SAME, BUT DIFFERENT: TESTING GOVERNMENT EXPENDITURE SHOCK MEASURES

A Thesis

by

ASHKAN HASSANI

Submitted to the Graduate and Professional School of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Chair of Committee,	Sarah Zubairy
Committee Members,	Tatevik Sekhposyan
	Mark Lawley
Head of Department,	Steve Puller

December 2021

Major Subject: Economics

ABSTRACT

In this study we test whether different government expenditure innovation measures proposed in the literature are good proxies for government spending shocks. We compare five government spending shock measures: forecast error of survey of professional forecasters (SPF), forecast error of forecast of Federal Reserve published in Greenbook (GB), military news, Fisher-Peters measure, Ben Zeev-Pappa measure. To compare these measures, we utilize the reliability estimator of a proxy, R^2, and F statistics which are proposed by [1]. We find that SPF and GB constitute the best proxy variable for government spending shocks with regard to different reliability estimators. SPF and military news proxies have the highest multipliers comparing to the others. Their both multipliers are greater than 1. Multiplier for Ben Zeev-Pappa and Fisher-Peters measures are less than zero. GB also has a positive multiplier but less than one.

DEDICATION

I dedicate this thesis to the memory of my loving baby, Ali, whom we lost while working on this research, and to his little brother, Amirali, who gave a new meaning to our life. I will carry your love forever in my heart.

ACKNOWLEDGMENTS

I would like to thank my wife, Maryam, for her incredible support during my difficult times. I would not have finished this without her. I would also like to thank my parents, Helen and Jahangir, for their unconditional and endless support, love, and encouragement.

Foremost, I would like to express my sincere gratitude to my advisor Dr. Sarah Zubairy for being an excellent advisor and for her continuous support, understanding, and encouragement during this process. I appreciate the countless hours she spent educating me and for always believing in me. I would like to thank my committee members Dr. Sekhposyan, and Dr. Lawley for their time and expertise.

TABLE OF CONTENTS

LIST OF FIGURES

FIGURE Page

LIST OF TABLES

TABLE Page

1. INTRODUCTION

A classic question in macroeconomics is: how does the economy respond to a change in government spending? Surprisingly, there is little consensus on the answer to this question. It derives from the fact that different approaches to identify government spending shocks lead to different answers. All approaches have some shortcomings. By government spending, we mean government purchases, i.e. G in the NIPA identity.

There are many government spending shocks measures proposed in the literature. We compare the five most important ones. These measures are: a) Military news shock measure proposed by [2], b) forecast error of Survey of Professional Forecasters (SPF) published by the Philadelphia Federal Reserve, c) defense spending shocks proposed by [3], d) excess returns of large US military contractors proposed by [4], and e) forecast error of Forecast of Federal Reserve provided by Philadelphia Fed's Greenbook (GB).

Interestingly, as Figure 1 shows, while all studies measure the same structural shock, these measures are very different from each other.

Typically, the literature on government spending has sought to address two main questions: (1) Which identification scheme is the preferred methodology for identify government spending shocks? (2) What are the government spending multipliers? We try to answer these questions in this paper by comparing the aforementioned proxies. In order to do so, our analysis is based on a reliability measure which is developed by [1], R^2, and F statistics. Using these tests, we find that SPF and GB forecast errors are the best government spending shock measures. In our analysis, we employ the Proxy-SVAR model, which has been developed by [5] and [1].

The rest of this study is organized as follows. Chapter 2 reviews different government spending shock measures. Chapter 3 test the strength of various proxies proposed in the literature for government spending shocks. Chapter 4 presents the results, and Chapter 5 concludes and provides future research direction.

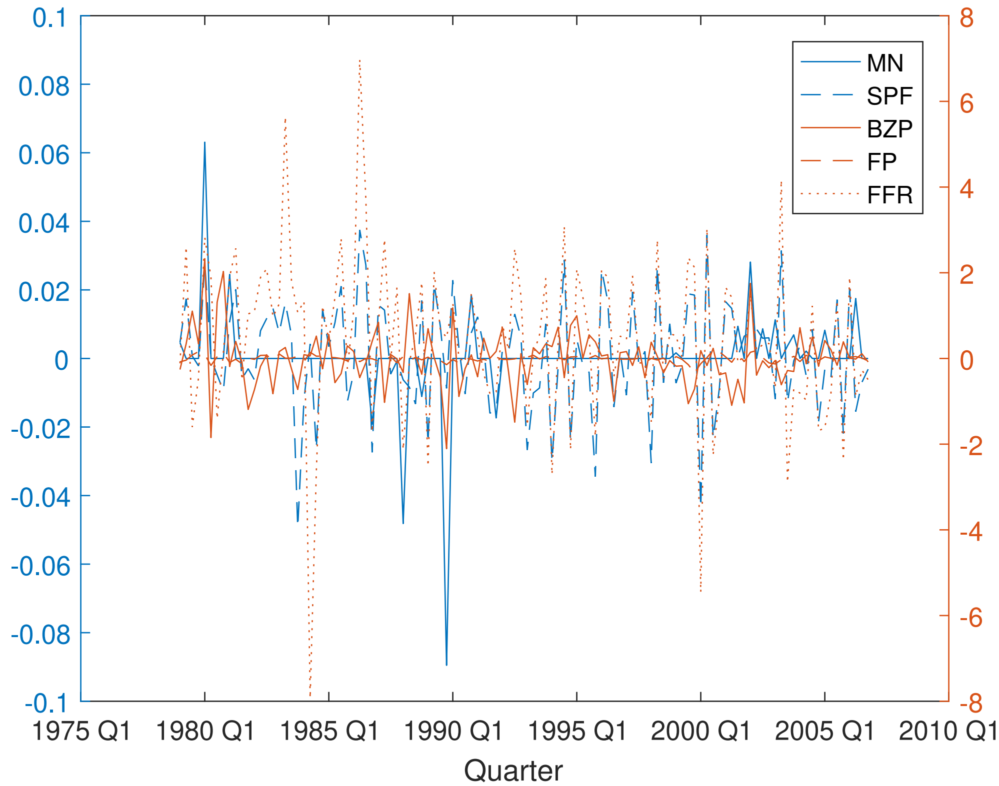


Figure 1.1: Comparison of different government spending shocks measures

2. GOVERNMENT SPENDING SHOCK MEASURES

The five main shock measures that we study in this paper are i) narrative military news shock, updated in [6]; (ii) forecast errors of Survey of Professional Forecasters (SPF) published by the Philadelphia Fed (iii) forecast error of Forecast of Federal Reserve provided by Philadelphia Fed's Greenbook; (iv) Ben Zeev and Pappa's defense news shock identified using [3] medium run horizon method; and (v) Fisher and Peters's military contractor excess returns in [4].

(i) Military news:

[2] used narrative approach to create a dummy variable capturing major military buildups. They read through Business Week to find the political events that led to the buildups. They were looking for a series that was exogenous to the current state of the economy. They used this narrative approach to ensure that the shock was unanticipated. They argue that there are three dates when persistent increases in US military spending were anticipated. By viewing the world events precipitating the military spending as exogenous, these dates are used to identify government spending shocks. The principal advantage of the narrative approach is the exogeneity of the spending episodes. This approach has some shortcomings: a) the small number of observations, that the episodes only involve spending increases, b) it assumes that the spending is known with certainty, and c) the selection of the dates is inherently subjective. These war dates correspond to the Korean War (1950:3), the escalation of the Vietnam War (1965:1), and the Carter-Reagan military build-up (1980:1).

[7] sought to reconcile why the war dates were producing different results from the SVARs that used Cholesky decompositions. Ramey argued that most government spending is anticipated at least several quarters in advance, so that the standard SVAR method was not identifying unanticipated shocks. In support of this idea, she showed that the shocks from an SVAR were indeed Granger-caused by the [2] war dates. To create a richer narrative variable to capture the "news" part of government spending shocks, she read Business Week

starting in 1939 and created a quantitative series of estimates of changes in the expected present value of government purchases, caused by military events. She then embedded the news series in a standard VAR, with the news ordered first in the Cholesky decomposition. In that work, she found results that were broadly consistent with the estimates based on the simple war dates. In follow-up work, [8] and [6] extended the military news series back to 1889. The military news variable tends to have low instrument relevance for samples that begin after the Korean War, though. I use this military news as one of the measures that I test.

(ii) Survey of Professional Forecasters (SPF):

The agents' expectations about future government spending growth are reported in Survey of Professional Forecasters (SPF) published by the Philadelphia Federal Reserve. The SPF reports forecasts for an ample set of macroeconomic and financial variables on a quarterly basis. This includes projections on government consumption and investment. While the Ramey shocks are based only on military expenditure, SPF uses projections for the largest available aggregate for government spending, comprising government consumption and investment at the federal, state and local level. Rather than only considering military expenses, what matters are projections on overall spending. Moreover, the SPF comprises forecasts for the aggregates of government consumption and investment, but not forecasts for subcategories such as military spending. In the SPF, forecasts for government consumption and investments are expressed in real terms and seasonally adjusted at annual rates. Following Ramey, we use the forecast errors of SPF to do the second analysis.

(iii) Forecast of Federal Reserve published in Greenbook data set (GB):

The Greenbook is produced before each meeting of the Federal Open Market Committee. The research staff at the Board of Governors prepares projections about how the economy will fare in the future. These projections are made available to the public after five years. This data set contains Greenbook projections for many of the variables also forecast in the Federal Reserve Bank of Philadelphia's SPF. The variables include real GDP and its components, nominal GDP, five measures of inflation, unemployment, industrial production, and housing starts. Altogether, the Philadelphia Fed's Greenbook Data Set includes Greenbook projections for 15 variables. Greenbook also forecasts government spending and we use forecast errors, similarly to SPF as another measure.

(iv) Ben Zeev and Pappa (BZP):

They use the medium-horizon identification methods of [3] to identify news shocks to defense spending from a time series model. In particular, [9] identify defense spending news as a shock that (i) is orthogonal to current defense spending; and (ii) best explains future movements in defense spending over a horizon of five years. This shock best explains future movements in defense spending over a horizon of five years.

(v) Fisher-Peters (FP):

They identify government spending shocks with statistical innovations to the excess returns of large US military contractors. This approach is used to estimate the dynamic responses of output, hours, consumption and real wages to a government spending shock. They find that positive government spending shocks are associated with increases in output, hours and consumption. Real wages initially decline and then rise after a year. [4] proposed an alternative series of news based on the excess returns of defense contractor stocks for the period starting in 1958.

3. METHODOLOGY

In this section we describe the data set, how we set up the empirical model, reliability measure and the test statistics.

3.1 The Proxy-SVAR Model

Throughout the paper we consider the time span for which all five series are available. In our analysis, we employ the Proxy-SVAR model, which has been developed by [5] and [1]. A useful interpretation of the proxy variables is as imperfect measurements of latent structural shocks. This interpretation makes sense in applications where the proxies are specified as narratively identified monetary or fiscal policy changes. Narratives of economic policy are constructed from historical sources that are used to summarize information about the size, timing, and motivation of policy interventions. But historical records sometimes contradict each other and calls for judgment. Another common issue of narrative shock series is that many observations are assumed to zero. These measurement problems invalidate the use of the narratives as direct observations of structural shocks. This also avoids bias estimates in simple regressions.

The VAR model contains the endogenous variables [Govt spending, Taxes, GDP] and the proxy will be chosen from a set of proxies including military news variable, SPF and etc. When we run VAR with Choleski, then we will have [Proxy variable, Govt spending, Taxes, GDP].

Choleski method is the most commonly used identification method in macroeconomics, which imposes alternative sets of recursive zero restrictions on the contemporaneous coefficients. This method was introduced by Sims (1980a). The "external instrument," or "proxy SVAR," method is a promising new approach for incorporating external series for identification. This approach takes advantage of information developed from "outside" the VAR, such as series based on narrative evidence, shocks from estimated DSGE models, or highfrequency information. The idea is that these external series are noisy measures of the true shock.

Identifying structural shocks in vector autoregressions (VARs) is important for research in macroeconomics. Consider the $n \times 1$ vector of time series variables, denoted by Y_t, that follows

$$Y_t^{'} = A_0 + Y_{t-1}^{'} A_1 + Y_{t-p}^{'} A_p + u_t^{'}, \tag{3.1}$$

where u_t is the $n \times 1$ vector of VAR innovations. Let v_t be the $n \times 1$ vector of structural shocks, which are related to the VAR innovations by

$$u_t = B v_t, \tag{3.2}$$

The objective of the structural VAR literature is mostly to estimate the column of B that corresponds to the structural shock of interest. For ease of exposition, we assume that the relevant column of B is the first, B_1.

Without loss of generality, I order v_t so that its first element is the structural shock of interest. Then, Equation (2) is

$$u_t = [B_1 B_2] v_t, \tag{3.3}$$

$v_{1,t}$ is the shock of interest, and $v_{2,t}$ contains the other structural shocks. Here, the vector B_1 determines how $v_{1,t}$ impacts Y_t , and estimating this vector is the focus of this paper. With this in mind, I make the following common assumptions about the properties of the structural VAR model.

Assumption 1:

a) The lag order p is known and the VAR is stationary.

b) B is invertible.

c) $\mathbb{E}(v_t) = 0$

d) $\mathbb{E}(v_t v_t^{'}) = \Sigma_v$, where Σ_v is positive definite and

$$\Sigma_v = \begin{pmatrix} \sigma_{v_1}^2 & 0 \\ 0 & \sigma_{v_2}^2 \end{pmatrix} \tag{3.4}$$

It is clear from the above that $\mathbb{E}(v_t v_s^{'}) = 0$ for $t \neq s$

There exists a time series variable, denoted by z_t, which is used as a proxy for $v_{1,t}$. Accordingly, we make the following assumptions:

Assumption 2:

a) z_t has finite mean, denoted by μ_z.

b) z_t has finite mean, denoted by σ_z^2.

c) z_t is a relevant proxy for $v_{1,t}$

$$\mathbb{E}(v_{1,t} z_t) = \phi \neq 0, \text{with } \phi \text{ finite} \tag{3.5}$$

d) z_t is exogenous from the structural shocks $v_{2,t}$.

$$\mathbb{E}(v_{2,t} z_t) = 0, \tag{3.6}$$

Assumptions 2.a and 2.b ensure that the proxy is well behaved. Assumptions 2.c and 2.d make it useful for identification.

3.2 The Reliability Measure

According to [1], an estimator of the reliability of the proxy is given by:

$$\Lambda = (\phi^2 \sum_{t=1}^{T} 1_T (\epsilon_t^T)^2 + \sum_{t=1}^{T} 1_T (m_t - \phi \epsilon_t^T)^2)^{-1} \phi^2 \sum_{t=1}^{T} 1_T (\epsilon_t^T)^2 \tag{3.7}$$

where 1_T is an indicator function for a nonzero observation of m_t. They show how an estimate of ϕ can be obtained using the restrictions implied by the estimated covariance matrix of the VAR residuals u_t.

The resulting reliability Λ lies between zero and one with larger values indicating a higher correlation between the proxy and the true underlying tax shock. The statistic is asymptotically equivalent to the R^2 statistic of the regression of the nonzero observations of the proxy variable against the corresponding structural tax shocks identified by the proxy VAR. In [10] to test for the presence of a proxy variable, he used the F statistic from the regression of the proxy variable onto the VAR errors.

4. RESULTS

In this section, we provide results for correlation between aforementioned proxies, impulse response functions (IRF) and mulitpliers for proxies. We also compare these proxies with regard to the reliability measure provided by [1], R^2, and F statistics.

4.1 Proxy Comparison

Figure 2 depicts the nonzero observations of the proxy variable against the corresponding structural government spending shocks identified by the aforementioned proxy VARs. These plots show that nonzero observations for SPF and GB are well spread over the regression line.

Table 1 presents the correlation between each pair of proxies.

Table 4.1: Correlation between proxies (1979-2006)

	Military	SPF	BZ&P	F&P	GB
Military	1				
SPF	0.1565	1			
Ben Zeev-Pappa	0.4053	-0.0919	1		
Fisher-Peters	0.1767	-0.1767	0.3804	1	
GB	0.1163	0.7272	-0.1638	-0.2233	1

As expected, SPF and GB have the highest correlation among the other pairs. The reason is that professional forecastors and Federal Reserve economists are both forecasting government spending in real time based on developments in economy.

In the following table, we summarize the reliability estimate, R^2 statistic, and F-statistics for different proxies for government spending shocks:

GB has the highest value for the reliability measure. However, with regard to other measures, SPF constitutes the best proxy variable for government spending shocks. Thus,

Table 4.2: Reliability and F-statistics for proxies for government spending shocks

	Reliability	R^2	F	Sample Period
SPF	0.32361	0.46727	42.395	1969-2006
GB	0.98543	0.27171	14.177	1979-2008
Military News	0.092332	0.14021	26.146	1890-2015
BZP	0.162	0.1	8.719	1948-2007
FP	0.01387	0.01559	1.1243	1947-2008

this table confirms that these two measures are similar and accordingly work better.

4.2 Multipliers under different proxies

Plots in Figure 3 show the normalized dynamic effect of a positive one standard deviation shock using aforementioned proxies. The 90% confidence intervals are the percentile intervals from a residual based moving block bootstrap algorithm described in [11] with 10,000 bootstrap replications. This shock causes an increase in GDP growth. IRF's for SPF and GB are very similar. Taxes and GDP start rising after the shock, but fall slowly thereafter. Government spending is also falling after the shock. IRF for military news is different from the aforementioned IRF's. Taxes and GDP will rise after period 0. For BZP, taxes and GDP on impact start with negative value but rises in the future periods. IRF for FP shows taxes increase on impact but falls in the future periods.

Figure 4 shows the multipliers for proxies. SPF and military news proxies have the highest multipliers comparing to the others.

For robustness check, we compare the IRF's for cases including consumption and investment. We add consumption and investment to the VAR model. The VAR model will be as: [Govt spending, Taxes, GDP, Consumption] and [Govt spending, Taxes, GDP, Consumption, Investment]. Figure 5 presents the results for the case with consumption and investment. SPF and GB IRF's are very similar for both data sets. Taxes and GDP are falling for the first 4-6 quarters, but they rise slowly thereafter. Government spending is falling except a rise at around the 4th quarter. Consumption rises smoothly, but investment falls in the beginning quarters,

We do the same exercise as in Table 2 for the overlapping sample for robustness check. The results for the overlapping period is almost the same. There is no significant difference between the results in Table 2 and Table 3. It confirms that SPF and GB are better proxies for government spending shocks.

Table 4.3: Reliability and F-statistics for proxies for government spending shocks

	Reliability	R^2	F	Sample Period
SPF	0.3588	0.5466	42.58	1979-2006
Fisher-Peters	0.04847	0.0493	1.831	1979-2006
Military News	0.0044	0.0678	0.8937	1979-2006
Ben Zeev-Pappa	0.1641	0.1205	4.839	1979-2006
GB	0.9889	0.2667	12.851	1979-2006

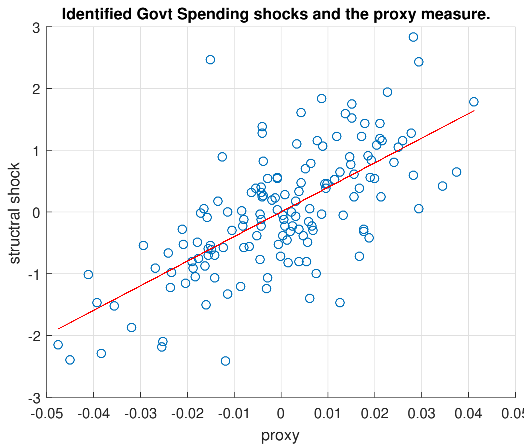
4.3 Considering SPF and GB forecast errors further

Table 4 presents the result of GDP multiplier for both SPF and GB proxies. The government expenditure multiplier is, thus, the ratio of change in income to a change in government spending. The reason behind this expansionary effect of government spending on income is that the increase in public expenditure constitutes an increase in income, thereby triggering successive increases in consumption, which also constitutes increase in income. For each of them, GDP multiplier has been calculated in three cases of with consumption, with investment, and with both consumption and investment in the VAR.

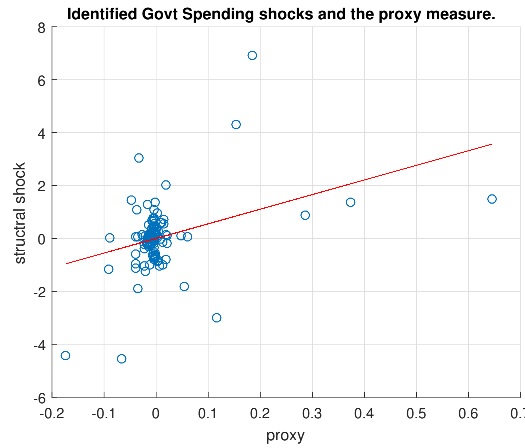
According to the Table 4, the government spending multiplier is larger for SPF, which means that income is expected increase with this proxy. Figure 5 confirms this expectation. The multiplier is less than 1 means $1 increase in government spending leads to less than $1 increase in GDP but greater than 0. By comparing figures 4 and 5, I recognized that the IRF's for both SPF and GB have not been changed significantly by adding consumption and investment to the VAR, specifically for government spending.

Table 4.4: GDP Multiplier (consumption and investment are additional variables in the VARs)

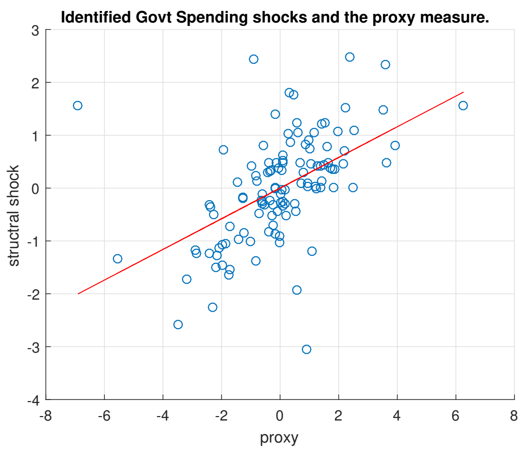
		Impact	4th	8th	12th
SPF	consumption-investment	0.7942	-0.086774	-0.584	-0.57772
	investment	0.89502	0.72147	0.69114	0.80566
	consumption	0.87044	-0.035091	-0.32752	-0.16826
GB	consumption-investment	0.20488	0.1506	0.21692	0.24562
	investment	0.48953	0.39791	0.46023	0.42652
	consumption	0.22152	0.038246	0.23559	0.27536



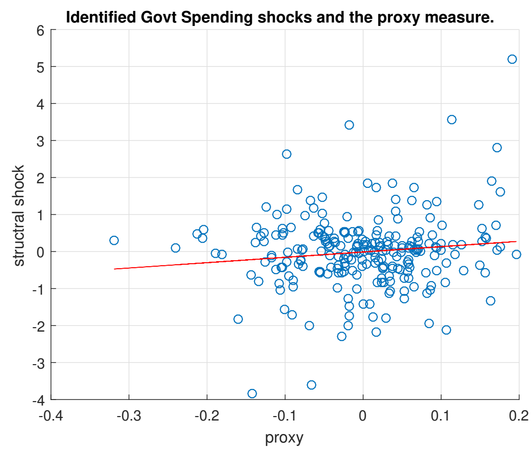
(a) SPF



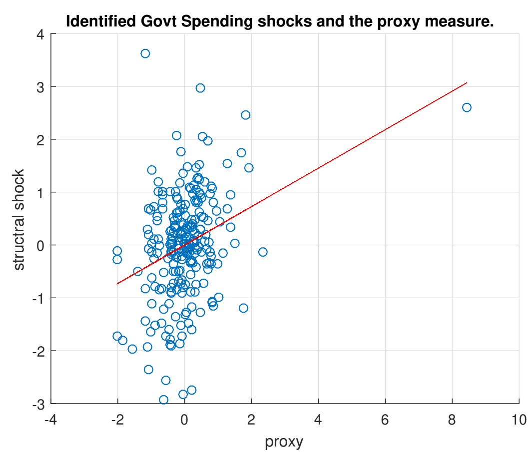
(b) Military News



(c) GB

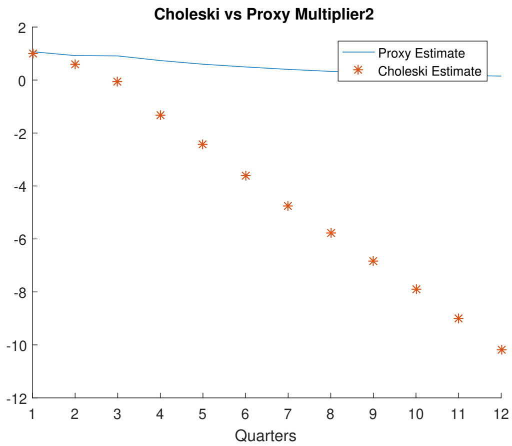


(d) FP

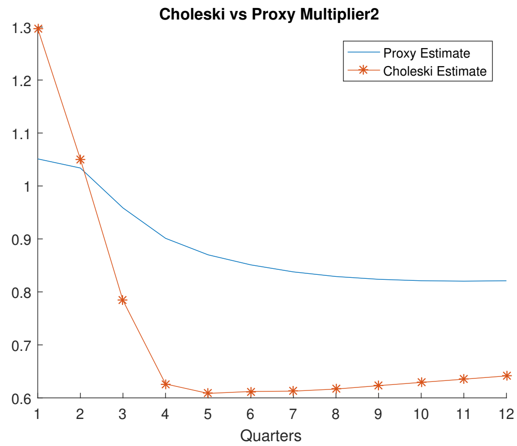


(e) BZP

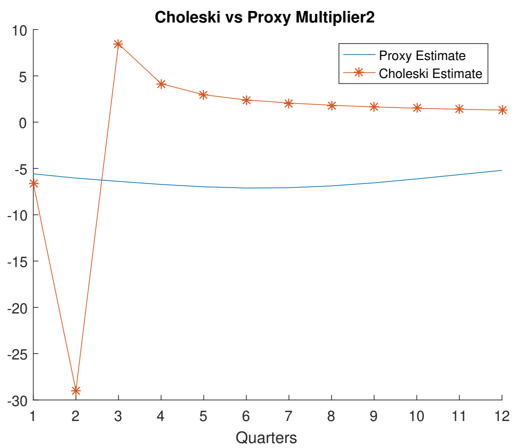
Figure 4.1: IRF's for proxies



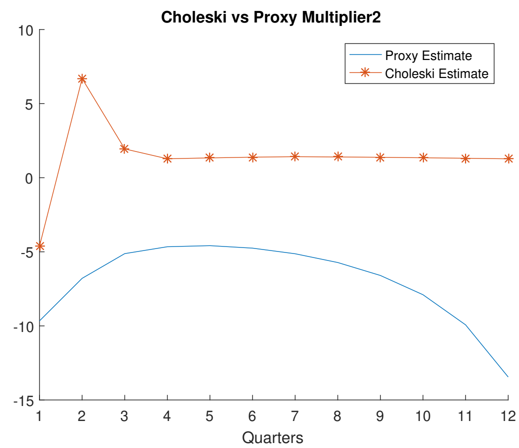
(a) SPF Multiplier



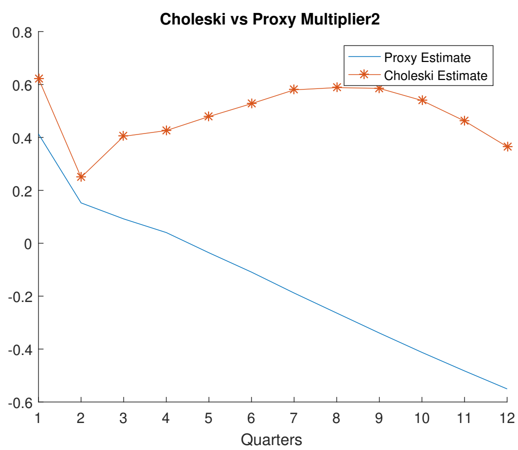
(b) Military News Multiplier



(c) BZ&P Multiplier

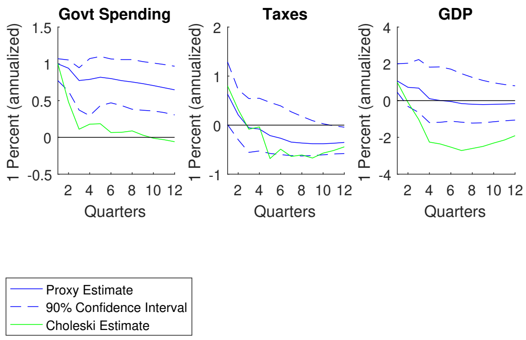


(d) F&P News Multiplier

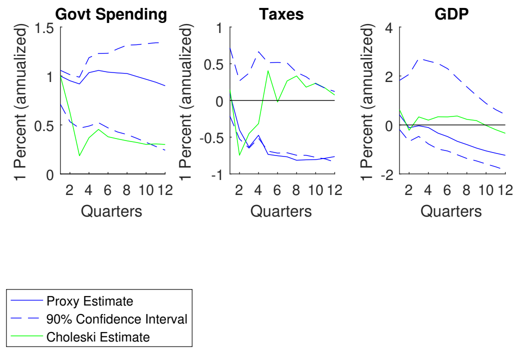


(e) GB Multiplier

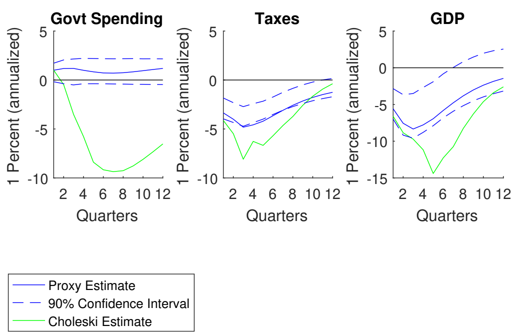
Figure 4.2: Multipliers



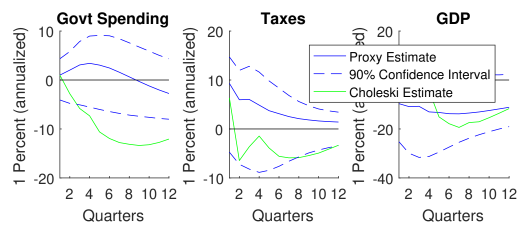
(a) Normalized IRF for SPF



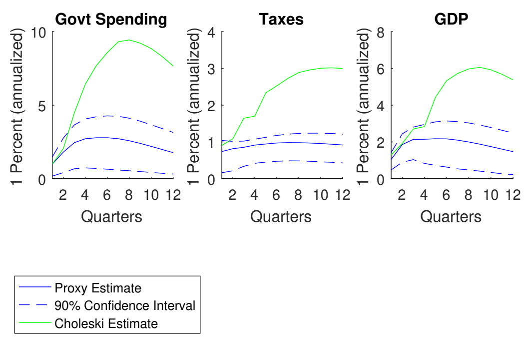
(b) Normalized IRF for GB



(c) Normalized IRF for BZP

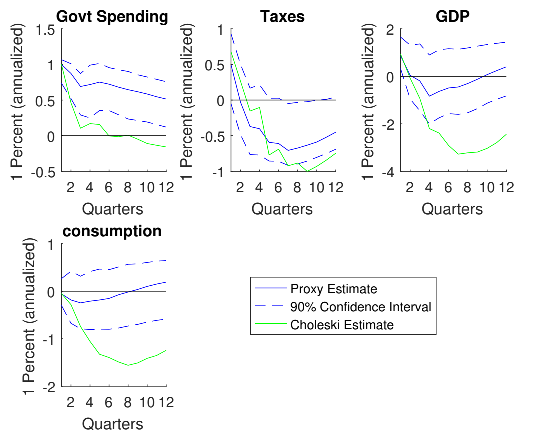


(d) Normalized IRF for FP

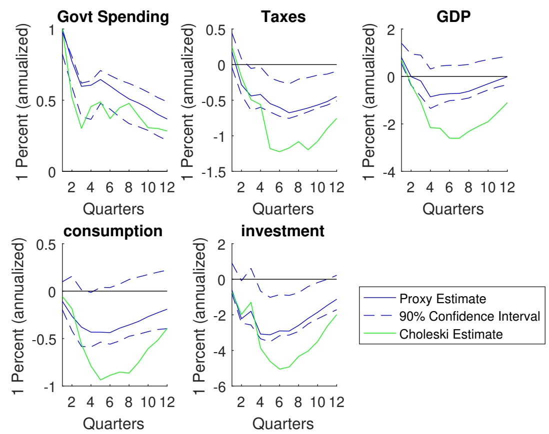


(e) Normalized IRF for Military News

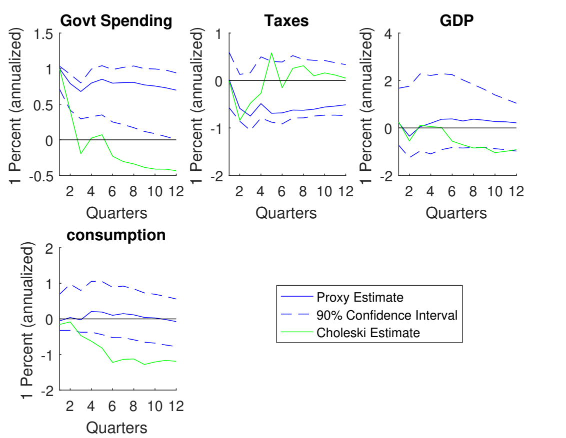
Figure 4.3: IRF's for proxies



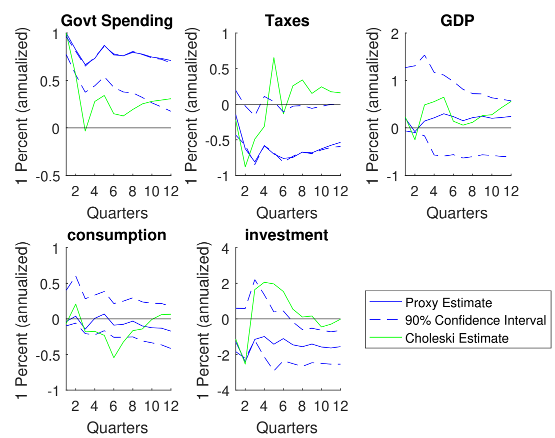
(a) Normalized IRF for SPF with consumption



(b) Normalized IRF for SPF with investment



(c) Normalized IRF for GB with consumption



(d) Normalized IRF for GB with consumption and investment

Figure 4.4: IRF with Consumption and Investment

5. CONCLUSIONS AND FUTURE RESEARCH

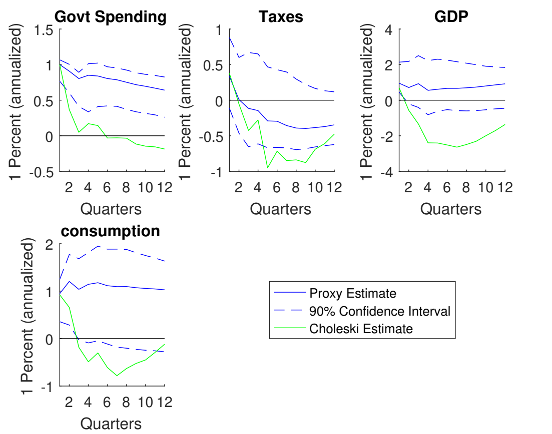
In this paper, we have analyzed which shock measure constitutes the best proxy variable to analyze the effects of government spending shocks. Our results are based on the reliability measure proposed by [1], R^2, and F statistics. We also check the IRF's and multipliers. We find that the measure by SPF and GB constitute the best proxy variable for government spending shocks and should thus preferably be employed in studying the effect of fiscal policy.

6. APPENDIX

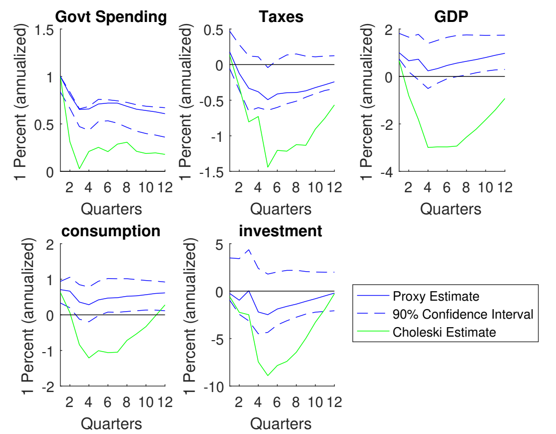
Here, we compare the multipliers and IRF's for the another data set for consumption and investment. The new data set has consumption and investment deflated by their own respective deflators. Results are almost the same as the other set. There is not significant difference between the results for both data sets of consumption and investment. Table 5 presents the government spending multiplier for the case with consumption and investment. Figure 6 is used to for robustness check.

Table 6.1: GDP Multiplier for a new data set

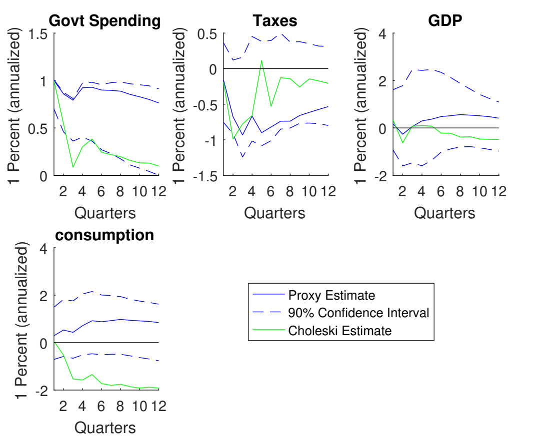
		Impact	4th	8th	12th
SPF	consumption-investment	1.0022	0.84281	0.77794	0.9444
	investment	0.9535	0.76924	0.62434	0.62045
	consumption	0.95635	0.88452	0.85898	0.962
GB	consumption-investment	0.21082	-0.32685	0.1813	0.11821
	investment	0.34121	-0.012981	0.071036	0.12331
	consumption	0.17392	0.074319	0.30563	0.39672



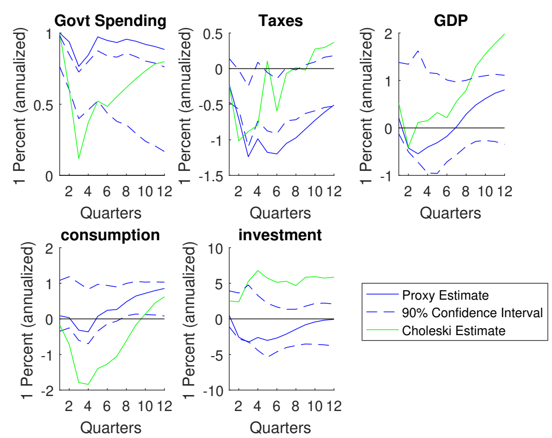
(a) Normalized IRF for SPF with consumption



(b) Normalized IRF for SPF with consumption



(c) Normalized IRF for GB with consumption



(d) Normalized IRF for GB with consumption and investment

Figure 6.1: IRF with Consumption and Investment- New data set

REFERENCES

[1] K. Mertens and M. O. Ravn, "The dynamic effects of personal and corporate income tax changes in the united states," *American economic review*, vol. 103, no. 4, pp. 1212–47, 2013.

[2] V. A. Ramey and M. D. Shapiro, "Displaced capital," tech. rep., National Bureau of Economic Research, 1998.

[3] R. B. Barsky and E. R. Sims, "Information, animal spirits, and the meaning of innovations in consumer confidence," *American Economic Review*, vol. 102, no. 4, pp. 1343–77, 2012.

[4] J. D. Fisher and R. Peters, "Using stock returns to identify government spending shocks," *The Economic Journal*, vol. 120, no. 544, pp. 414–436, 2010.

[5] J. H. Stock and M. W. Watson, "Generalized shrinkage methods for forecasting using many predictors," *Journal of Business & Economic Statistics*, vol. 30, no. 4, pp. 481–493, 2012.

[6] V. A. Ramey and S. Zubairy, "Government spending multipliers in good times and in bad: evidence from us historical data," *Journal of Political Economy*, vol. 126, no. 2, pp. 850–901, 2018.

[7] V. A. Ramey, "Identifying government spending shocks: It's all in the timing," *The Quarterly Journal of Economics*, vol. 126, no. 1, pp. 1–50, 2011.

[8] M. T. Owyang, V. A. Ramey, and S. Zubairy, "Are government spending multipliers greater during periods of slack? evidence from twentieth-century historical data," *American Economic Review*, vol. 103, no. 3, pp. 129–34, 2013.

[9] N. Ben Zeev and E. Pappa, "Chronicle of a war foretold: The macroeconomic effects of anticipated defence spending shocks," *The Economic Journal*, vol. 127, no. 603,

pp. 1568–1597, 2017.

[10] K. G. Lunsford, "Identifying structural vars with a proxy variable and a test for a weak proxy," 2015.

[11] C. Jentsch and K. G. Lunsford, "Proxy svars: asymptotic theory, bootstrap inference, and the effects of income tax changes in the united states," 2016.