Questioning the puzzle: fiscal policy, real exchange rate and inflation*

Laurent Ferrara
SKEMA Business School

Luca Metelli
Banca d'Italia

Filippo Natoli
Banca d'Italia

Daniele Siena
Banque de France

October 2019

Abstract

The paper re-investigates the effects of government spending shocks on the real exchange rate, inflation and the trade balance. In contrast with previous *puzzling* results, we find that an increase in government spending appreciates the real exchange rate, is inflationary, induces a trade balance deficit and decreases consumption. The difference with the existing literature lies in the identification of fiscal shocks: embedding a narrative approach in a proxy-SVAR drives our new results. We then show that the empirical findings are consistent with a standard estimated open economy model.

JEL classification: E62, F41, C36

Keywords: fiscal shocks, military narrative shock, proxy SVAR, real exchange rate, inflation

*We thank Fabio Canova, Efrem Castelnuovo, Ethan Ilzetzki, Alex Kriwoluzky, Enrique Mendoza, Gernot Müller, Johannes Pfeifer, Bruce Preston, Morten Ravn, Soren Ravn and participants to the PSE-BdF Workshop, the AMSE-BdF International Macro Workshop, the CEF 2019 and the Melbourne Institute Macroeconomic Policy Meetings for their helpful comments and suggestions. The views expressed in this paper are those of the authors and do not necessarily reflect those of the Banque de France or the Bank of Italy. All the remaining errors are ours. Email contacts: laurent.ferrara@skema.edu, luca.metelli@bancaditalia.it, filippo.natoli@bancaditalia.it, daniele.siena@banque-france.fr.

1 Introduction

Since the Great Recession, the discussion on the role of fiscal policy has gained traction, as discretionary fiscal measures have started afresh to serve as policy tools in advanced economies. The renewed interest in fiscal policy has spurred considerable academic research on its effects, especially on domestic variables such as output and inflation. Less attention has been devoted to the international aspect of these policies, even though, in open economies, real exchange rates and trade balances are important transmission channels. This paper is an attempt to bridge the closed and open economy literature, re-examining the role of fiscal policy shocks in an open economy environment.

Despite the importance of the question, it is still unclear how inflation and the real exchange rate respond to a fiscal shock. According to standard theoretical frameworks, whether Real Business Cycle or old and new-Keynesian theories, inflation should increase and the real exchange rate should appreciate in response to an increase in (unproductive) government spending. However, the empirical literature finds mixed results. On inflation, while Edelberg et al. (1999) and Zeev and Pappa (2017) find that a government spending shock is inflationary, Fatás and Mihov (2001), Mountford and Uhlig (2009), Jorgensen and Ravn (2018) and D'Alessandro et al. (2019) find that the same shock decreases prices. On the real exchange rate, Kim and Roubini (2008) found that fiscal expansions depreciate the real exchange rate and improve the trade balance. This result has then been confirmed by Ravn, Schmitt-Grohé, and Uribe (2006), Monacelli and Perotti (2010), Enders, Muller, and Scholl (2011), Ravn, Schmitt-Grohe, and Uribe (2012) and Ilzetzki, Mendoza, and Vegh (2013). An exception are Born et al. (2013), that finds that under a fixed exchange rate regime the real exchange rate appreciates, and Ilzetzki and Jin (2013) that finds that the response of the real exchange rate depends on the sample considered. In line with different conditional responses, recently, Kim (2015), Auerbach and Gorodnichenko (2016), Forni and Gambetti (2016), Miyamoto, Nguyen, and Sheremirov (2019), Boehm (2019), Lambertini and Proebsting (2019) and Born et al. (2019) argued that the response of the exchange rate might depend on country characteristics, like the stage of economic development, or on the timing of the fiscal shock (namely if it is anticipated or not), or on the sign and type of fiscal shock (government consumption or investment) and on the exchange rate regime.

This paper resumes the debate by employing a different identification scheme to estimate the impact of fiscal spending shocks on inflation and the real exchange rate. The military narrative series constructed by Ramey (2011) and Ramey (2016a) is used as an instrument into a Vector Auto-Regression, using the proxy-SVAR methodology developed

by Mertens and Ravn (2013) and Stock and Watson (2008). We show that, by simply doing this, all puzzling results vanish: government spending shocks are indeed inflationary, appreciate the real exchange rate, worsen the trade balance and decrease private consumption. These dynamics are aligned with theoretical predictions. The responses of a standard estimated real business cycle small open economy model match fairly well empirical impulse-responses. We see this paper as a reconciliation of empirical results with standard theories.

The proxy-SVAR is estimated on quarterly United States data using a Bayesian approach over the 1964Q1-2015Q4 period. Two important aspects should be emphasised at the very outset. First, even though Ramey (2016a) has constructed a narrative series to instrument both contemporaneous and anticipated government spending, we use it here to instrument only unanticipated shocks. We show that this series is indeed a valid and good instrument for contemporaneous government spending in the 1964-2015 period, satisfying both the relevance and the exclusion restrictions. Second, given the importance of the time-frame for fiscal estimates, we pick as a baseline the 1964Q1-2015Q4 period to use the official real effective exchange rate data from the Bank of International Settlement (BIS), available at the earliest from 1964. However, multiple robustness checks, enlarging the sample or excluding the Bretton-Woods' and Great Recession periods, are also included.

Our focus on surprise government spending shocks, as opposed to anticipated shocks, is chosen to make our results comparable with the existing, puzzling, evidence on inflation and the real exchange rate. We show that our instrument is indeed capturing current and not future, anticipated, movements in government spending: we test the relevance condition of our instrument using standard F-tests, as in Ramey (2016b), but including as control the lagged variables of the baseline SVAR specification. In the baseline proxy-SVAR we also make sure to include a variable which responds on impact only in the presence of new information, i.e. the market value of military firms proxied by the stock price of Boeing, one of the leader manufacturers of defense airplanes and second US Federal government contractor in 2015.

Theory-consistent responses to a positive government spending shock are also found when the estimation is carried out in the post-1976 sample (which excludes the Bretton-Wood's period, Kim et al., 2017) or when we exclude the Great Recession period (i.e., with sample ending in 2006). Using nominal exchange rates or using a different definition of inflation – based on the consumer price index instead of the personal consumption expenditure index – does not change the results. Moreover, using defense government investment in the place of Ramey (2016a)'s narrative series as an instrument for government spending as in Miyamoto et al. (2019) confirms our results (Section 4).

Finally, in order to check the theoretical coherence of our empirical results, we set up a standard two-good RBC small open economy model and we show that the empirical and theoretical impulse-responses to a government spending shock are indeed coherent. This is true for both international and domestic variables. The model is estimated through an impulse-response matching procedure to set three critical parameters governing the response of the real exchange rate, consumption and the trade balance (see Corsetti et al., 2008 and Monacelli and Perotti, 2010): the trade elasticity, the persistency of the shocks and the wealth elasticity of labor supply.

Related Literature — Our paper draws on different strands of literature. First, it is closely related to the literature analyzing the empirical effects of fiscal policy on the real exchange rate and inflation. The seminal paper focusing on exchange rate responses is Kim and Roubini (2008), where the authors document a US real exchange rate depreciation following a positive US fiscal shock, at odds with what the theory predicts. They also document a counterintuitive reaction of the trade balance, which improves instead of deteriorating. Such puzzling results ignited a stream of the literature which mainly confirmed these empirical regularities. Monacelli and Perotti (2010) find that, in the US and other advanced economies, a rise in government spending induces a depreciation of the CPI real exchange rate and a trade balance deficit. They also find that private consumption rises in response to a government spending shock, in line with Blanchard and Perotti (2002). Ravn et al. (2012) use a panel structural VAR analysis to document that an increase in government purchases raises output and private consumption, deteriorates the trade balance, and depreciates the real exchange rate, both in the US and in other four industrialized countries. Enders et al. (2011) find, using sign restrictions, that the exogenous expansions of government spending depreciates the real exchange rate and the terms of trade. Ilzetzki et al. (2013) concentrate on the output effect of fiscal policy, but it highlights the same puzzling response of the real exchange rate, using a panel of 44 countries. More recently, Kim (2015) investigated again the question, examining 19 OECD countries. The author finds that current account worsens and real exchange rate appreciates in the majority of the countries, but various country characteristics (e.g. trade openness, capital mobility, etc.) are driving the result. Similarly, Miyamoto et al. (2019) explore the response of the exchange rate to a government spending shock differentiating between advanced and emerging countries. They identify the shock using annual military expenditures and find an appreciating (depreciating) exchange rate in emerging (advanced) economies. Even if focusing on a different aspect, Boehm (2019) shows that a government investment shock, and not a government consumption shock, can slightly

appreciate the real exchange rate when the country has a floating nominal exchange rate (Ilzetzki et al., 2017). Contrastingly, Lambertini and Proebsting (2019) find an appreciating exchange rate in a monetary union. Finally, Born et al. (2013) and Born et al. (2019) show an asymmetric reaction of real exchange rate responses due to the exchange rate regime and to the sign of government spending shocks.

The two papers closest to our findings in terms of exchange rate responses are Auerbach and Gorodnichenko (2016) and Forni and Gambetti (2016). Auerbach and Gorodnichenko (2016) use daily data on U.S. defense spending and documents that the dollar immediately and strongly appreciates after announcements of future government spending. On the contrary, when actual payments are made, spending variations have no significant effects on the exchange rate. Forni and Gambetti (2016) use the Survey of Professional Forecasters to account for both government spending anticipated and surprise shocks. They estimate the effects of both types of shocks using a quarterly VAR from the 80's, finding that anticipated shocks generate an appreciation of the real exchange rate, while surprise shocks generate a depreciation. Differently, our paper focuses on unanticipated government spending shocks, where the puzzling results were first found and are still topical, evaluating its effects on all those variables for which previous studies found empirical dynamics at odds with standard theory, jointly at the domestic and international level.

Focusing on the effects of fiscal policy on inflation, results are also mixed. Edelberg et al. (1999) and Zeev and Pappa (2017) find that a government spending shock is inflationary. Other studies, like Nakamura and Steinsson (2014) and Perotti (2005), find either a non-significant response or at least mixed evidence. However, a larger and more recent set of papers, like Fatás and Mihov (2001), Mountford and Uhlig (2009), Jorgensen and Ravn (2018) and D'Alessandro et al. (2019) find that a government spending shock decreases prices. In particular, Jorgensen and Ravn (2018), using data from the 80's and adopting various identification schemes, document that in response to an increase in government spending, inflation falls. They rationalize the negative behavior of inflation by showing that a fiscal shock increases domestic productivity, hence generating a supply side boost which more than compensate the increase in aggregate demand. Similar results are found by D'Alessandro et al. (2019), which develops a quarterly Bayesian VAR including fiscal and TFP variables for the period 1954Q3-2007Q4, finding that inflation turns negative after a positive fiscal shock.

One common feature over most of the aforementioned papers is the identification methods adopted in order to recover the structural fiscal shock. Indeed, such empirical investigations are based either on the Blanchard and Perotti (2002) restrictions on the

variance-covariance matrix, or on sign restrictions, or on the use the narrative series on a recent sample. Consistently, using Blanchard and Perotti (2002) identification scheme on our sample we also confirm *puzzling* results (Figure 2). However, we highlight here that combining a rigourous narrative method with the SVAR structure on a sufficiently long sample solves the puzzles.

Clearly, our paper is also related to the literature on the estimation methods of fiscal policy shocks. In particular, our paper draws on the proxy-SVAR methodology, developed independently by Mertens and Ravn (2013) and Stock and Watson (2008) and on the narrative approach of Ramey (2011) and Ramey (2016a).[1]

Our focus on international relative prices makes our findings relevant for the literature studying fiscal spillovers as well (Faccini et al. (2016), Corsetti et al. (2009), Corsetti et al. (2011), Corsetti and Muller (2013) and Auerbach and Gorodnichenko (2013)). In fact, in order to understand the transmission of fiscal shocks across countries, real exchange rate responses are obviously crucial.

Last, our paper is related to the theoretical literature analyzing the economic effects of fiscal policies. A standard closed economy neo-classical model (Baxter and King, 1993) would suggest that an increase in unproductive government spending would generate a fall in private consumption (via a negative wealth effect due to the increase in the present value of taxes to be paid) and an increase in prices. Empirically, however, most of the evidence pointed towards an increase in private consumption and a fall in prices in response to a positive government spending shock. This mismatch between theory and empirics has been shaping theoretical studies, which tried to rationalize the empirical findings (see, for example, Basu and S. Kimball, 2003, Linnemann, 2006, Ravn et al., 2006, Galí et al., 2007 and more recently Jorgensen and Ravn, 2018 and D'Alessandro et al., 2019). A similar contrast between theoretical predictions and empirical evidence drove also the theoretical literature looking at the impact of fiscal policy in open economies. A benchmark general equilibrium open economy model featuring complete financial markets would imply that an increase in government spending would generate an appreciation of the exchange rate, a fall in the trade balance and a fall in consumption. Empirically, however, the evidence was pointing towards a depreciation of the real exchange rate, an increase in the trade balance and an increase in consumption. Monacelli and Perotti (2008) and Monacelli and Perotti (2010) describe well the empirical vs theoretical inconsistencies: benchmark open economy models including the wealth effect of government spending and perfect risk-sharing across countries cannot rationalize simultaneously the effects on quantities and relative prices, and even more so if government spending is intensive in

[1] See Angelini et al. (2019) for an application using different instruments in a proxy-SVAR approach.

non-traded goods. To solve those issues, two theoretical solutions have been proposed: first, counteract the negative wealth effect coming from government spending by assuming non-separable utility or equilibrium variable markups (Monacelli and Perotti, 2010); second, calibrate the model with a low trade elasticity (Enders et al., 2011). In order to compare our results with this literature, our model will account for the possibility of these two features and will estimate their relevance via an impulse-responses matching procedure.

The rest of the paper is organized as follows. Section 2 briefly describes the proxy-SVAR methodology, the identification strategy, the data and specification adopted in the paper. Section 3 presents the empirical results. Section 4 shows the robustness of the results to different specifications. Section 5 matches the theoretical and empirical impulse-responses by estimating a standard two-good RBC small open economy model. Finally, Section 6 concludes.

2 Empirical model and identification strategy

In this Section we introduce our empirical model and the identification strategy. First, we briefly describe the proxy-SVAR methodology. Second, we present our set of target variables. Third, we discuss the use of the military narrative series as an instrument for unanticipated government spending shocks.

2.1 The proxy-SVAR framework

Consider the following Vector AutoRegressive (VAR) model:

$$X_t = c_0 + \sum_{k=1}^{P} A_k X_{t-k} + u_t \qquad u_t \sim N(0, \Sigma_u) \tag{1}$$

where X_t is a vector of endogenous variables, A_k are the matrices containing the reduced-form parameters, u_t is the vector of reduced-form residuals and Σ_u is the covariance matrix of the reduced-form shocks. In order to identify structural shocks in the VAR, one needs to specify a matrix P_0 that pre-multiplying Equation 1 yields:

$$P_0 X_t = P_0 c_0 + P_0 \sum_{k=1}^{P} A_k X_{t-k} + \epsilon_t \tag{2}$$

where $\epsilon_t = P_0 u_t$ is the vector of structural shocks with mean zero and covariance matrix

Σ_ϵ. To construct the matrix P_0 to identify fiscal shocks in the United States, we use the proxy-SVAR methodology, developed independently by Mertens and Ravn (2013) and Stock and Watson (2008). Restrictions on P_0 are obtained by making use of a proxy of the true latent exogenous variable. We employ a narrative measure m_t to proxy for the unobserved fiscal shock $\epsilon_{f,t}$, where we assume $E(m_t) = 0$; In addition, denoting the non-fiscal US shocks as $\epsilon_{nf,t}$, our narrative measure needs to satisfy the following two conditions:

$$E[m_t, \epsilon_{f,t}] = \gamma \tag{3}$$

$$E[m_t, \epsilon_{nf,t}] = 0 \tag{4}$$

This means that our proxy m_t is correlated with the unobserved fiscal policy shock but it is orthogonal to the remaining shocks. This methodology provides the restrictions for the columns of the matrix P_0 related to the fiscal variable. To obtain them, we follow the standard two-step procedure for proxy-SVARs: first, we run a two-stage least squares (2SLS) estimation of all non-fiscal residuals in the US model ($u_{nf,t}$) on the fiscal ones, using m_t as an instrument for $u_{f,t}$: the estimated coefficients represent each variables' restrictions up to a scale factor; second, we impose covariance restrictions to identify each element in the l^{th} column of P_0. Details on the proxy SVAR procedure are reported in Mertens and Ravn (2013).

2.2 Data and specification

Narrative measures of fiscal policy changes, both for taxes and spending shocks, have been constructed in the literature from historical sources. These measures are imperfectly correlated with latent structural policy shocks, mostly because of measurement errors: historical records sometimes contradict each other and narrative series typically disregard minor policy changes. Using a proxy-SVAR approach has the advantage of extending the use of proxy series to cases where we know that these are measured with errors. In the literature this methodology has been used to proxy tax shocks while here we extend it to identify unanticipated government spending shocks (see Section 2.3). Against this backdrop, we use the military spending narrative series constructed by Ramey (2016a).

The baseline specification of our VAR model encompasses the following US variables: real government spending G_t, real GDP y_t, tax revenues tax_t, real private consumption c_t, inflation π_t, total factor productivity (TFP) tfp_t, trade balance (in percent of GDP) TB_t, the stock price of Boeing (proxying the market value of the military firms sector) s_t and

the narrow real effective exchange rate of the dollar $reer_t$.[2] With the only exception of inflation and the trade balance-to-GDP ratio, all other variables are taken in logs. Inflation is computed on an annual basis using the personal consumption expenditure (PCE) deflator. Real government spending, real GDP and real tax revenues are obtained by deflating nominal variables using the GDP deflator; differently, private consumption is deflated using the PCE deflator. The TFP variable is taken from Fernald (2012), and the military narrative series stems from Ramey (2016a). Nominal GDP, government spending and tax revenues are taken from Ramey and Zubairy (2018). Stock prices are taken from Yahoo!Finance. The real effective exchange rate, as well as the nominal exchange rate used in a robustness check, are taken from the BIS database. Data on nominal defense government investment, used in another robustness section, are deflated with the defense consumption and investment deflator; both the defense and deflator variables are taken from the FRED database.

We estimate the model on quarterly data and, as it is standard in the literature, we include the constant and four lags of the endogenous variables. The baseline estimation sample ranges from 1964Q1 to 2015Q4.[3] The model is estimated using Bayesian techniques, performed via a block MCMC algorithm. We use the dummy method of Del Negro and Schorfheide (2011) and Caldara and Kamps (2017) and we impose a Minnesota prior on the reduced-form VAR parameters; in addition, we choose the hyper-parameters governing the prior distributions in order to impose relatively weak priors.

2.3 The narrative series in the proxy-SVAR framework

Puzzling results on the response of inflation and exchange rate have be found in the literature in reaction to surprise, contemporaneous, government spending shocks. We therefore want to focus on these shocks. Consistently, the use of a proxy-SVAR allows us to extract from a narrative series the information to instrument current (and not future) fluctuations.

However, to proxy contemporaneous US government spending shocks we use the Ramey (2016a) narrative series which has been constructed to capture the net present value of both current and expected military expenditures (i.e. surprise plus anticipated movements). We therefore need to explain and support our choice. The argument prompted by Ramey (2016a) and Ramey and Zubairy (2018) is that during the largest military episodes

[2] The broad effective exchange rate is only available starting in 1994.

[3] As already anticipated, such sample interval is the widest possible given the constraints on data availability: data on real effective exchange rate starts in 1964Q1 and the narrative military series ends in 2015Q4. We exclude from the sample the years of the recent financial crisis, i.e. from 2007Q3 to 2009Q4.

of the recent US history (i.e. World War I, World War II and the Korean War), long lasting spendings have been anticipated by information reported by the press: these are therefore positive, anticipated, government spending shocks. As a result, the historical narrative series including war episodes should be viewed as information on current and future government expenditures. In line with this reasoning, Ramey (2016b) and Ramey and Zubairy (2018) show that for the narrative series to be a good instrument of anticipated government spending episodes, the war episodes need to be included in the sample. An F-test obtained by including the military narrative variable in a regression model of government spending over controls shows exactly this (Ramey, 2016b) (replicated in the red-dotted line of Figure 1).

Provided that real effective exchange rates for the United States are not available before 1964, we are forced to focus our attention on a shorter sample, which does not include war episodes. This, alone, should push towards the use of the narrative instrument as an instrument for unanticipated shocks. However, we repeat the F-tests to check the strength of our instrument for unanticipated shocks on our sample period. We proceed as follow: we regress cumulated spending on the military narrative series at time t and four lags of control variables. This regression can be written as

$$\sum_{j=0}^{h} g_{t+j} = \gamma_h + m_h \text{narrative}_t + \phi_h(L) z_{t-1} + \omega_{t+h} \tag{5}$$

where $\sum_{j=0}^{h} g_{t+j}$ is the sum of current and future government spending, narrative $_t$ is the military narrative series and z_{t-1} is the set of lagged controls.[4] The test is computed against an alternative specification which excludes the narrative series from the set of regressors, i.e.

$$\sum_{j=0}^{h} g_{t+j} = \gamma_h + \phi_h(L) z_{t-1} + \omega_{t+h}. \tag{6}$$

We run three specifications of the F-test: first, on the 1964-2015 sample using only tax revenue and GDP as controls, as in Ramey and Zubairy, 2018; second, on the same 1964-2015 sample but enriching the set of controls with all variables that are present in the baseline SVAR; third, on the Ramey's post World War II sample (1947-2015).

Results of the F-tests, each of them conducted with $h = 20$ (i.e., from 0- to 20-quarter horizon), are displayed in Figure 1. The Figure reports the F-statistics minus the appropriate critical value threshold. Indeed, according to whether the residuals of Equation

[4]The findings from the F-tests are robust to the alternative specification of the dependent variable as g_{t+j} instead of $\sum_{j=0}^{h} g_{t+j}$.

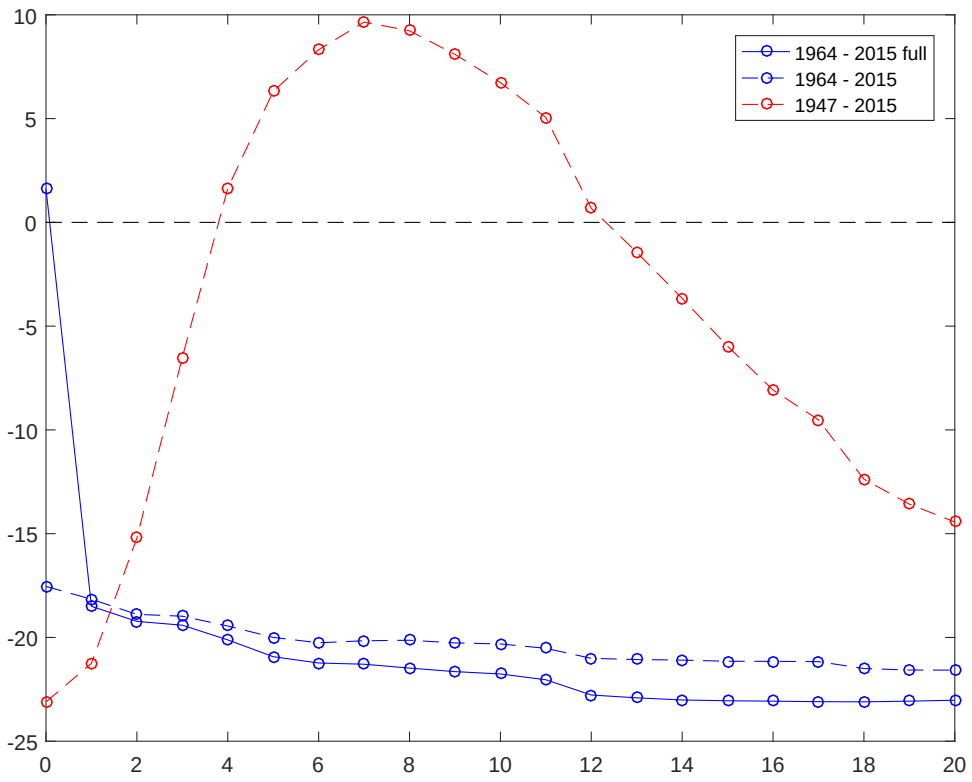

Figure 1: **Deviations over horizons of F-statistics from their critical values**. First-stage F-statistics for government spending shocks. The F-statistics are based on the regression of the sum of government spending from t to $t+h$ on the military narrative series at t, plus 4 lags of the spending and narrative variables and of additional control variables. Controls for the 1964-2015 and 1947-2015 specifications (blue and red dotted lines) are tax revenue and GDP, while the baseline 1964-2015 specification (blue solid line) has additional controls (inflation, TFP, consumption, short term interest rate and the real exchange rate). The horizontal dashed line at zero is the weak instrument threshold. A value above zero indicates that the test accepts the instrument to be a valid one.

(5) are autocorrelated or not, the corresponding F-test have different critical values, as Montiel Olea and Pflueger (2013) clarifies. In order to take this into account, we run the Ljung-Box Q-test for residual autocorrelation on the three regression specifications, one for each h series of residuals ω_{t+h}. Every test indicate serial correlations from 4 to 20 lags. Results, available upon request, show that, as expected, residuals of h-quarter ahead predictive regressions (with $h > 0$) are all autocorrelated. This is because control variables do not include time $t+h-1$ observations. Concerning contemporaneous regressions (i.e. $h = 0$), residuals are still autocorrelated when controlling for tax revenues and GDP only. However, they become *not* autocorrelated when controlling for the full set of variables in

our baseline specification. Therefore, the (lower) critical value for serially uncorrelated error terms is considered only for impact F-statistics (i.e., $h = 0$) in the baseline 1964-2015 specification with a full set of controls; in all other cases, the significance of the F-test is judged with respect to the threshold of autocorrelated error terms.[5]

By reporting *relative* F-statistics, the zero line in Figure 1 represents the meaningful threshold to test for weak instrument. Therefore, a value above zero indicates that the test accepts the instrument to be a valid one. Two facts emerge from the analysis: (1) in the 1964-2015 sample, the narrative series is a good instrument only when instrumenting *contemporaneous* government spending (first point of the solid blue line);[6] moreover, this is true only in the larger (full) specification and not in the three-variable model (dashed blue line). This confirms the need for a larger set of relevant control variables to improve the test precision (Stock and Watson, 2018);[7] (2) the narrative series is a good instrument for anticipated government spending only when including also the Korean-War (dashed red line) and this is not anymore the case for the 1964-2015 sample. Over this period, the narrative series has no relevance at all at future horizons.

One possible reason for this finding, in contrast with the existing literature, could be related to a decrease in the implementation lag of some types of military expenditures outside war episodes. All in all, our results suggest that the military narrative series constructed in Ramey (2016a) is a valid instrument for future spending when the sample starts before the Korean-war but it is a relevant instrument for current government spending when considering a more recent sample.

To exclude the possibility that our proxy is a good instrument also for variables other than government spending, we repeat our F-tests by substituting g with one of the other variable at a time on the left hand side of Equations 5 and 6. Results, displayed in Figure 7 in Appendix A, show that the F-test fails for all variables at all horizons but for contemporaneous government spending.[8]

Finally, another potential issue lying in our identifcation strategy relates to non fundamentalness, as highlighted by Forni and Gambetti (2014). Indeed, if a VAR model does not contain sufficient information, it is not possible to recover the true structural shocks.

[5]For the serially uncorrelated case, we apply the threshold of Montiel Olea et al. (2018) - i.e. 3.84. For the other cases we use the one proposed by Montiel Olea and Pflueger (2013), and used in Ramey and Zubairy (2018), which is 23.1085.

[6]The same result holds when using only the post Bretton-Woods period, which ensures that our instrument is a good one also when considering the period with a flexible exchange rate regime (the f-test result is available upon request).

[7]Results hold by also including in the set of controls the principal components extracted from the dataset of macroeconomic variables of McCracken and Ng (2016), with the value of the F-statistic being even larger.

[8]We also implement the testing procedure proposed by Lunsford (2016), projecting the proxy variable on the VAR reduced-form residuals, and we obtain an F-statistic of 2.7.

Forni and Gambetti (2014) show the necessary and sufficient conditions under which the VAR is invertible and propose a test to detect non-fundamentalness.[9] The idea of the test rests on the assumption that structural shocks ϵ_t, estimated in the following section, cannot be Granger-caused by any other variable. In the spirit of Forni and Gambetti (2014) we project the recovered structural spending shock on the lagged principal components extracted from a large dataset of macro variables (McCracken and Ng, 2016), which summarize the information set of the econometrician. In order to assess fundamentalness in our environment we test whether the coefficients ψ in the following regression are jointly significant:

$$\epsilon_t = \delta + m_h \sum_{j=1}^{nPC} \psi_j \mathrm{PC}_{j,t-1} + \phi_t \tag{7}$$

where δ is a constant, PC stands for the principal components and nPC is the number of PC considered. The F-statistic is 0.0188, failing to reject the null hypothesis of fundamentalness.

3 Empirical Results

This section presents the main results from the empirical analysis. First, we compare the responses of all variables, with a particular focus on the exchange rate and inflation, to a government spending shock identified through the proxy-SVAR methodology with responses stemming from a standard Cholesky identification method. Second, we provide empirical evidence on the robustness of our results.

3.1 Impulse response functions

We start by showing standard *puzzling* results. Figure 2 reports responses to a one standard deviation positive shock to US government spending, using the recursive Cholesky identification method on the 1964-2015 sample. The real exchange rate depreciates (here defined as number of foreign goods for domestic ones), inflation falls, trade balance improves and consumption increases.

Figure 3 displays instead the responses of the same variables (on the same sample) when the fiscal shock is identified using the proxy-SVAR methodology. We find that the

[9] Canova and Sahneh (2018) propose an alternative method to test for non-fundamentalness in small-scale SVAR.

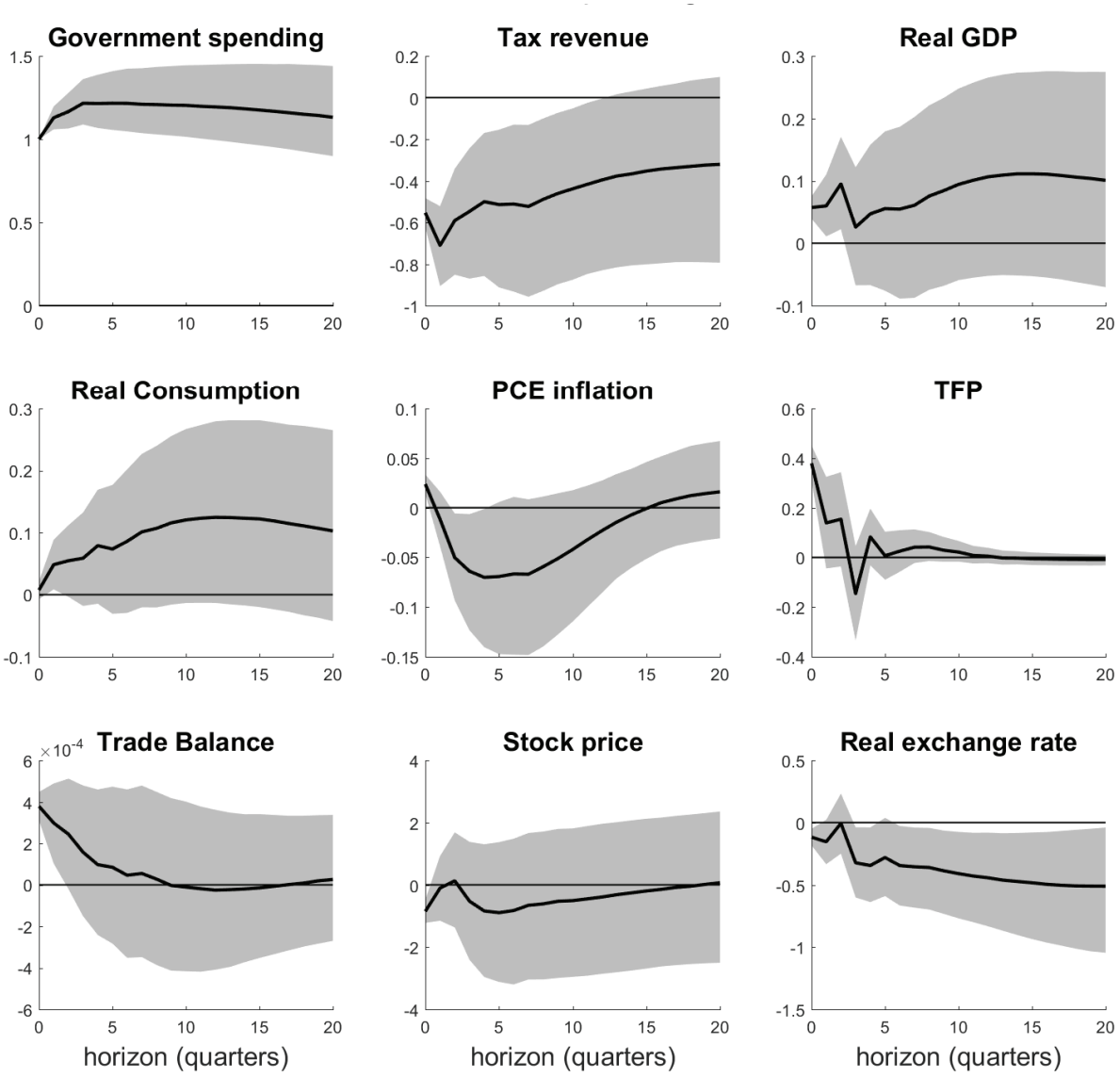

Figure 2: **Cholesky identification.** Impulse responses from a one standard deviation government spending shock. Target variables are tax revenues, Real GDP, real private consumption, PCE inflation, total factor productivity, trade balance, short-term interest rate and real effective exchange rate. The real effective exchange rate is defined as the weighted basket of foreign goods to domestic goods: a decrease stands for a depreciation. The impulse responses are obtained in a VAR framework with the spending shock identified through the Cholesky scheme. Shaded bands denote the 68% pointwise credible sets.

real exchange rate appreciates, inflation increases, trade balance deteriorates and consumption falls.

Dissecting the result, we notice that the real exchange rate appreciation is driven both by the response of inflation and by the nominal effective exchange rate (see Figure 4), in

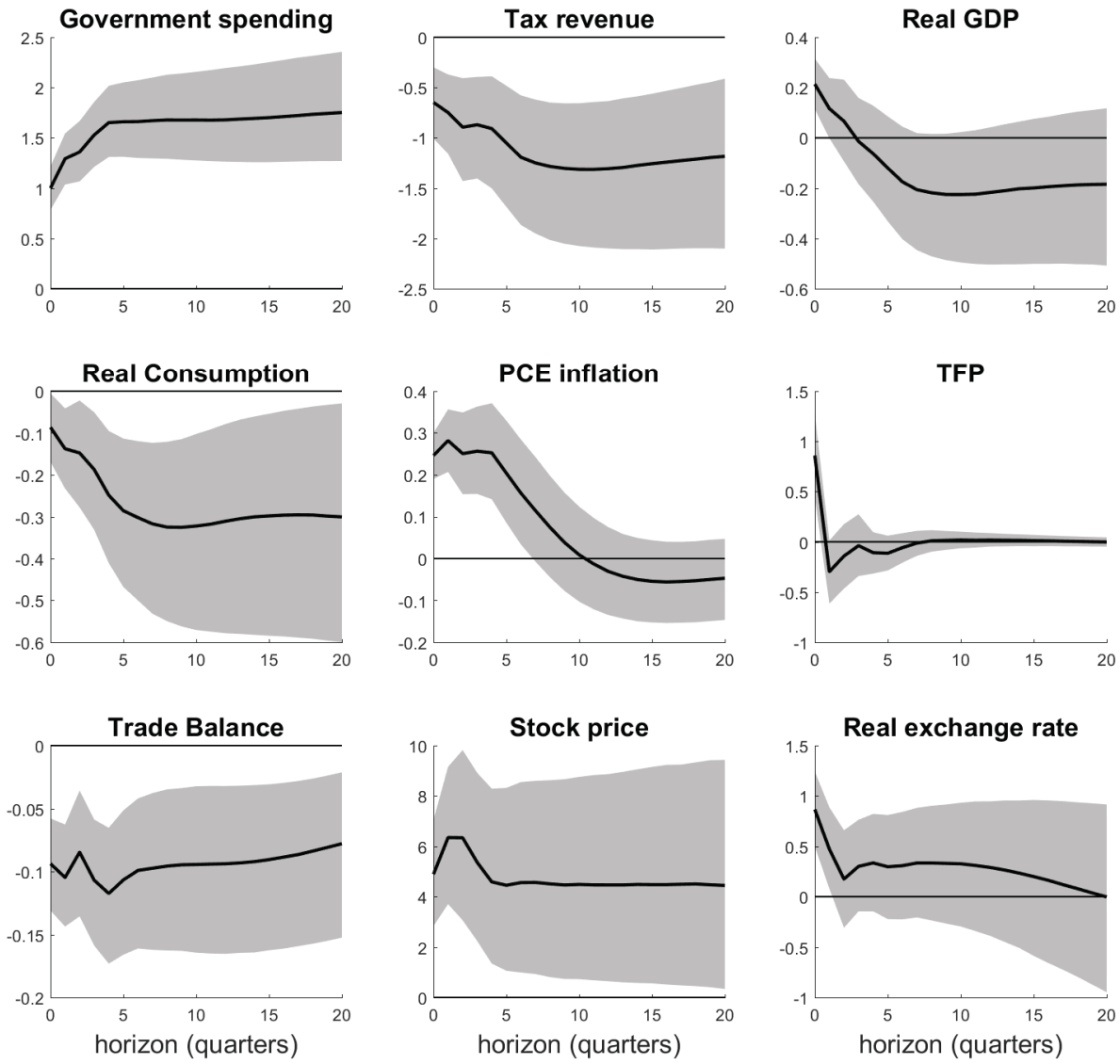

Figure 3: **Proxy-SVAR narrative identification.** Impulse responses from a one standard deviation government spending shock. Target variables are tax revenues, Real GDP, real private consumption, PCE inflation, total factor productivity, trade balance, short-term interest rate and real effective exchange rate. The real effective exchange rate is defined as the weighted basket of foreign goods to domestic goods: a decrease stands for a depreciation. The impulse responses are obtained in a proxy-SVAR framework in which government spending is instrumented with the military narrative series of Ramey (2016a). Shaded bands denote the 68% pointwise credible sets.

line with Mussa (1986). Inflation increases on impact and becomes not significant after a few quarters. The fall in trade balance supports the twin deficit hypothesis, coherently with the appreciated real exchange rate, and contrasts the alternative twin divergence hypothesis (Kim and Roubini, 2008). Consumption decreases, in line with Ramey (2011),

confirming the crowding-out effect due to the increase in the present value of taxes to be paid. TFP, in line with Jorgensen and Ravn (2018), augments. However, differently from Jorgensen and Ravn (2018), the increase in supply, due to the TFP increase, does not overcome the positive increase in demand from government spending and therefore prices increase.

The remaining variables show a standard behavior. Economic activity increases on impact, implying a fiscal multiplier slightly below 1, and then becomes insignificant. The stock price index of military firms also increases, confirming the non-anticipated component in the identified shock. Finally, tax revenues decrease.

In the next section we show that results are robust to various alternatives. Excluding the Bretton-Woods and the Great Recession period delivers similar results. Defining inflation as the consumer price index or focusing on the nominal exchange rate does not alter results. Also using a different narrative series to instrument the government spending shocks, in particular quarterly changes in government defense investment, confirms our results.

4 Robustness

This section reports additional evidence to support our baseline result, i.e. that the real exchange rate appreciates, inflation reacts positively and the trade balance deteriorates after a spending shock. We propose two additional sets of impulse responses. The first one is constructed using the same identification scheme of the baseline but changing samples length or variables specification. The second one uses defense government investment, instead of the narrative military series of Ramey, 2016a, to instrument surprise government spending shocks.

4.1 Other VAR specifications

Figure 4 shows the impulse-responses to our proxy-SVAR re-estimated on four different model specifications: 1) Excluding the Bretton-Woods period - 1976Q1-2015Q4. The sample now coincides with the period of the fully floating exchange rate regime and it has the advantage of being directly comparable with Kim and Roubini (2008); 2) Excluding the Great Recession - 1964Q1-2006Q4. This allows us to exclude the financial crisis and its, maybe, different behavior; 3) Including in the VAR the nominal effective exchange rate instead of the real one; 4) Replacing our Personal Consumption Consumer Index with a more standard Consumer Price Index to measure inflation. For conciseness, Figure 4

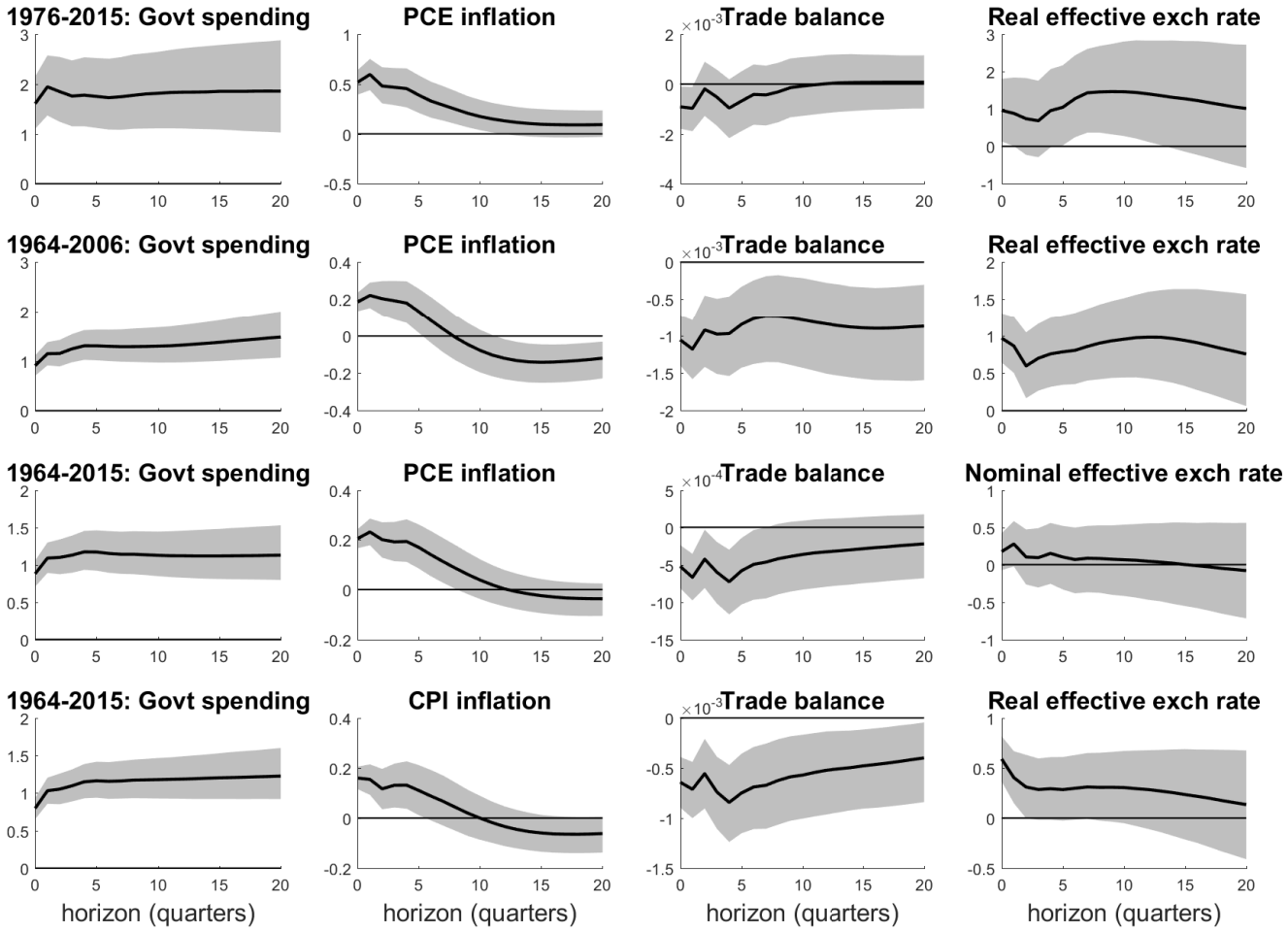

Figure 4: **Robustness using narrative shocks.** Impulse responses of government spending, inflation, trade balance and real effective exchange rate across different samples or variables specifications. The real effective exchange rate is defined as the weighted basket of foreign goods to domestic goods: a decrease stands for a depreciation. Line 1: fully flexible exchange rate sample (1976Q1-2015Q4). Line 2: pre-crisis sample (1964Q1-2006Q4). Line 3: full sample, nominal (instead of real) effective exchange rate. Line 4: full sample, CPI (instead of PCE) inflation. The estimation sample is 1964Q1-2015Q4. Shaded bands denote the 68% pointwise credible sets.

reports only the main variables of interest, i.e. the spending shock, the exchange rate, the trade balance and inflation. All other variables are available upon request.

The shock is inflationary, appreciates the real (or nominal) exchange rate and deteriorates the trade balance in all four specifications.

4.2 Government defense investment as an instrument

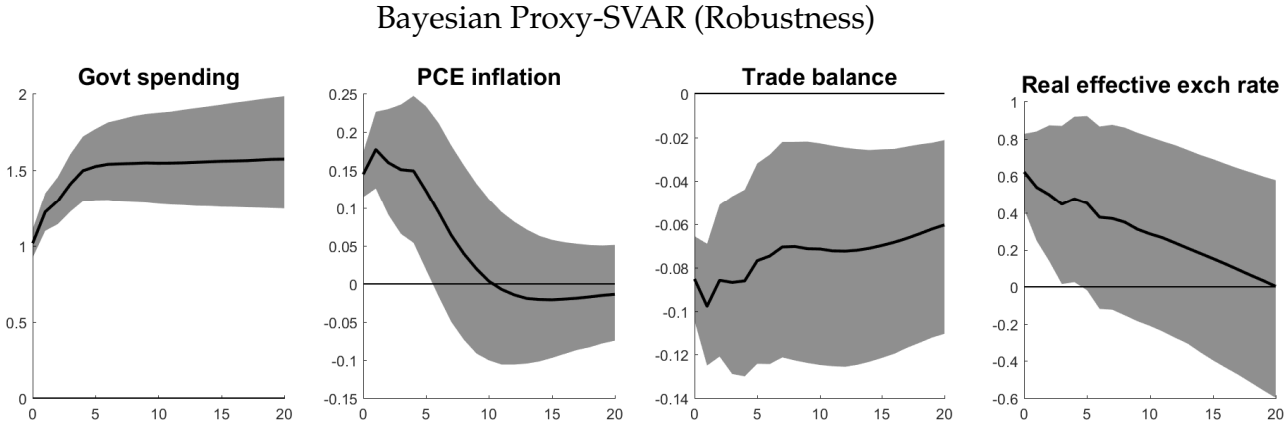

Figure 5: **Robustness using military instruments.** Impulse response functions constructed using defense investment as instrument for total government spending. The estimation sample is 1964Q1-2015Q4. The real effective exchange rate is defined as the weighted basket of foreign goods to domestic goods: a decrease stands for a depreciation. Shaded bands denote the 68% pointwise credible sets.

In order to be sure that our results are not only driven by the Ramey (2016a) narrative series, we analyse the impulse responses to the same shock identified through a different instrument: the quarterly changes in government defense investment (see Miyamoto et al., 2019). Being aware that government consumption and investment shock might have different characteristics (see Boehm, 2019), we need to test the relevance of this series as an instrument for overall spending. We perform a similar F-test to the one performed for the Ramey (2016a) narrative series. We obtain values of the F-statistics that are well above the thresholds, confirming the goodness of this additional instrument for surprise government spending shocks.[10] We then proceed, as before, estimating the Proxy-SVAR maintaining the same specification of the baseline model in terms of variables and estimation sample (1964Q1-2015Q4).

Figure 5 reports the results, confirming that the real exchange rate appreciates, inflation increases and the trade balance deteriorates.

5 Theory and Empirical results - solving the puzzles

To check the theoretical coherence of our empirical results, we build a two-goods RBC small open economy model. We then estimate its parameters to see how far we can go with a simple model in matching our empirical estimates. The results are quite striking:

[10] Results are available upon request.

the simplest model does a fair job in accounting for a broad range of macroeconomic responses to a government spending shock.

The model is the standard Small Open Economy Real Business Cycle model (see Mendoza, 1991 and Schmitt-Grohé and Uribe, 2017) enriched with multiple goods, as in Galí and Monacelli (2005), and a utility specification accounting for different degrees of the wealth effect of government spending, as in Jaimovich and Rebelo (2009). The economy is small, does not affect world prices and takes the world interest rate as given. It has three agents, household, firms and the government. International financial markets are incomplete and there are no nominal frictions. Households consume a composite of domestic and foreign goods, supply labor and save/borrow using a single internationally traded asset. They own the physical capital, rent it to firms and take investment decisions, which is subject to adjustment costs. Domestic firms produce a tradable good using capital and labor, selling it domestically and abroad. Movements in the terms of trade determine the competitiveness of the domestic sector, taking world demand as exogenous. The government purchases domestic goods raising funds through taxes, running a balanced fiscal budget (Monacelli and Perotti, 2010). The independence of the non-stochastic steady state from initial conditions is ensured through an endogenous discount factor, as in Schmitt-Grohé and Uribe (2003). Details on the standard model are available upon request.

Table 1. Estimated parameter values

Parameter	Value	Standard Error
Trade elasticity	0.694	0.096
Capital adjustment cost	0.714	0.032
Wealth Elasticity	0.894	0.1
AR 1	1.41	0.081
AR 2	-0.416	0.1
Inter-temporal elasticity of substitution	0.644	0.042
Home bias in consumption	0.798	0.012

To talk to the existing literature trying to reconcile theory and empirical findings and to see how far our simplest possible framework can go to account for the empirical findings, we estimate seven crucial parameters:[11] (1) the trade elasticity - governing (often together with the persistency of shocks) wealth effects determining the response of the households' demand and therefore relative prices; (2) capital adjustment cost - hindering the evolution of capital, affecting the correlation of macro variables' responses and of the

[11] We tried also to estimate the inverse of the Frisch elasticity and habits in consumption but these parameters are not identified using the IRF-matching procedure.

trade balance; (3) wealth elasticity of the labor supply - controlling the elasticity of the labor supply to wealth movements, setting the crowding out of government spending shocks; (4-5) Second order autoregressive process - shaping the response of government spending to its shock; (6) Inter-temporal elasticity of substitution - defining the inter-temporal behavior of consumption; (7) home bias in consumption - setting the share of domestic goods consumed in the basket of the households.

Model and empirical responses to a 1% increase in government spending (% deviations)

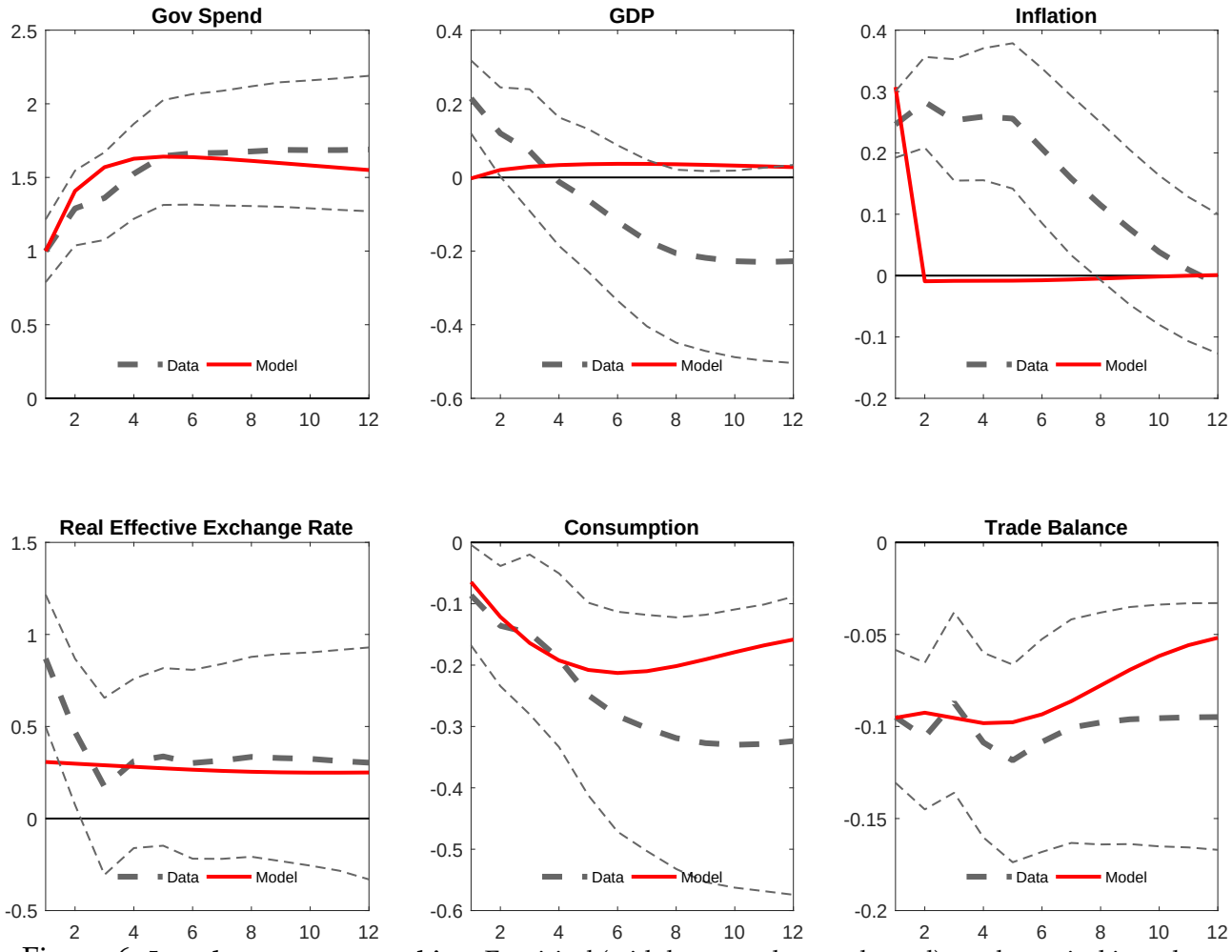

Figure 6: **Impulse-response matching.** Empirical (with lower and upper bound) vs. theoretical impulse-responses to one standard deviation shock to an unanticipated (unproductive) government spending shock.

We estimate these parameters by matching the impulse-responses of six variables: government spending, GDP, inflation, real exchange rate, trade balance and consumption. The estimated values of the parameters and their standard errors are reported in Table 5.[12]

[12]Standard errors are computed using Altig et al. (2011) procedure.

Three are the main findings of the impulse-matching procedure. First, the trade elasticity is low and consistently with the international macro literature, is below one. Second, households' preferences display a quite large wealth effect in order to be consistent with the crowding out of consumption. Third, the labor elasticity, habits in consumption and the home bias are not the most relevant parameters driving the responses of the model to a government spending shock.

Figure 6 compares the impulse-responses of the estimated model with the empirical ones. The model matches, on impact, all the signs of the responses and for all, with the exception of inflation, quite well also the dynamics. An increase in government spending is inflationary, appreciates the real exchange rate and, while increasing aggregate output, generates a current fall in aggregate consumption. Focusing on inflation, the model is unable to explain a persistent inflation dynamics. It is important here to remember that this model features flexible prices and the absence of nominal rigidities is such that prices adjust immediately and almost once for all.

6 Conclusions

Starting from an extensive closed and open economy literature showing puzzling effects of government spending shock on inflation, real exchange rate, trade balance and consumption, we show that properly identified unanticipated changes in government spending have effects in line with standard theoretical results. In practice, the paper reinvestigates the effects of government spending shocks embedding a narrative approach in a proxy-SVAR framework. It does that by showing that the Ramey (2016a) military instrument, when used after the Korean War, is still a valid instrument, but only for unanticipated shocks. We find that an increase in government spending appreciates the real exchange rate, increases inflation, induces a trade balance deficit and generates a fall in consumption. These results are consistent with a simple two-good standard small open economy RBC model; this simple model estimated to match empirical impulse responses does a good job in explaining the sign and dynamics of macro responses to a surprise government spending shock.

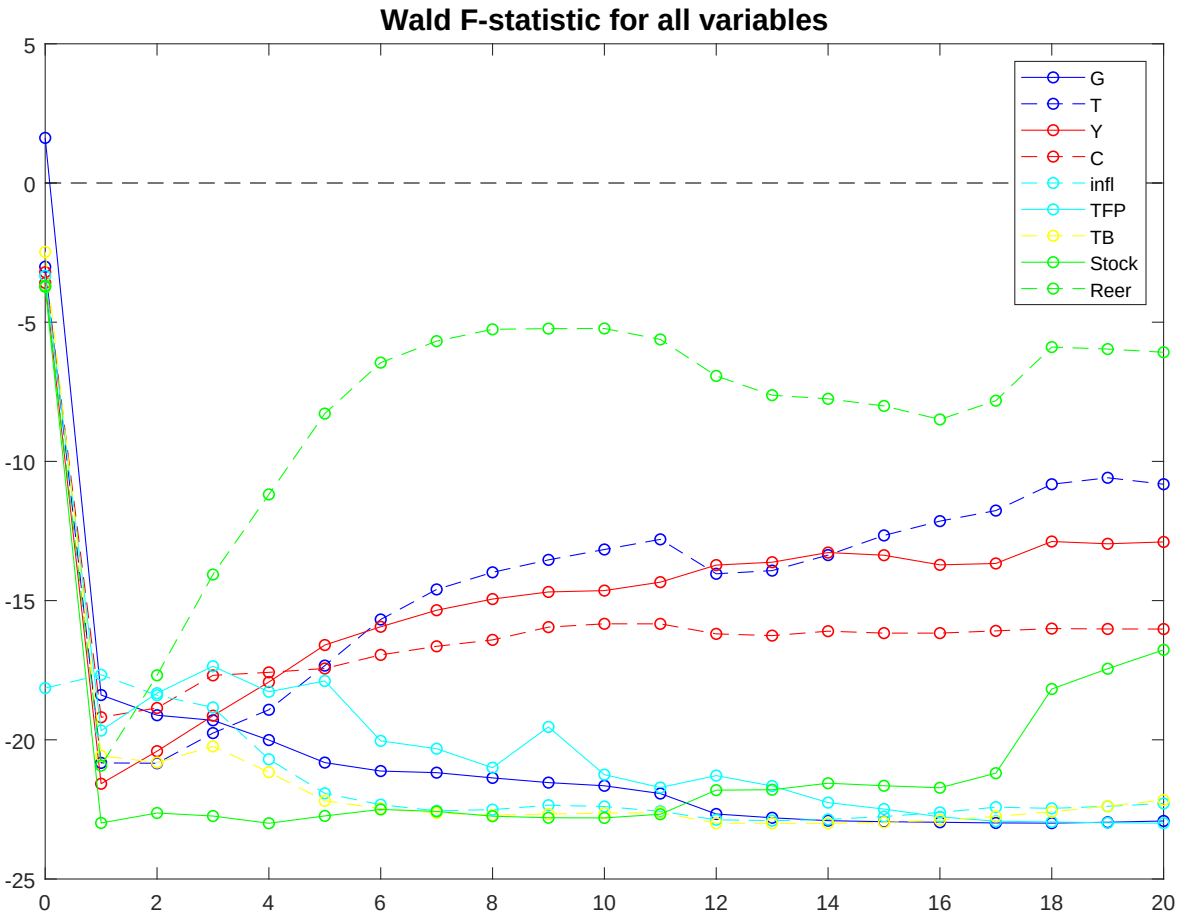

Figure 7: **F-statistics of the tests conducted on all variables of our VAR specification.**

Appendix

A Test the narrative series as an instrument

In this section we report the results of the F-tests of our military series on all variables in our VAR. The results are displayed in Figure 7.

References

Altig, D., Christiano, L. J., Eichenbaum, M., Lindé, J., 2011. Firm-specific capital, nominal rigidities and the business cycle. Review of Economic Dynamics 14 (2), 225 – 247.

Angelini, G., Caggiano, G., Castelnuovo, E., Fanelli, L., Jun. 2019. Fiscal Multipliers: Playing with Multiple Instruments. working paper, mimeo.

Auerbach, A. J., Gorodnichenko, Y., May 2013. Output Spillovers from Fiscal Policy. American Economic Review 103 (3), 141–146.

Auerbach, A. J., Gorodnichenko, Y., May 2016. Effects of Fiscal Shocks in a Globalized World. IMF Economic Review 64 (1), 177–215.

Basu, S., S. Kimball, M., 01 2003. Investment planning costs and the effects of fiscal and monetary policy. mimeo.

Baxter, M., King, R. G., June 1993. Fiscal Policy in General Equilibrium. American Economic Review 83 (3), 315–334.

Blanchard, O., Perotti, R., 2002. An Empirical Characterization of the Dynamic Effects of Changes in Government Spending and Taxes on Output. The Quarterly Journal of Economics 117 (4), 1329–1368.

Boehm, C. E., 2019. Government consumption and investment: Does the composition of purchases affect the multiplier? Journal of Monetary Economics.

Born, B., D'Ascanio, F., Müller, G. J., Pfeifer, J., Feb. 2019. Government Spending, Downward Nominal Wage Rigidity, and Exchange Rate Dynamics. working paper, mimeo.

Born, B., Juessen, F., Müller, G. J., 2013. Exchange rate regimes and fiscal multipliers. Journal of Economic Dynamics and Control 37 (2), 446–465.

Caldara, D., Kamps, C., 2017. The Analytics of SVARs: A Unified Framework to Measure Fiscal Multipliers. Review of Economic Studies 84 (3), 1015–1040.

Canova, F., Sahneh, M. H., 2018. Are Small-Scale SVARs Useful for Business Cycle Analysis? Revisiting Nonfundamentalness. Journal of the European Economic Association 16 (4), 1069–1093.

Corsetti, G., Dedola, L., Leduc, S., 04 2008. International risk sharing and the transmission of productivity shocks. Review of Economic Studies 75 (2), 443–473.

Corsetti, G., Kuester, K., Muller, G. J., 2011. Floats, pegs and the transmission of fiscal policy. Tech. rep.

Corsetti, G., Meier, A., Muller, G., Nov. 2009. Cross-border spillovers from fiscal stimulus. CEPR Discussion Papers 7535, C.E.P.R. Discussion Papers.

Corsetti, G., Muller, G. J., March 2013. Multilateral Economic Cooperation and the International Transmission of Fiscal Policy. In: Globalization in an Age of Crisis: Multilateral Economic Cooperation in the Twenty-First Century. NBER Chapters. National Bureau of Economic Research, Inc, pp. 257–297.

D'Alessandro, A., Fella, G., Melosi, L., Jul. 2019. Fiscal Stimulus with Learning-By-Doing. International Economic Review (826).
URL https://ideas.repec.org/p/qmw/qmwecw/826.html

Del Negro, M., Schorfheide, F., 2011. Bayesian Macroeconometrics. Handbook of Bayesian Econometrics Geweke.

Edelberg, W., Eichenbaum, M., Fisher, J. D., January 1999. Understanding the Effects of a Shock to Government Purchases. Review of Economic Dynamics 2 (1), 166–206.

Enders, Z., Muller, G. J., Scholl, A., January 2011. How do fiscal and technology shocks affect real exchange rates?: New evidence for the United States. Journal of International Economics 83 (1), 53–69.

Faccini, R., Mumtaz, H., Surico, P., 2016. International fiscal spillovers. Journal of International Economics 99, 31 – 45.

Fatás, A., Mihov, I., Apr. 2001. The Effects of Fiscal Policy on Consumption and Employment: Theory and Evidence. CEPR Discussion Papers 2760, C.E.P.R. Discussion Papers.
URL https://ideas.repec.org/p/cpr/ceprdp/2760.html

Fernald, J., 2012. A quarterly, utilization-adjusted series on total factor productivity. Federal Reserve Bank of San Francisco Working Paper Series 09.

Forni, M., Gambetti, L., 2014. Sufficient information in structural VARs. Journal of Monetary Economics 66 (C), 124–136.
URL https://ideas.repec.org/a/eee/moneco/v66y2014icp124-136.html

Forni, M., Gambetti, L., 2016. Government spending shocks in open economy VARs. Journal of International Economics 99 (C), 68–84.

Galí, J., López-Salido, J. D., Vallés, J., March 2007. Understanding the Effects of Government Spending on Consumption. Journal of the European Economic Association 5 (1), 227–270.

Galí, J., Monacelli, T., 07 2005. Monetary policy and exchange rate volatility in a small open economy. Review of Economic Studies 72 (3), 707–734.

Ilzetzki, E., Jin, K., Feb. 2013. The Puzzling Change in the International Transmission of US Macroeconomic Policy Shocks. working paper, mimeo.

Ilzetzki, E., Mendoza, E. G., Vegh, C. A., 2013. How big (small?) are fiscal multipliers? Journal of Monetary Economics 60 (2), 239–254.

Ilzetzki, E., Reinhart, C. M., Rogoff, K. S., Feb. 2017. The Country Chronologies to Exchange Rate Arrangements into the 21st Century: Will the Anchor Currency Hold? NBER Working Papers 23135, National Bureau of Economic Research, Inc.

Jaimovich, N., Rebelo, S., September 2009. Can news about the future drive the business cycle? American Economic Review 99 (4), 1097–1118.

Jorgensen, P. L., Ravn, S. H., 2018. The Inflation Response to Government Spending Shocks: A Fiscal Price Puzzle? mimeo.

Kim, S., 2015. Country characteristics and the effects of government consumption shocks on the current account and real exchange rate. Journal of International Economics 97 (2), 436–447.

Kim, S., Roubini, N., March 2008. Twin deficit or twin divergence? Fiscal policy, current account, and real exchange rate in the U.S. Journal of International Economics 74 (2), 362–383.

Kim, S.-H., Moon, S., Velasco, C., 2017. Delayed overshooting: Is it an 80s puzzle? Journal of Political Economy 125 (5), 1570–1598.

Lambertini, L., Proebsting, C., Feb. 2019. Fiscal Policy, Relative Prices and Net Exports in a Currency Union. working paper, mimeo.

Linnemann, L., October 2006. The Effect of Government Spending on Private Consumption: A Puzzle? Journal of Money, Credit and Banking 38 (7), 1715–1735.

Lunsford, K. G., Dec. 2016. Identifying Structural VARs with a Proxy Variable and a Test for a Weak Proxy. Working Papers (Old Series) 1528, Federal Reserve Bank of Cleveland.

McCracken, M. W., Ng, S., October 2016. FRED-MD: A Monthly Database for Macroeconomic Research. Journal of Business & Economic Statistics 34 (4), 574–589.

Mendoza, E. G., September 1991. Real business cycles in a small open economy. American Economic Review 81 (4), 797–818.

Mertens, K., Ravn, M. O., 2013. The dynamic effects of personal and corporate income tax changes in the united states. American Economic Review 103(4), 1212–1247.

Miyamoto, W., Nguyen, T. L., Sheremirov, V., 2019. The effects of government spending on real exchange rates: Evidence from military spending panel data. Journal of International Economics 116, 144 – 157.
URL http://www.sciencedirect.com/science/article/pii/S0022199618304458

Monacelli, T., Perotti, R., 04-05 2008. Openness and the Sectoral Effects of Fiscal Policy. Journal of the European Economic Association 6 (2-3), 395–403.
URL https://ideas.repec.org/a/tpr/jeurec/v6y2008i2-3p395-403.html

Monacelli, T., Perotti, R., 05 2010. Fiscal Policy, the Real Exchange Rate and Traded Goods. Economic Journal 120 (544), 437–461.

Montiel Olea, J. L., Pflueger, C., 2013. A robust test for weak instruments. Journal of Business & Economic Statistics 31 (3), 358–369.

Montiel Olea, J. L., Stock, J. H., Watson, M. W., 2018. Inference in structural vector autoregressions identified with an external instrument. mimeo.

Mountford, A., Uhlig, H., 2009. What are the effects of fiscal policy shocks? Journal of Applied Econometrics 24 (6), 960–992.

Mussa, M., January 1986. Nominal exchange rate regimes and the behavior of real exchange rates: Evidence and implications. Carnegie-Rochester Conference Series on Public Policy 25 (1), 117–214.
URL https://ideas.repec.org/a/eee/crcspp/v25y1986ip117-214.html

Nakamura, E., Steinsson, J., March 2014. Fiscal Stimulus in a Monetary Union: Evidence from US Regions. American Economic Review 104 (3), 753–792.

Perotti, R., 2005. Estimating the effects of fiscal policy in OECD countries. Proceedings.

Ramey, V. A., 2011. Identifying Government Spending Shocks: It's all in the Timing. The Quarterly Journal of Economics 126 (1), 1–50.

Ramey, V. A., 2016a. Defense news shocks, 1889-2015: Estimates based on news sources. Unpublished manuscript.

Ramey, V. A., Feb. 2016b. Macroeconomic Shocks and Their Propagation. NBER Working Papers 21978, National Bureau of Economic Research, Inc.

Ramey, V. A., Zubairy, S., 2018. Government spending multipliers in good times and in bad: Evidence from us historical data. Journal of Political Economy 126 (2), 850–901.

Ravn, M., Schmitt-Grohé, S., Uribe, M., 2006. Deep Habits. Review of Economic Studies 73 (1), 195–218.

Ravn, M. O., Schmitt-Grohe, S., Uribe, M., 2012. Consumption, government spending, and the real exchange rate. Journal of Monetary Economics 59 (3), 215–234.

Schmitt-Grohé, S., Uribe, M., October 2003. Closing small open economy models. Journal of International Economics 61 (1), 163–185.

Schmitt-Grohé, S., Uribe, M., 2017. Open Economy Macroeconomics. Princeton Press.

Stock, J. H., Watson, M. W., 2008. Lecture 7: Recent Developments in Structural VAR Modeling. Presented at the National Bureau of Economic Research Summer Institute Minicourse: What's New in Econometrics: Time Series, Cambridge, MA.

Stock, J. H., Watson, M. W., 2018. Identification and estimation of dynamic causal effects in macroeconomics using external instruments. The Economic Journal 128 (610), 917–948.

Zeev, N. B., Pappa, E., August 2017. Chronicle of a War Foretold: The Macroeconomic Effects of Anticipated Defence Spending Shocks. Economic Journal 127 (603), 1568–1597.