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.

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.

Paul, Pascal. 2015. The Time Varying Transmission of Monetary Policy Surprises. University of Oxford. Unpublished.

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

Ravn, Morten O., and Martin Sola. 2004. Asymmetric Effects of Monetary Policy in the United States. *Federal Reserve Bank of St. Louis Review* 86(5): 41–60.

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.

Weise, Charles L. 1999. The Asymmetric Effects of Monetary Policy: A Nonlinear Vector Autoregression Approach. *Journal of Money, Credit and Banking* 31(1): 85–108.

Williams, John C. 2016. Monetary Policy in a Low R-star World. FRBSF Economic Letter 2016-23.

Wooldridge, Jeffrey M. 2015. Control Function Methods in Applied Econometrics. *Journal of Human Resources*, 50(2): 420–445.

Appendices

A. Selection on observables or conditional ignorability

This section discusses ideal conditions under which covariate control strategies (such as typical structural VAR settings) would allow identification of the average treatment effect (ATE) of monetary policy.

In particular, suppose that, conditional on $\boldsymbol{x}$ the value of r is determined independently of the potential outcome $\boldsymbol{y}(r) = (y(0,r), y(1,r), ..., y(H-1,r))$. In other words, if r reflects the choice of interest rates made by a policymaker, the assumption would imply that $\boldsymbol{x}$ contains all the information used by the policymaker in determining r up to random variation, and therefore the value of r conditional on $\boldsymbol{x}$ is as if it had been randomly assigned. If r were a discrete variable that took on only a small number of values, such as $r \in \{-1, 0, 1\}$, then we would be in the familiar territory of the policy evaluation literature (see, e.g., Angrist 1995) and the full panoply of matching estimators would be at our disposal (for examples based on inverse propensity score weighting see Angrist, Jordà, and Kuersteiner (forthcoming) and Jordà and Taylor (2016)). Here we assume that r takes on continuous values and therefore we seek identification through a different avenue.

Traditional identification of causal effects in monetary economics relies implicitly on assumptions that can be helpfully compared with the potential outcomes paradigm. The first assumption that we state is common in cross-sectional studies. In such settings it is assumed that there is no interference between units (here country-year pairs). That is, the potential outcomes of one unit are independent of the treatment assignment to another unit. This assumption is commonly referred to as *Stable Unit Treatment Value Assumption* or SUTVA (Cox 1958; Rubin 1978) and is generally untestable.

In an international panel setting such as ours, this assumption deserves further discussion. SUTVA is a strong assumption to make regarding *treatment* assignment. The choice of policy interest rates in one country could affect the potential outcomes of another (specially those of a close trading partner). Violations of SUTVA would require explicit modeling of these interactions. The approach that we take here is more pragmatic. In addition to allowing for heterogeneity, we try to control for global factors that would tend to make policy choices correlated. Moreover, we correct standard errors for possible clustering as further insurance. That said, SUTVA is less problematic when evaluating the *instrument* assignment.

Economies that peg their exchange rate are usually small relative to the base economy. In such cases, a reasonable working assumption is that *pegs* have little effect on the policy choices made by *bases*. Moreover, despite having common base economies, pegging countries display considerable heterogeneity in their interest rate choices, as shown in Figure 1. Inference on violations of SUTVA and its consequences undoubtedly deserve further exploration, but that exploration is beyond the scope of this paper. In fact, there is relatively little research available that we can cite on this issue. For all these reasons, we will not explore potential violations of SUTVA further.

We assume that the process determining r can be written as $r(\boldsymbol{x}, \boldsymbol{\psi}, v)$, where the parameter vector $\boldsymbol{\psi} \in \Psi$ is assumed to be fixed (in order to avoid variation in policy due to structural breaks in the policy reaction function rather than through random chance via v). A more detailed discussion of this assumption is provided in Angrist, Jordà, and Kuersteiner (forthcoming). In other words, here we focus on policy interventions that do not come from changes in the policy reaction function.

Here, we state these ideas somewhat formally with the following assumption, which in one form or another, has characterized most empirical work on monetary policy to date. Based on the definitions for $\boldsymbol{y}(r), \boldsymbol{x}$, and r, the selection-on-observables condition (sometimes also called conditional ignorability, or conditional independence) can be stated as:

Assumption 5 (Selection on observables, or conditional ignorability)**.**

$$\boldsymbol{y}(r) \perp r \mid \boldsymbol{x} \qquad \textit{for } \boldsymbol{\psi} \in \Psi \textit{ fixed},$$

where $\perp$ indicates conditional statistical independence and $y(r)$ refers to the set of potential outcomes y could take for a given value of r. We say that perturbations in r given $\boldsymbol{x}$ are independent of the **potential** *outcome $\boldsymbol{y}(r)$ for $\boldsymbol{\psi} \in \Psi$ fixed.*

At first glance, the statement of Assumption 5 may seem peculiar since it appears to indicate that the outcome is orthogonal to the treatment given controls. However, notice that $\boldsymbol{y}(r)$ refers to the *potential* outcomes $\boldsymbol{y}$ can take, not to its realization, $\boldsymbol{y}$. That is, in ideal random assignment of treatment, potential outcomes should not affect the treatment chosen. We are not assuming that the treatment effect is null.

Selection-on-observables requires that any observable information relevant in determining the treatment to be included in the conditioning set. This ensures that any remaining variation in treatment assignment is as good as random. Responses estimated by local projections can be specified and estimated as single equations with rich conditioning sets (in terms of explanatory variables), but in a parametrically conservative manner from the point of view of the number of lags required for consistency (see Jordà 2005 for an explanation of this point). In a panel context, where VAR specifications can be parametrically taxing, single equation panel estimation is particularly convenient.

Notice that Assumption 5 implies that any available information, especially that contained in contemporaneous variables, should be included in the conditioning set. To put it in the parlance of orderings in Cholesky decompositions from VARs, selection on observables would suggest that to calculate the responses to a given shock, the variable being shocked should preferably be ordered last. And if one were interested in calculating the responses to another shock, it would be desirable to reorder the variables again so that the corresponding variable is ordered last as well.

Instead, Assumption 5 can be more naturally accommodated using local projections. Rather than trying to zero out some variables, one can simply include them in the conditioning set. Since each variable can be analyzed separately, there is no need to exclude any contemporaneous information in any of the equations (other than the left hand side variable, naturally). Assumption 5 emphasizes the desirability to condition on as much relevant information as possible to ensure that any remaining variation in the treatment variable is as close to random as possible.

B. Potential violations of the exclusion restriction

Detailed derivations of the results in Section 6 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 (10), extended to include potential spillovers:

$$\begin{aligned} r_{i,t} &= \mu_i + z_{i,t}\theta + \boldsymbol{x}^*_{i,t}\boldsymbol{\phi} + \epsilon_{i,t}\,, \\ y_{i,t+h} &= \alpha_i^h + \hat{r}_{i,t}\gamma_h + z_{i,t}\delta + \boldsymbol{x}^*_{i,t}\boldsymbol{\beta}^*_h + \eta_{i,t+h} \qquad h = 0,1,...,H\,. \end{aligned} \tag{15}$$

In our application we have a scalar instrument, which we use to make the derivations more accessible. The more general case with multiple instruments is easily accommodated, but would require that we use matrix notation formally. We make standard regularity assumptions about the error processes and assume, in particular, that the following moment conditions hold $E(z_{i,t}\epsilon_{i,t}) = E(z_{i,t}\eta_{i,t+h}) = 0$. That is, $z_{i,t}$ 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 z and r. Pre-multiplying expression (15) by $M_{i,t}$ allows us to focus on the coefficients of interest, θ, γ^h and δ. For example, using this projection matrix, note that the OLS estimator of θ is simply

$$\hat{\theta} = \left(\frac{1}{NT}\sum_{t=1}^{T}\sum_{i=1}^{N} z_{i,t}M_{i,t}z_{i,t}\right)^{-1}\left(\frac{1}{NT}\sum_{t=1}^{T}\sum_{i=1}^{N} z_{i,t}M_{i,t}r_{i,t}\right).$$

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

$$S(z,\epsilon) = \left(\frac{1}{NT}\sum_{t=1}^{T}\sum_{i=1}^{N} z_{i,t}M_{i,t}\epsilon_{i,t}\right) \xrightarrow{p} E(z_{i,t}M_{i,t}\epsilon_{i,t})$$

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

$$\hat{\gamma}_h = \frac{S(z,r)\,S(z,y(h)) - S(z^2)\,S(r,y(h))}{S(z,r)^2 - S(z^2)\,S(r^2)}\,, \tag{16}$$

$$\hat{\delta} = \frac{S(z,r)\,S(r,y(h)) - S(r^2)\,S(z,y(h))}{S(z,r)^2 - S(z^2)\,S(r^2)}\,. \tag{17}$$

These moment conditions rely on the usual OLS assumptions $E(r_{i,t}M_{i,t}\eta_{i,t+h}) = 0$, and $E(z_{i,t}M_{i,t}\eta_{i,t+h}) = 0$. 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 $r_{i,t}$. In that case, OLS estimates of $\hat{\gamma}_h$ and $\hat{\delta}$ based on (16) and (17) will be biased.

In particular, we can define the bias term

$$\hat{\lambda} = \frac{S(z^2)\,S(r,\eta(h))}{S(z,r)^2 - S(z^2)\,S(r^2)} \xrightarrow{p} \lambda\,.$$

It is then straightforward to show using expressions (16) and (17) that

$$\begin{aligned} \hat{\gamma}_h &\xrightarrow{p} \gamma - \lambda\,, \\ \hat{\delta} &\xrightarrow{p} \delta + \theta\lambda\,, \end{aligned}$$

where we made use of the standard OLS result

$$\hat{\theta} = \frac{S(z,r)}{S(z^2)} \xrightarrow{p} \theta \,. \tag{18}$$

Notice that under the assumptions of the model, this estimator is consistent for θ. Next, without loss of generality, let $\gamma_h = \alpha\delta$ for α an unrestricted parameter. Then, it is easy to show that

$$\hat{\delta}(\alpha) = \frac{\hat{\theta}\hat{\gamma}_h + \hat{\delta}}{1 + \alpha\hat{\theta}} \xrightarrow{p} \delta(\alpha) \,. \tag{19}$$

In practice the true value of α is unknown. The approach that we follow here is to calculate expression (19) for a range of values of $\alpha \in [\underline{\alpha}, \overline{\alpha}]$. Because α is unknown, we do not attempt to characterize the asymptotic distribution of our spillover corrected estimator.

The next step consists of incorporating $\hat{\delta}(\alpha)$, 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:

$$\begin{aligned}
y^*_{i,t+h} &= r_{i,t}\gamma_h + \boldsymbol{x}^*_{i,t}\boldsymbol{\beta}^*_h + u_{i,t+h} \,, \\
y^*_{i,t+h} &\equiv (y_{i,t+h} - z_{i,t}\hat{\delta}(\alpha)) \,, \\
u_{i,t+h} &\equiv \eta_{i,t+h} + z_{i,t}(\hat{\delta}(\alpha)) - \delta(\alpha))) \,.
\end{aligned}$$

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

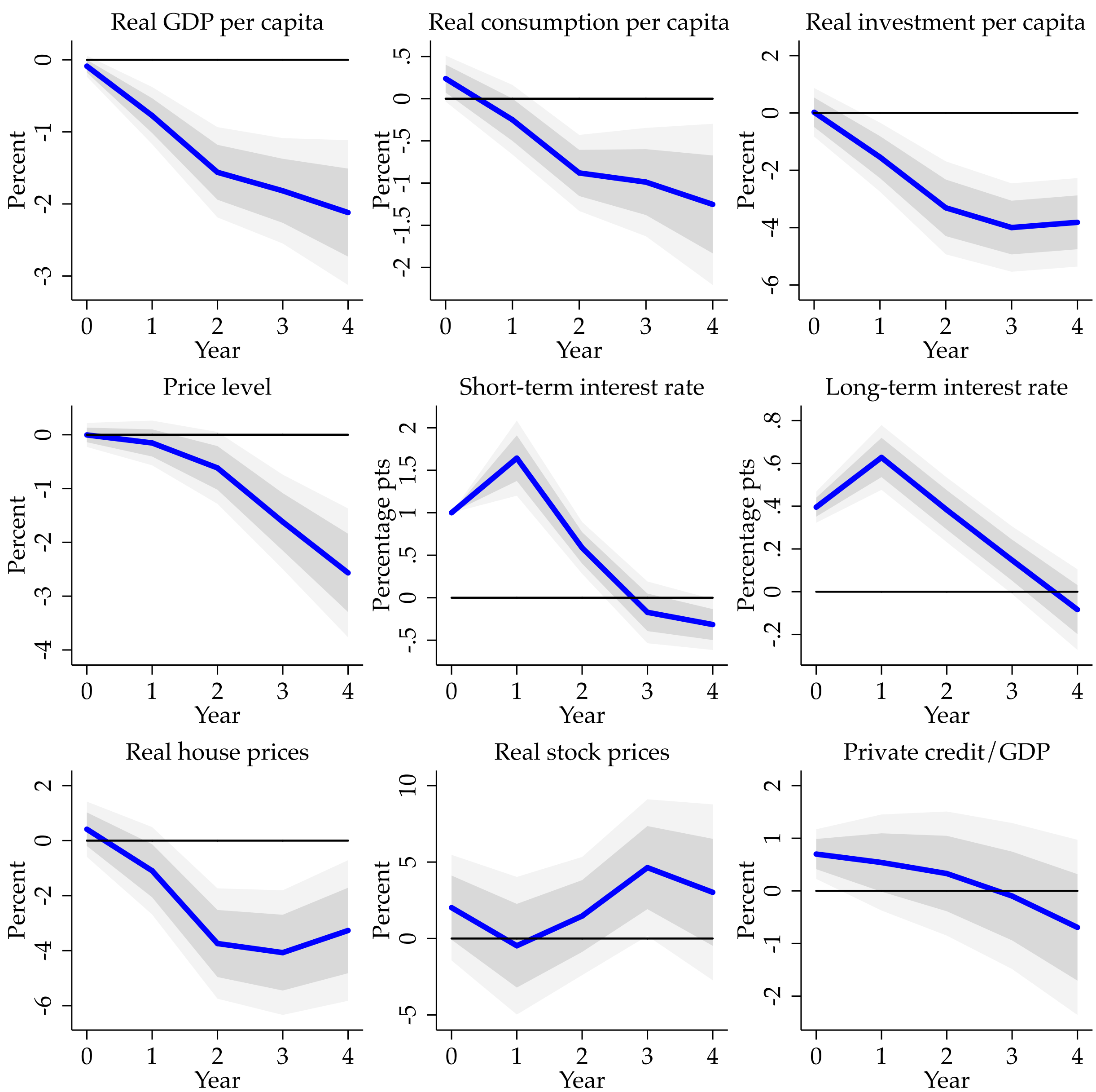
$$\left(\frac{1}{NT} \sum_{t=1}^{T} \sum_{i=1}^{N} z_{i,t} M_{i,t} z_{i,t} \right) (\hat{\delta}(\alpha)) - \delta(\alpha)) \;\xrightarrow{p}\; 0 \quad \text{if} \quad E(z\,M\,z) < \infty \,.$$

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

C. Robustness checks

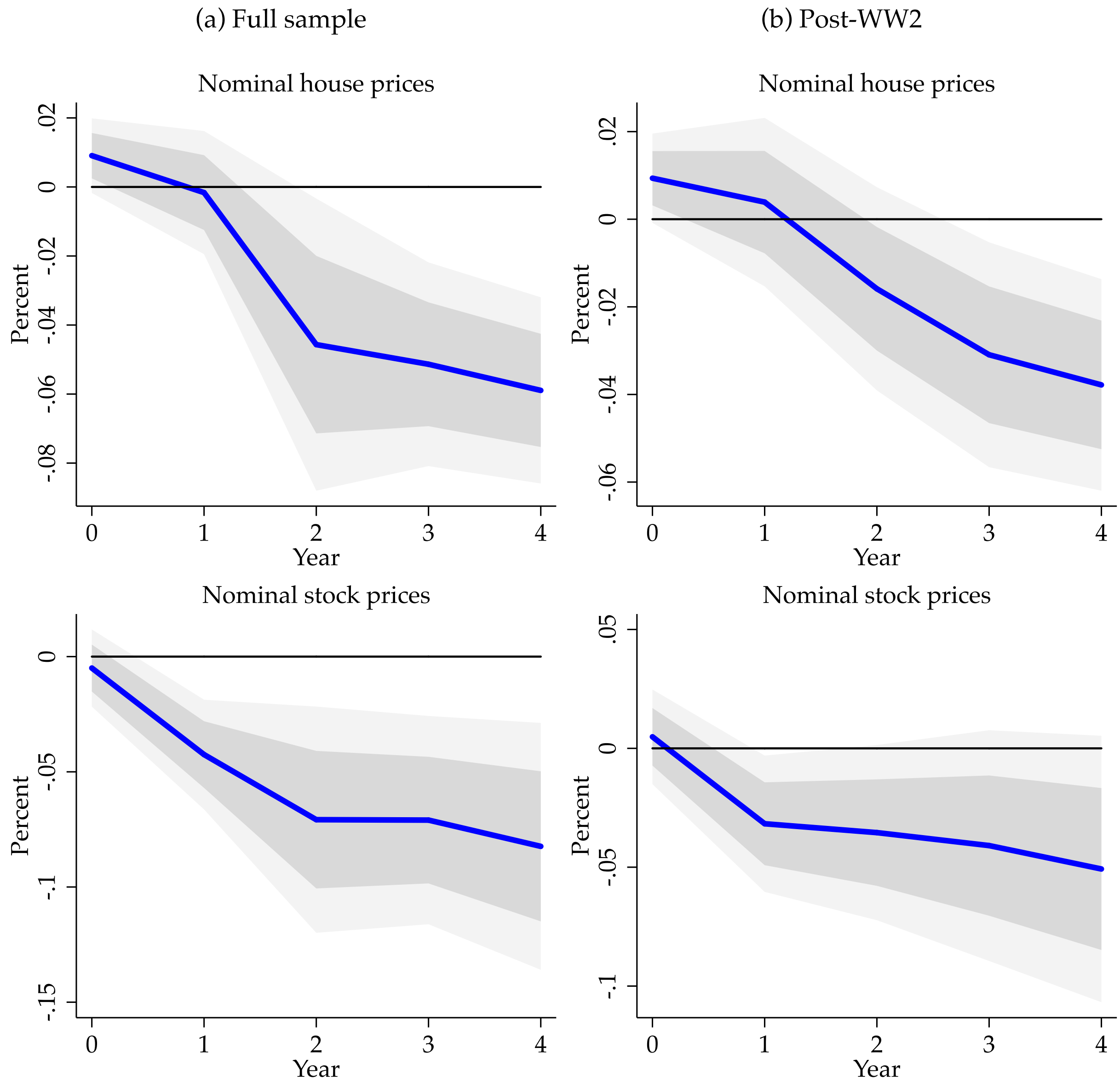
Figures A.1 through A.4 in this appendix contain a series of robustness check which are mentioned in the main text.

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



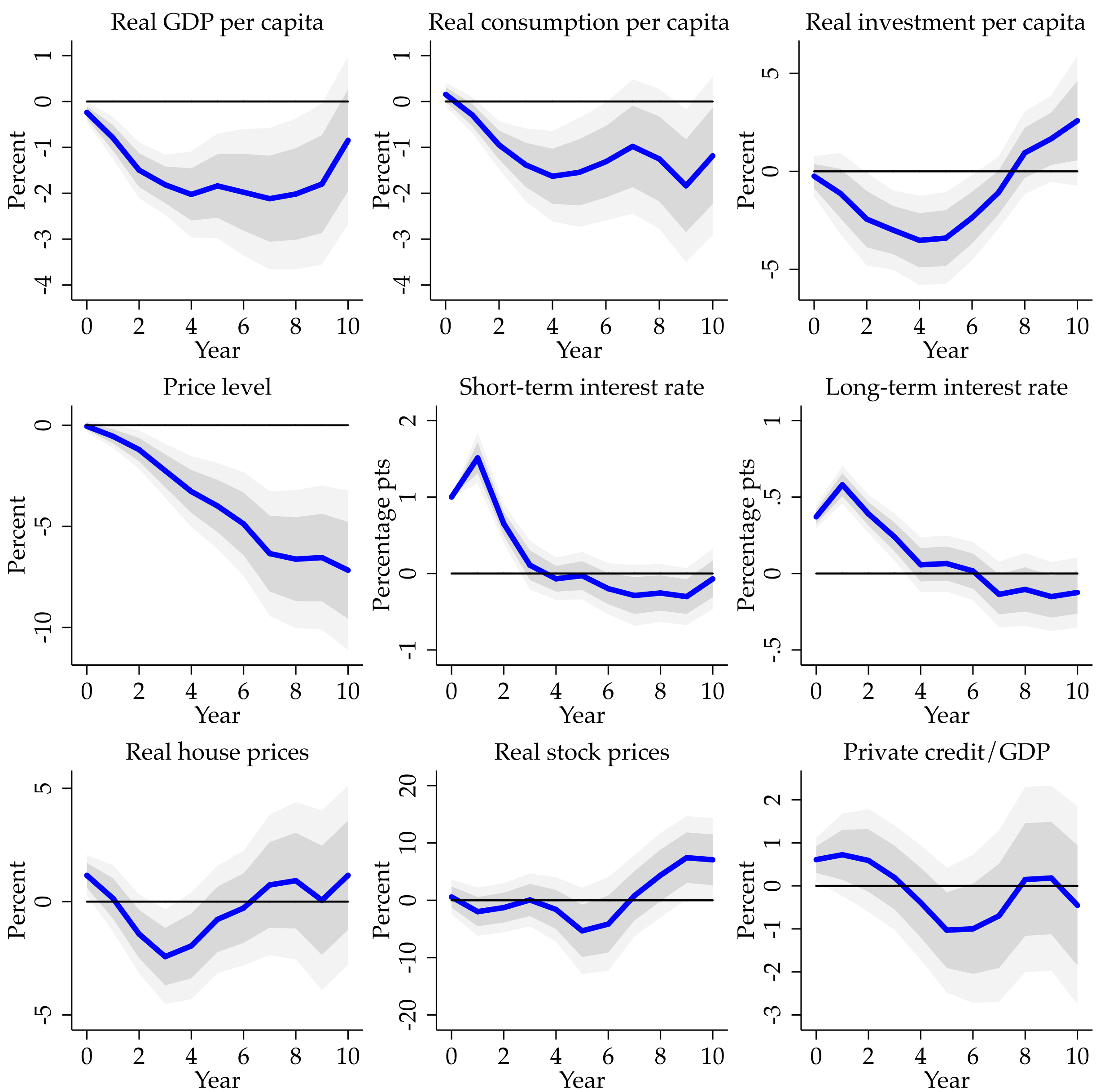
Notes: Post-WW2 sample: 1948–2013. LP-IV estimates displayed with a solid blue line and 1 S.D. and 90% confidence bands. See text.

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



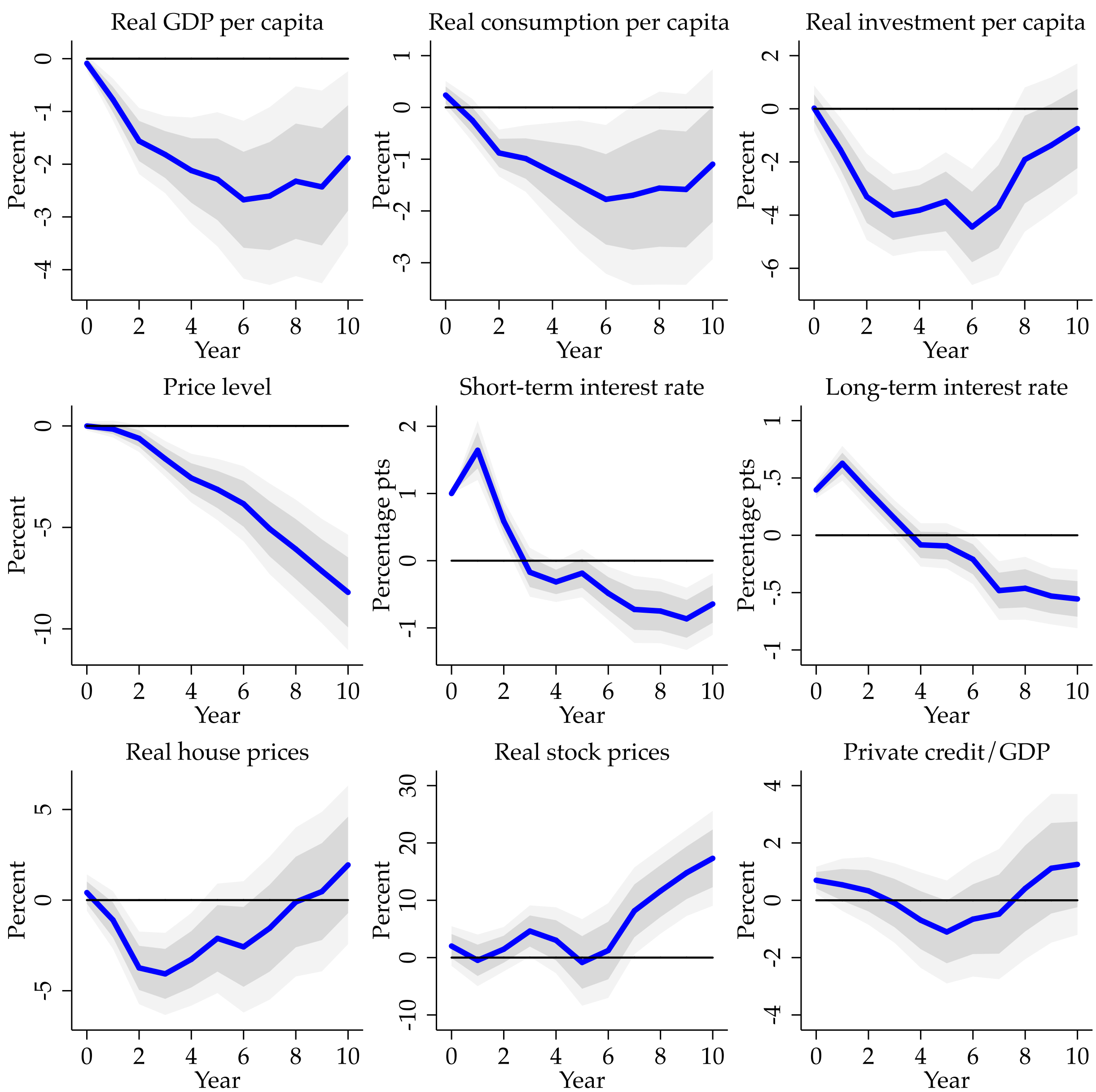
Notes: Full sample: 1870–2013. Post-WW2 sample: 1948–2013. LP-IV estimates displayed with a solid blue line and 1 S.D. and 90% confidence bands. See text.

Figure A.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 A.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.