The Economic Effects of Government Spending*
(First Draft)

Matthew Hall and Aditi Thapar
University of Michigan
August 15, 2016

Abstract

We create a forecast-based measure of government spending shocks from the Federal Reserve Board's Greenbook forecasts, which we use to estimate the effects of fiscal policy on output, consumption, investment, wages, and inflation. By using professional forecasts to obtain the unanticipated changes in government spending and to estimate the impulse response functions, we can very efficiently control for the information set of agents. We find that the government spending multiplier is close to one on impact and increases to 1.6 after about two years. Consumption and wages both increase, while inflation falls. Our findings lend support to structural VAR estimation results and imply that neither timing issues nor anticipation of government spending are driving the VAR results.

JEL Codes: E37, E62

*We would like to thank Charles Brown, Susan Collins, Gabriel Ehrlich, George Fulton, Chris House, and John Leahy for helpful comments. Special thanks to George Mytilineos for outstanding research assistance. All errors are our own. Please contact the authors by e-mail at mathall@umich.edu or athapar@umich.edu.

1 Introduction

Since the beginning of the Great Recession, especially once short-term interest rates hit the zero-lower-bound, there has been a resurgence of interest in the effectiveness of fiscal policy. The macroeconomic effects of fiscal policy are highly contested on both theoretical and empirical grounds. Researchers disagree on the direction and magnitude of the effects of fiscal policy. Traditional Keynesian models predict that output, consumption, hours worked, and real wages all increase in response to increased government spending. In contrast, neoclassical models imply that consumption and real wages fall in response to positive government spending shocks. Due to the differing effect on consumption, Keynesian models typically imply larger government spending multipliers than the neoclassical models. The empirical evidence is also mixed. Results from calibrated structural models tend to depend on the set of assumptions on which they are built.[1] There is conflicting evidence from the two primary reduced-form empirical techniques as well: Vector Autoregressions (VARs) largely support the Keynesian view while the narrative-based war-dates analysis supports a more classical one. In general, but not always, structural Vector Autoregressions (VARs) tend to show larger multipliers, while the narrative-based analyses find evidence for more modest ones.[2] The differences in the results between the two reduced-form approaches has been attributed to VARs being unable to adequately control for anticipated future policy actions.

A fundamental issue in estimating the effects of government policy is determining changes in government spending or taxes that are unanticipated. The challenge researchers face is that many current changes in government policy were either planned in the past or reflect the anticipated response of government spending to other variables. Vector Autoregression (VAR) based analyses of fiscal policy have been criticized on the grounds that they incorrectly attribute anticipated changes in policy as shocks since a standard VAR analysis is unable to account for future changes in policy

[1]For example, Cogan et al. (2010) employ a DSGE model with sticky prices and wages and estimate a fiscal multiplier that is significantly less than unity. In contrast, Gali et al. (2007) modify a new-Keynesian model with a large proportion of rule-of-thumb consumers and find a multiplier of two.

[2]See Blanchard and Perotti (2002), Perotti (2007), and Gordon and Krenn (2010) for examples of VAR approaches yielding a multiplier greater than one, and Mountford and Uhlig (2009) for an example of a VAR approach with a multiplier below one. For narrative-based approaches, including use of war dates, see Ramey and Shapiro (1998), Eichenbaum and Fisher (2005), and Ramey (2011a) for estimates of a government spending multiplier less than one.

that are announced but are not visible in the data.

There are two goals to this paper. First, we control for expected changes in future policy actions by using professional forecasts to estimate the effects of changes in government spending on output, consumption, investment, wages, and inflation. Professional forecasts implicitly contain a very rich information set, which includes anticipated changes in fiscal policy and other economic and policy variables. By building our estimation methodology on forecast data, we can appropriately time any changes in fiscal policy as well as disentangle the effects of government spending from other sources on output. Second, we study whether the VAR results of the effects of fiscal policy are biased due to their inability to control for anticipated changes in fiscal policy.

We use one-quarter ahead forecast errors on government spending from the Federal Reserve Board's Greenbook forecasts to obtain structural shocks to fiscal policy. We then project longer-term (one through seven quarter ahead) forecast errors of GDP, consumption, business fixed investment, residential construction, wages and inflation on to the structural shocks to government spending to obtain the impulse response functions. Since professional forecasts provide conditional expectation based on all information that is available at the time, by using forecast errors to estimate impulse response functions we can control for all the information that is available.

We find that in response to an increase in government spending, GDP, consumption, real wages, and residential construction increase; inflation falls; while business fixed investment is largely unchanged. These results hold for different recursive ordering restrictions, as well as when we restrict government spending to only defense spending. The government spending multiplier is close to one on impact and increases to 1.6 after about two years. We compare our methodology to other approaches and demonstrate the importance of effectively controlling for the information set of agents. Our methodology allows us to obtain estimates that are more precise than other leading methodologies in the field. Our results tend to support the standard VAR estimates. Since we have very effectively controlled for anticipated changes in fiscal policy, we conclude that the differences in the results between VAR-based and other narrative-based war-dates studies is not caused by the inadequate control of anticipated policy changes in VARs.

Standard VAR analyses for estimating the effects of government spending, follow the lead of Blanchard and Perotti (2002) who estimate a structural VAR with government spending, taxes, and GDP.[3] The main advantage of VARs is their ability to capture complex patterns in the data with very few structural assumptions. The main disadvantage of a VAR is that, due to the small set of included variables, the identified shocks are not truly exogenous innovations, since they do not adequately control for future policy actions even when that information is publicly available. Many anomalous results in the VAR literature have been attributed to inadequate controls for information. Ramey (2011b) finds that changes in defense and non-defense spending are anticipated by private agents a few quarters prior to the actual changes in spending. She finds that professional forecasts predict fiscal shocks from VARs, and concludes that VAR shocks are not true innovations.

The standard narrative approach tends to follow the Ramey and Shapiro (1998) war-dates analysis.[4] This approach searches for events such as wars which are likely to be exogenous to the economy and analyzes their effects. The main advantage of the narrative approach is that the shocks are credibly exogenous to the business cycle. The main disadvantage of this approach is that the results are driven by the military buildup of World War II and the Korean War, and there are only four war-dates in the baseline case which leads to an extremely small sample of innovations. It is also not clear that military spending is representative of all government spending.

Recently, researchers have begun including one-quarter ahead professional forecasts in their reduced-form empirical analysis as a way to control for anticipated changes in fiscal policy. The forecasts are typically included in an ad hoc manner, as proxies for the multitude of variables that cannot be added in a VAR, and to control for expectations of anticipated changes. Ramey (2011b) finds that narrative-based shocks and professional forecasts Granger-cause VAR shocks. She estimates a VAR augmented by the one-period-ahead forecast error in government spending growth based on the Survey of Professional Forecasters. She finds that temporary rises in govern-

[3]For other examples of VAR-based analyses of the effects of fiscal policy see Perotti (2007); Mountford and Uhlig (2009); and Auerbach and Gorodnichenko (2012).

[4]For other examples of studies using the war-dates see Barro and Redlick (2011); Ramey (2011b); and Ramey and Zubairy (2014). Favero and Giavazzi (2012) use the Ramey (2011b) war-dates in a VAR. Romer and Romer (2010) use the Federal Reserve's narrative record to estimate shocks to taxes. Ramey (2016) provides an excellent overview of the literature.

ment spending do not stimulate the economy, as opposed to VARs without forecast errors which tend to find the opposite. Auerbach and Gorodnichenko (2012) also incorporate expectations of professional forecasts in a VAR to estimate government spending multipliers in expansions and recessions. They find that including forecast errors for government spending in their VAR specification tends to exacerbate effects of government spending shocks. Both studies only include the one-step ahead forecast errors as a way of controlling for expected changes in policy when estimating the shocks to government spending. The innovation of our paper is using longer term forecasts to control for information in estimating impulse response functions and government spending multipliers, which other studies do not do.

The rest of this paper is organized as follows: Section 2 details our empirical approach and contrasts it with those of the VAR and standard linear projection techniques. Section 3 describes the forecast data from our three sources - the Federal Reserve Board's Greenbook forecasts, the University of Michigan's Research Seminar in Quantitative Economics forecasts, and the Survey of Professional Forecasters - and assesses the relative performance of professional forecasts versus those generated from VARs. In Section 4 we discuss the basic empirical results of our methodology in terms of impulse response functions and multipliers using Greenbook forecasts. We perform various robustness checks in Section 5, including (i) comparing our empirical results with those using standard linear projection, (ii) using defense shocks as exogenous drivers of government spending shocks, (iii) altering our sample period, and (iv) comparing Greenbook results to those generated by other data sources. Section 6 concludes.

2 Econometric Framework

In this section, we discuss the econometric framework used in this paper. To motivate our methodology, we begin by first describing the basic framework of the VAR approach as applied to the estimation of the effects of fiscal policy.

2.1 Vector Autoregression

Consider an $n \times 1$ vector of economic variables, $X_t = \{G_t, T_t, Y_t\}$, where G_t represents real government spending, T_t represents real tax revenues, and Y_t is real GDP. All variables are in logs. X_t is determined by the history of changes to its components ε according to some functional form:

$$X_t = F(X_0, \varepsilon_t, \varepsilon_{t-1}, ..., \varepsilon_0). \tag{1}$$

A structural VAR with p lags represents an approximate linearization of this relationship,

$$AX_t = \Gamma + B(L)X_{t-1} + e_t^{VAR}, \tag{2}$$

where A is a $n \times n$ matrix, Γ is a $n \times 1$ vector, $B(L)$ is a lag polynomial of order p, and e_t^{VAR} is a $n \times 1$ vector of structural errors. The reduced-form VAR associated with this structural VAR is defined as:

$$\begin{aligned} X_t =& A^{-1}\Gamma + A^{-1}B(L)X_{t-1} + A^{-1}e_t^{VAR} \\ =& \widetilde{\Gamma} + \widetilde{B}(L)X_{t-1} + u_t. \end{aligned} \tag{3}$$

We can represent the reduced-form errors (u_t) as the one-step ahead forecast errors generated by the linear VAR model,

$$u_t^{VAR} = X_t - \mathbb{E}_{t-1}[X_t | X_{t-1}, X_{t-2}, ..., X_{t-p}]. \tag{4}$$

The vector of reduced-form residuals u_t is typically used to identify the structural shocks e_t. The structural errors in the VAR are related to the reduced-form errors by the following relation:

$$e_t^{VAR} = Au_t^{VAR}. \tag{5}$$

The structural shocks are recovered from the reduced-form errors by making assumptions about the structure of the A matrix. A standard assumption in the VAR literature is to identify structural errors by assuming that A is lower-triangular and that the structural shocks are independent. This recursive identification assumption, also known as a Cholesky decomposition, implies that any contemporaneous covariance between variables is attributed to the variables ordered earlier in the VAR. We describe the identification of government spending shocks in greater detail in the context of our methodology in the next subsection.

To obtain the impulse response functions, note that we can represent the linear VAR with p lags as an $MA(\infty)$ process as given by,

$$X_{t+s} = \sum_{i=0}^{\infty} D_i e_{t+s-i}^{VAR} = D_0 e_{t+s}^{VAR} + D_1 e_{t+s-1}^{VAR} + ... + D_{s-1} e_{t+1}^{VAR} + ... \tag{6}$$

The $s-$step ahead forecast error from the VAR can be represented as,

$$X_{t+s} - E_t X_{t+s} = \sum_{i=0}^{s-1} D_i e_{t+s-i}^{VAR} = D_0 e_{t+s}^{VAR} + D_1 e_{t+s-1}^{VAR} + ... + D_{s-1} e_{t+1}^{VAR}. \tag{7}$$

Since the e_{t+s}^{VAR} are assumed to be independent, D_{s-1} can be interpreted as the effect of a structural shock next period to the variable of interest s periods later.

2.2 Professional forecasts approach

As opposed to the standard VAR methodology, we use the forecasts of professional forecasters to obtain the forecast errors required for the identification of the structural shocks to fiscal policy as well as to estimate the impulse response functions via local projection methods.[5] We also obtain empirical estimates of short- and long-run fiscal multipliers.

In this paper, we apply the basic approach developed in Thapar (2008) to the estimation of the effects of fiscal policy.[6] Using professional forecasts, we do not need to estimate the forecasts of

[5]See Jordà (2005) for a generalized discussion of these methods. Thapar (2008) and Cochrane and Piazzesi (2002) use financial market information to estimate the effects of monetary policy via local projecions. Ramey (2016) summarizes developments in the literature on the propagation of macroeconomic shocks.

[6]Thapar (2008) used futures markets and professional forecasts to estimate the effects of monetary policy. In

the variables of interest since we treat them as data. Therefore, we bypass estimation of equations (3) and (4) and instead calculate the one-step ahead forecast errors directly, expressed as

$$u_t^f = X_t - \mathbb{E}_{t-1}[X_t | I_{t-1}^f].$$

In this formulation X_t is the actual value of X in period t and $E_{t-1}[X_t | I_{t-1}^f]$ is the period $(t-1)$ expectation of X_t given I_{t-1}^f, the information set available to forecaster f at time $t-1$.[7] The advantage of our approach is that I_{t-1}^f is presumably much larger than the information set of the VAR. By construction, VAR forecasts can only contain information from the small subset of variables that are included in the system.

Following the basic VAR approach, the structural shocks (ε_t) can be backed out from the one-step ahead forecast errors via the following relation:

$$\varepsilon_t^f = A u_t^f. \tag{8}$$

Although our methodology is flexible with respect to identification technique, note that in this paper we focus only on the Blanchard and Perotti (2002) strategy to identify structural shocks (ε_t^f) from our one-step-ahead forecast errors.

We assume for simplicity that we have only three variables in the system: taxes (T), government spending (G), and output (Y), so $X_t = (G_t, T_t, Y_t)'$.[8] In this case, the system of equations in

general, the estimation of the effects of (monetary or fiscal) policy are complicated due to the fact that policy changes affect the economy and that policy-makers respond to economic developments. Using professional forecasts allows one to disentangle these circular effects, when analyzing the effects of either monetary policy or fiscal policy.

[7] In the VAR above, $I_{t-1}^f = X_{t-1}, X_{t-2}, ..., X_{t-p}$ and $\mathbb{E}_{t-1}[X_t | X_{t-1}, X_{t-2}, ..., X_{t-p}] = \widetilde{\Gamma} + \widetilde{B}(T) X_{t-1}$, or the best linear fit.

[8] Any additional variables of interest would be ordered after government spending and not affect identification of the government spending shocks. We therefore ignore these variables in the present discussion.

(8) can be written equivalently as

$$u_t^{f,\tau} = \alpha_{11} u_t^{f,g} + \alpha_{12} u_t^{f,y} + \varepsilon_t^{f,\tau} \tag{9}$$

$$u_t^{f,g} = \alpha_{21} u_t^{f,\tau} + \alpha_{22} u_t^{f,y} + \varepsilon_t^{f,g} \tag{10}$$

$$u_t^{f,y} = \alpha_{31} u_t^{f,\tau} + \alpha_{32} u_t^{f,g} + \varepsilon_t^{f,y}. \tag{11}$$

Three identifying restrictions are necessary in order to obtain uncorrelated structural shocks $\varepsilon_t = (\varepsilon_t^{f,\tau}, \varepsilon_t^{f,g}, \varepsilon_t^{f,y})'$ from the system of equations above. Following Blanchard and Perotti (2002), we obtain these restrictions in the manner described below.

The first identifying restriction sets $\alpha_{22} = 0$, since it typically takes policy-makers longer than one quarter to identify and respond to shocks to economic activity. This restriction assumes that there is no contemporaneous feedback from output to government spending, in either the mind of the forecaster or the economy at large.

The second identifying restriction estimates α_{12}, the elasticity of taxes with respect to output, directly from external sources. As Blanchard and Perotti note, since tax revenues tend to adjust automatically to changes in GDP, we must construct a measure of cyclically-adjusted taxes $(u_t^{f,\tau} - \alpha_{12} u_t^{f,y})$ that are not correlated with innovations in output. This measure instruments $u_t^{f,\tau}$ in equation (11) to ensure that α_{31} is unbiased. A detailed derivation of the tax elasticity for our forecast data is available in the appendix, Section A.

While most researchers agree that fiscal policy does not respond contemporaneously to other variables in the system, there is no consensus on whether spending or tax decisions are made first. The third identifying restriction reflects this ambiguity. Assuming government spending decisions are made prior to tax decisions (tax innovations do not affect spending decisions contemporaneously) imposes $\alpha_{21} = 0,\ \alpha_{11} \neq 0$. This restriction, combined with $\alpha_{22} = 0$, implies that our forecast errors for government expenditures are equivalent to our shocks ($u_t^{f,g} = \varepsilon_t^{f,g}$). We refer to this as the "government spending ordered first" scenario.

Reversing the decision-making order and assuming tax decisions as being made prior to spending implies that $\alpha_{11} = 0, \ \alpha_{21} \neq 0$. Based on this assumption, we obtain a different series for government spending shocks, which we refer to as the "taxes ordered first" scenario. This restriction is equivalent to assuming that shocks to government spending do not affect taxes in the contemporaneous quarter except through the output channel via the elasticity α_{12}.

As discussed in Section (4) below, the ordering between taxes and spending has a negligible effect on our baseline results. Blanchard and Perotti (2002) also find that the assumption about the ordering of government spending and tax decisions does not affect their results.

Instead of tracing out an impulse response iteratively through the same linear model, as in a VAR, we project our forecast errors at each of the H horizons for which they exist, onto our structural shocks:

$$X_{t+s} - \mathbb{E}_{t-1}[X_{t+s}|I_{t-1}^{f}] = c + \beta^{s}\varepsilon_t + \theta_{t+s} \ \text{ for } s = 1, 2, ..., H, \tag{12}$$

where β^s represents the response of X, s periods in the future, to a shock to ε today. Once we have our series for structural shocks, the H equations in (12) can be estimated individually via least squares using a heteroskedasticity and autocorrelation consistent covariance matrix. The number of variables for which we can estimate impulse responses is constrained only by the availability of forecasts for the variable.

Estimating impulse response functions for a larger set of variables is possible in a VAR by increasing the block of equations (3) above. The number of required identifying restrictions r in (5), given the number of variables n, increases at the same rate as in our approach: $r = \frac{n(n-1)}{2}$. The more pressing problem in a VAR setting is that the number of degrees of freedom shrinks much more rapidly. For example, in a VAR with no constant, no time trend, n variables, and 4 lags, the number of coefficients to be estimated is $4 \times n^2$. Adding a 4^{th} variable increases the number of coefficients to be estimated in the first stage alone from 36 to 64, while adding a 5^{th} variable increases the number of coefficients to 100. This is the reason most VARs include only a very small subset of variables. In contrast, since our projection process uses forecasts as data, we can simply substitute the new variable into equation (12), so we obtain more precise impulse responses when dealing with smaller data sets.

By using forecast errors directly, we gain several additional advantages. First, our errors are not the result of the limited information set of a VAR, $\{X_{t-1}, X_{t-2}, ..., X_{t-p}\}$, but rather include all the information known to the forecasters at time $t-1$. This information set should include any changes in fiscal policy announced by the government, as well as other changes to the economy that are anticipated. Second, the VAR fits a single, time-invariant, linear model to the data. It cannot account for non-linearities or regime changes. Finally, by expanding the system of equations (9)−(11) to include other macroeconomic variables of interest, such as inflation or the instrument of monetary policy, we can generate additional impulse responses or simply control for more shocks and better isolate the effect of fiscal policy innovations on the variable of interest.

2.3 Comparison to other local projection methods

Our approach differs from a VAR, as described in detail above, but also from other local projection methods, as described in Jordà (2005), and applied by Auerbach and Gorodnichenko (2013), Ramey (2016), and Romer and Romer (2010). For a given series of structural shocks, one can obtain impulse response functions by estimating a sequence of regressions via a standard local projection method, given by

$$X_{t+s} = c + b^{X,s} e_{1t} + control\ variables + \eta_{t+s} \text{ for each } s = 1, 2, ..., h. \tag{13}$$

The innovation of our methodology is to use the expectations of the variable of interest and estimate impulse response functions

$$X_{t+s} - \mathbb{E}_{t-1}[X_{t+s}|I^f_{t-1}] = c + \beta^{X,s}\varepsilon_t + \theta_{t+s}, \text{ for each } s = 1, 2, ..., H, \tag{14}$$

where X_{t+s} is the variable of interest, such as log GDP, at time $t+s$, or the sum of log GDP between time t and $(t+s)$, and $b^{X,s}, \beta^{X,s}$ are the impulse responses of X at horizon s to a shock in period t from each of the two approaches. This method can be applied to shocks obtained from any source, such as a Cholesky decomposition, monetary policy shocks, or oil price shocks.

The control variables in equation (13) are added to account for information that is available to agents in period t, when the fiscal policy shock hits the economy.

Our methodology involves projecting forecast errors onto structural shocks, as given by equation (14). The advantage of our approach is that including the forecasts ($\mathbb{E}_{t-1}[X_{t+s}|I^f_{t-1}]$) in the estimation should provide a far superior control of the information available at the time than any subset of variables that researchers would include.

Equations (13) and (14) ought to be asymptotically equivalent. We should expect our approach, however, to be more efficient, since it controls for a larger subset of information. We test this proposition in Section 2.3 and find that although the point estimates from both approaches are similar, there are strong efficiency gains from using our methodology.

2.4 Multipliers

The impulse response functions estimated above provide us the elasticity of output with respect to government spending at various horizons. To convert the elasticities to multipliers, we follow Mountford and Uhlig (2009) and calculate multipliers as the integral or present value of the response of GDP divided by the integral response of government spending.[9] The present value multiplier (ϕ) at horizon s is given by

$$\phi^s = \frac{\sum_{i=0}^{s} (1+r)^{-i} \beta^{Y,i}}{\sum_{i=0}^{s} (1+r)^{-i} \beta^{G,i}} \times \frac{Y}{G}\,, \tag{15}$$

where $\beta^{j,i}$ refers to the impulse response of variable $Y = \{GDP\}$ at horizon $i = 1, 2, ..., N$ and r is the real interest rate. The real interest rate is assumed to be the average real interest rate over the relevant sample period.

[9]The impulse response at horizon s, β^s, is the percent response of Y to a one percentage point shock to Y. $\beta^s = \frac{dY/Y}{dG/G}$. The multiplier, on the other hand, refers to $\phi^s = \frac{dY}{dG} = \beta^s \times \frac{Y}{G}$.

3 Data

In this section, we describe the data that is used in this paper. To implement our approach we use forecasts (and forecast errors) of GDP, government spending and tax and transfer information, all on a quarterly basis. We focus on the Federal Reserve Board's Greenbook (GB) forecasts in this paper. For the purpose of comparison we also use quarterly forecasts from the University of Michigan's Research Seminar in Quantitative Economics (RSQE) and the Survey of Professional Forecasts (SPF), conducted by the Federal Reserve Bank of Philadelphia.

3.1 Greenbook forecasts

The Greenbook is the colloquial name given to the official report titled "Current Economic and Financial Conditions – Summary and Outlook" that is produced by the research staff at the Board of Governors of the Federal Reserve System. This report is prepared for the Federal Open Markets Committee (FOMC) prior to every scheduled FOMC meeting. It includes the forecasts of the US economy and is available to the FOMC members six days prior to every scheduled meeting. Before 1981 there was a regularly scheduled meeting every month. Since 1981 there have been eight scheduled meetings every year, approximately once every six weeks.

The Greenbooks are publicly available, but with a six year lag, which constrains the end of our sample period to the end of 2010. Although Greenbook forecasts are available from 1966 onwards, the maximum forecast horizon was increased from five quarters ahead to seven quarters ahead in 1979. Eight-quarter-ahead forecasts are available beginning in 1988. In this paper, we use data beginning in 1978.[10]

Some of the forecasts of the variables of interest are available in levels, others in growth rates, and a select few are in both levels and growth rates. Still others, like GDP and the GDP deflator,

[10]Until 1974, the forecasts for the first ten months of the year were only available for the current calendar year, while in the last two months each year the forecast horizon was extended to the next calendar year. As a result of this forecasting convention, the Greenbooks contain forecasts between one and four quarters ahead, depending on the quarter. Due to the short forecast horizon of the data before 1979, we focus on the post-1978 sample period for this paper.

which had been available in both levels and growth rates, began to be published only in growth rate terms beginning in 2005. This implies that whether we decide to use data in levels or growth rates, we would need to extrapolate some data to ensure that we have a consistent time series for each variable. In this paper, we decided to use growth rates for all variables of interest.

The Greenbook forecasts estimate the growth of future real government expenditures on consumption and investment, including federal, state, and local expenditures. We use this set of estimates as our measure of government spending. Our expected growth rate of government receipts is a gross measure, for want of forecast data on transfers from the government to persons. In addition, state and local receipts are not included in the Greenbook forecasts, so all analyses of Greenbook data that include taxes/receipts use forecast errors of federal gross receipts, indexed by the GDP deflator. All other variables have standard definitions.

We are interested in estimating the cumulative effects of government spending on variables of interest at different horizons. Since we are working with growth rates rather than levels, for the purpose of estimating impulse response functions and multipliers we calculate cumulative growth rates for each variable of interest (X), $g^X_{t+s} = \frac{X_{t+s}}{X_t}$, for each date t and horizon $s = 1, 2, ..., H$ for which we have forecast data available. As discussed above, data limitations imply that $H = 7$.

3.2 RSQE forecasts

The University of Michigan's RSQE (Research Seminar in Quantitative Economics), the original quarterly macroeconomic forecaster, began forecasting the U.S economy in 1952. They produce forecasts at regular intervals that have shifted slightly over time. Unfortunately, RSQE forecasts from 1952-1982 were destroyed in a fire. We use RSQE's one through seven quarter ahead forecasts of our variables of interest for the 1983-2010 sample period.[11]

While their current practice is to publish detailed write-ups of their forecasts in March, May, September, and November, in the past their August set of projections were their featured (and

[11] Although from 1983 onwards we have forecasts for at least ten quarters ahead, we use only up to the seven quarter ahead forecasts since the primary motivation for using the RSQE dataset is to compare the results using RSQE forecasts to those estimated using the Greenbook data.

sometimes only) 3rd-quarter forecast. We draw one forecast from each quarter between 1983 and 2010, according to the following conventions: we use the March, May, and November forecasts throughout our analysis, and we use the August forecast from 1983-2007, and the September forecast from 2008-2010.

For government receipts from RSQE, we construct a measure of real government net receipts, which includes state and local receipts as well as federal and non-federal transfers to persons (excluding interest payments). For all other variables the definitions are the same across the RSQE and Greenbook datasets. As with Greenbook forecasts, we work with cumulative forecast errors in growth rates for the RSQE forecasts as well.

3.3 SPF forecasts

The Survey of Professional Forecasters (SPF) is conducted once every quarter by the Federal Reserve Bank of Philadelphia. [12] We use SPF forecast data for the 1983-2010 period. The SPF dataset contains mean forecasts of approximately 30-40 anonymous private sector forecasters. Anonymity implies that concerns over reputation and a desire for publicity do not bias forecasts. The dataset includes forecasts of quarterly real GDP, real consumption, real business fixed investment, real investment in residential construction, and the GDP deflator. Information on wages was not collected as part of the survey. The survey is conducted in the middle month of each quarter, at which time the survey respondents report their forecasts for each of the next four quarters and for the subsequent calendar year. The fact that the longer-term forecasts are on a calendar-year basis leads to an asymmetry between the short-term and the longer-term forecasts. This gives rise to a complex set of correlations in the data. We therefore only use the one-to-four-quarters-ahead forecasts in our analysis.

[12] Prior to 1990 the survey was conducted by the American Statistical Association and the National Bureau of Economic Research and was called the ASA/NBER Economic Outlook Survey. Croushore (1993) discusses the survey in detail.

3.4 Real-time or current data

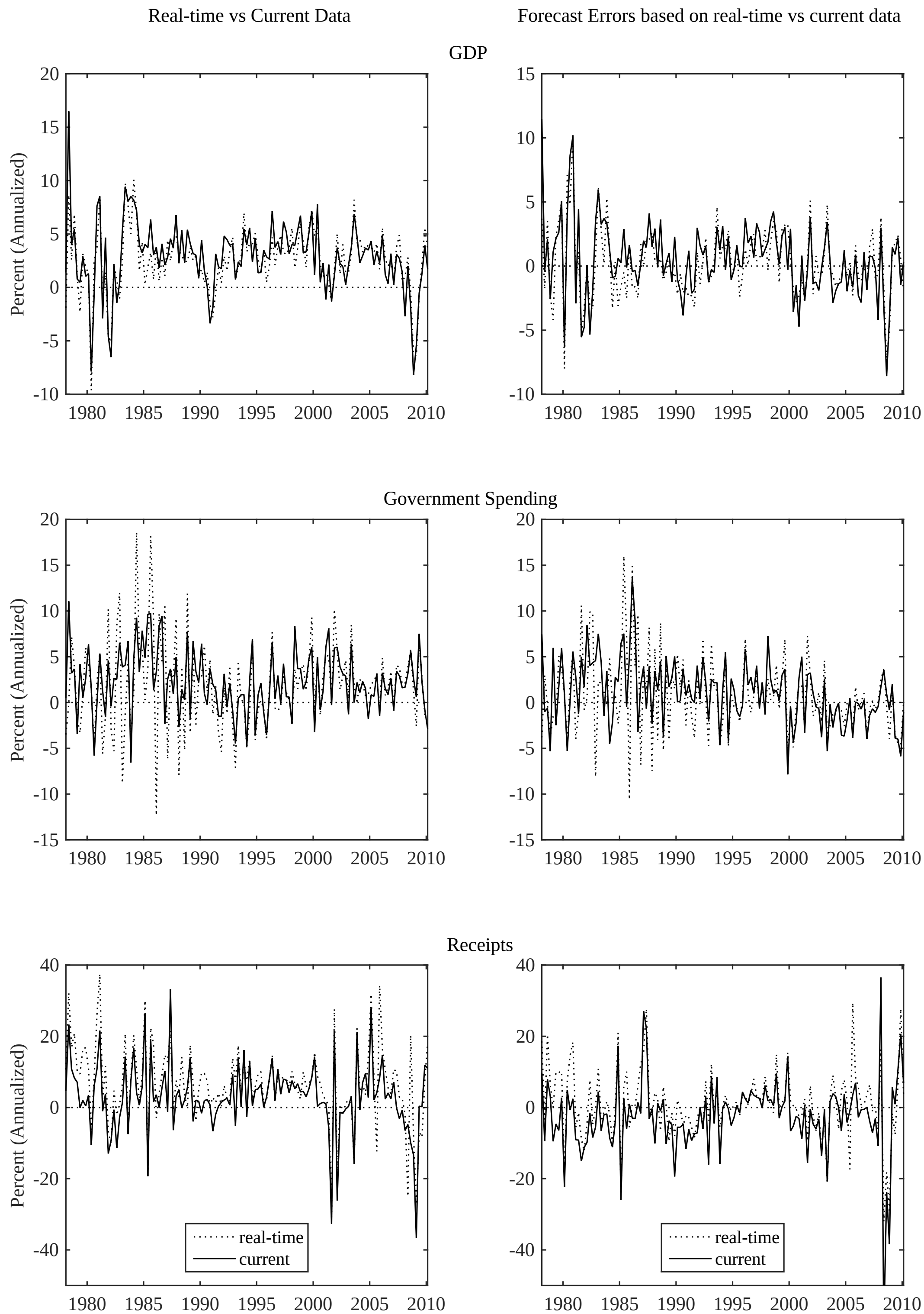
In the use of forecast data, a primary issue in defining forecast errors is whether to use the data as is currently available (current data) or to use the data that were available around the time the forecasts were made (real-time data).[13] Many studies using forecast data tend to use real-time data to construct forecast errors; even though more recent vintages of historical data tend to reflect our improved understanding of the true value of the variable over time. These studies prefer to match forecasts with the contemporaneous understanding of the underlying data. The main argument for basing forecast errors on real-time data is that calculating forecast errors based on the currently available data might artificially inflate the forecast errors and incorrectly attribute definition changes as shocks.

In this paper, we base forecast errors on current data for three reasons. First, since we are working with cumulative growth rates, estimating forecast errors from real-time data is not necessarily desirable. Cumulative growth rates are subject to revisions over progressively longer time horizons than single-quarter rates. In addition, if an annual or comprehensive revision occurs between the forecast and the realization of the variable of interest, then not only the level but also the definition of variables is subject to change. In this case it is not clear, even in theory, which real-time data concept is appropriate to compare to a forecast. Linked quarterly growth rates would reflect a blend of revisions and conceptual frameworks, while any other choice would be arbitrary and potentially vary according to the length of time over which the growth rate is calculated. Second, more recent vintages of historical data tend to reflect our improved understanding of the true growth rates over time which is what forecasters are likely trying to predict. Finally, since we are working with growth rates rather than levels definition changes are less likely to have large effects on the growth rate of the variables of interest.

Figure 1 compares the real-time and current data for the growth rate of the variable indicated.

[13]The Bureau of Economic Analysis (BEA) publishes three estimates for the NIPA accounts. The advance (first) estimate, the second estimate, and the third estimate are published one, two, and three months respectively after the end of any given quarter. Every summer the BEA conducts an annual revision of the data over the past three calendar year and approximately every five years the BEA conducts a comprehensive revision. The third estimate is changed any time the BEA makes major changes.

Figure 1: Growth Rates of Real-Time vs Current Data and Forecast Error of Growth Rates



Real-time refers to the data as it existed two quarters after the end of the reference quarter. Current refers to data as it currently exists. In the forecast errors column, real-time refers to errors constructed using the real-time data for actuals and current refers to current data being used for actuals.

The column on the left compares growth rates of the real-time data to those from current data for our three primary variables: GDP, government spending, and government receipts. The column on the right compares Greenbook forecast errors of the growth rate of the variable of interest, derived from real-time data vs using current data. An additional issue with using real-time data as the proxy for actual data in constructing forecast errors can be observed in the middle panel of the left column of the figure. Government spending tends to be revised significantly in later years. The period between 1981 and 1988 appears to be responsible for most of the subsequent revisions, and the current data is much less volatile than in the initial telling. GDP growth rates, however, are more similar between the real-time series and current data. Receipts are also subject to reasonably substantial revisions, but not as great as for GDP.

One of the most important steps in our analysis in this paper is estimating shocks to government spending. If we construct forecast errors from real-time data, then the revisions to government spending in the pre-1990 period are likely to be attributed as shocks. To avoid this misspecification of shocks as well as due to the complications with using real-time data when working with cumulative growth rates, in this paper we use current data to construct forecast errors.

3.5 Comparison of Forecast Errors

In this section we compare various performance measures of the Greenbook forecasts of GDP, government spending, and government receipts, the three main variables we need to estimate spending shocks. We compare the Greenbook forecast errors to the RSQE and SPF forecast errors, as well as those generated by several VAR specifications, all of which include 4 lags of each variable. The VAR specifications are chosen to highlight key points. There are two sets of choices involved. The first choice is whether to use the current data or to use real-time data. The second choice is whether to estimate only one VAR over the entire model, or to estimate a rolling VAR, where the VAR is estimated every period and its predictions for $(Y_{t+i}, G_{t+i}, T_{t+i})'$ are based on its simulated values of $(Y_{t+j}, G_{t+j}, T_{t+j})'$ for $j = 1, 2, ..., (i-1)$ rather than the actual values of those variables.

We obtain the current quarter and the one- to eight-quarter-ahead forecasts, and forecast errors, from four VARs. The first VAR specification $(VAR1)$ uses current data and is estimated only once over the entire sample period. This specification benefits from consistent definitions of each variable over the entire sample and from using the entire sample period to estimate the best linear relationship between the variables. Actual data is used when making out of sample forecasts. The second VAR specification $(VAR2)$ also uses current data, but it is estimated on a rolling basis and without access to the current period's output, receipts, and government spending (which would not be known at the time a forecast must be made). The VARs forecasts, rather than actual data, are used when estimating the $2, 3, ..., 7$ quarter ahead forecasts. The third VAR $(VAR3)$ is estimated only once over the entire sample period using real-time data. As in the first VAR, here too actual data is used for forecast horizons greater than one. The fourth VAR specification $(VAR4)$ is a rolling VAR that is estimated using real-time data, which most closely mimics the forecasts that would be made by a professional forecaster. The VAR forecasts are used for out-of-sample forecasting when the forecast horizon is greater than one. *Ex ante* we expected $VAR1$ to perform the best, since it uses current data and is estimated only once over the entire sample, and $VAR4$ to perform the worst.

Table 1 compares the root mean-squared error (RMSE), and mean absolute error (MAE) of each forecast series over two sample periods, 1983-2010 (the sample common to the available Greenbook, RSQE, and SPF forecasts) and 1990-2010. The latter sample period is included due to the irregularities discussed above of government spending growth in the real-time data vs the current data during the 1980s.

Among the VARs, in general $VAR1$, which uses current data and is estimated only once over the entire sample period, is the best performing of the VARs by most measures. Its MAE and RMSE are lower for all three variables over the short run for both sample periods. The second best, $VAR3$, performs nearly as well as $VAR1$ and manages to equal or beat it in GDP forecast errors in select periods. These two VAR specifications benefit from three advantages of over-fitting: 1) X_t data is assumed to be known in period t, when in reality only data up to X_{t-1} is available,

Table 1: Forecast Error Statistics

1983 - 2010												
	GDP								G		T	
	fe^{Y}_{t}	fe^{Y}_{t+1}	fe^{Y}_{t+2}	fe^{Y}_{t+3}	fe^{Y}_{t+4}	fe^{Y}_{t+5}	fe^{Y}_{t+6}	fe^{Y}_{t+7}	fe^{G}_{t}	fe^{G}_{t+1}	fe^{T}_{t}	fe^{T}_{t+1}
Root Mean Squared Error (percentage points)												
GB	1.93	2.32	2.45	2.53	2.58	2.59	2.54	2.80	3.29	3.59	9.52	11.28
RSQE	2.13	2.41	2.54	2.65	2.66	2.60	2.62	2.68	3.79	3.56	10.12	10.88
SPF	2.00	2.39	2.45	2.55	2.62	--	--	--	3.11	3.22	--	--
VAR1	2.58	2.43	2.40	2.39	2.38	2.38	2.39	2.39	3.41	3.39	13.68	13.45
VAR2	2.71	2.77	2.82	2.66	2.61	2.59	2.60	2.63	3.50	3.48	14.35	14.29
VAR3	2.58	2.44	2.42	2.42	2.36	2.34	2.30	2.30	4.98	4.98	15.05	14.88
VAR4	2.81	2.69	2.72	2.45	2.29	2.25	2.25	2.25	5.31	4.79	16.34	15.08
Mean Absolute Error (percentage points)												
GB	1.47	1.83	1.89	1.96	1.96	1.95	1.93	2.08	2.57	2.79	5.79	6.93
RSQE	1.69	1.91	1.87	1.93	1.97	1.84	1.86	1.95	2.81	2.77	7.35	7.94
SPF	1.57	1.84	1.83	1.86	1.88	--	--	--	2.46	2.48	--	--
VAR1	1.89	1.81	1.78	1.77	1.76	1.76	1.77	1.77	2.69	2.66	9.90	9.70
VAR2	1.99	2.00	2.03	1.86	1.83	1.81	1.83	1.86	2.75	2.74	10.59	9.60
VAR3	1.88	1.81	1.78	1.79	1.74	1.72	1.69	1.70	3.64	3.64	11.36	11.18
VAR4	2.11	2.01	2.09	1.83	1.67	1.65	1.66	1.65	3.77	3.50	11.99	10.55

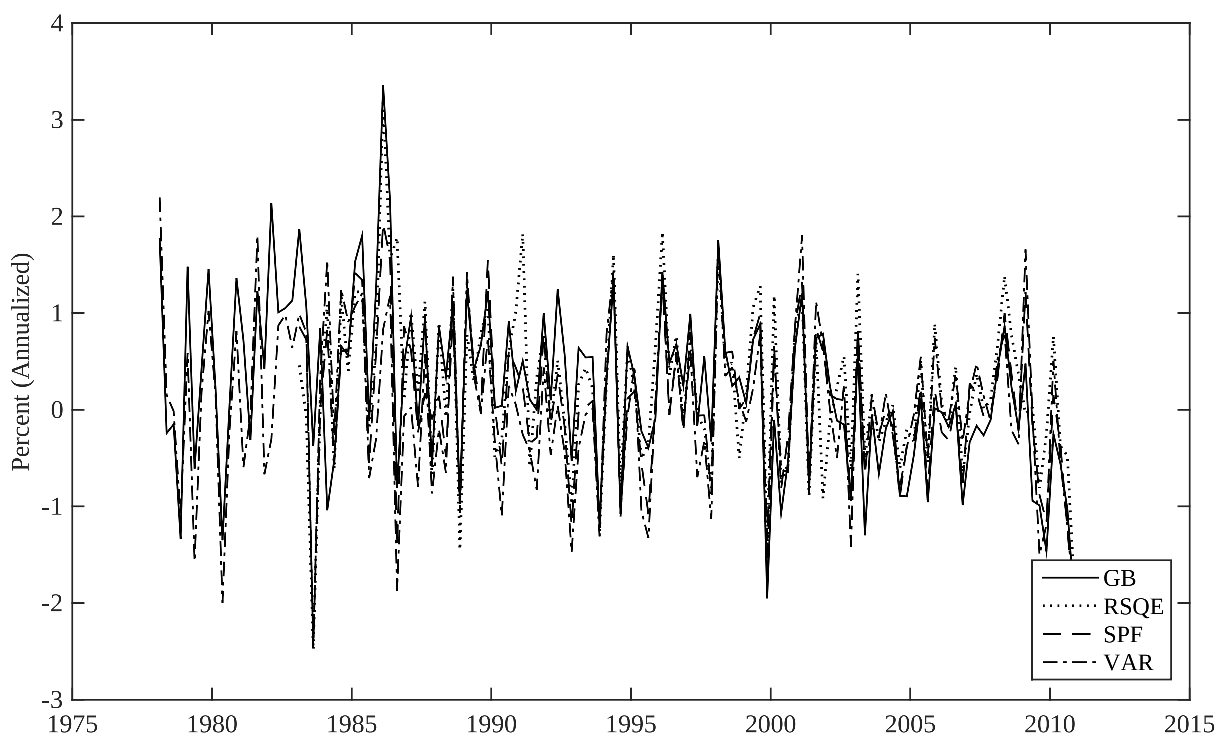
1990 - 2010												
	GDP								G		T	
	fe^{Y}_{t}	fe^{Y}_{t+1}	fe^{Y}_{t+2}	fe^{Y}_{t+3}	fe^{Y}_{t+4}	fe^{Y}_{t+5}	fe^{Y}_{t+6}	fe^{Y}_{t+7}	fe^{G}_{t}	fe^{G}_{t+1}	fe^{T}_{t}	fe^{T}_{t+1}
Root Mean Squared Error (percentage points)												
GB	1.96	2.34	2.55	2.68	2.72	2.70	2.63	2.83	2.62	3.02	10.32	11.48
RSQE	2.10	2.50	2.79	2.90	2.85	2.77	2.72	2.79	2.88	3.05	10.28	11.43
SPF	1.96	2.37	2.55	2.71	2.69	--	--	--	2.66	2.83	--	--
VAR1	2.61	2.61	2.59	2.50	2.47	2.46	2.47	2.49	3.16	3.15	13.92	13.92
VAR2	2.82	2.98	3.02	2.79	2.66	2.68	2.69	2.71	3.26	3.21	14.86	15.12
VAR3	2.52	2.52	2.52	2.50	2.44	2.43	2.43	2.43	3.55	3.48	14.66	14.66
VAR4	2.74	2.78	2.80	2.53	2.34	2.33	2.33	2.32	3.70	3.45	16.05	15.89
Mean Absolute Error (percentage points)												
GB	1.49	1.87	2.01	2.08	2.07	2.03	2.00	2.09	2.09	2.37	5.98	6.78
RSQE	1.71	2.00	2.08	2.11	2.09	1.96	1.92	2.02	2.21	2.44	7.68	8.37
SPF	1.55	1.83	1.91	1.97	1.93	--	--	--	2.07	2.16	--	--
VAR1	1.94	1.93	1.91	1.85	1.82	1.80	1.82	1.84	2.45	2.43	9.95	9.97
VAR2	2.12	2.14	2.16	1.96	1.87	1.89	1.90	1.91	2.54	2.52	10.89	10.07
VAR3	1.87	1.87	1.87	1.84	1.79	1.78	1.78	1.78	2.79	2.74	10.89	10.87
VAR4	2.05	2.07	2.13	1.89	1.71	1.70	1.70	1.69	2.93	2.70	11.38	10.86

Notes: VAR1 uses current vintage data, estimated over entire sample period. VAR2 uses current vintage data, estimated on a rolling basis without access to the current period's data. VAR3 uses real-time data, estimated over the entire sample period. VAR4 uses real-time data, estimated on a rolling basis without access to the current period's real-time data.

and 2) all future data is assumed to be known for purposes of estimating the VAR coefficients, and 3) $E_t X_{t+i}$ for $i \geq 2$ is estimated using actual data for periods $i = 1, 2...(i-1)$. $VAR2$ and $VAR4$ that are estimated on a rolling basis, perform significantly worse in both the near-term and over a longer horizon.

Comparing the VARs to the forecasters, in general forecasters perform better than the VARs that are estimated on a rolling basis. The five to seven quarter ahead Greenbook forecasts are not available at all horizons for our sample period, whereas the VARs statistics are based on data at all dates. This likely leads to the poorer performance of the Greenbook forecasts at longer horizons. The RMSE and MAE of the two rolling VAR's are typically higher than those for the professional forecasts. Since professional forecasters have access to a larger information set, they can make superior forecasts to those estimated by a simple VAR.

Figure 2: Structural errors (G ordered first) from different forecasters



In the 1990-2010 sample period, the Greenbook and RSQE forecast errors exhibit a lower

RMSE for all three variables than all our VAR specifications, and lower MAE in government spending and government receipts.

Figure 2 compares the structural errors for government spending from four different forcast sources: Greenbook, RSQE, SPF, and a VAR.[14] Although there are differences in the shocks generated by different forecast models, the correlation between each set of series lies between $0.7-0.9$.

4 Baseline Results

In this section we discuss the main results from the estimation of the effects of government spending shocks on some key macroeconomic variables in the United States. All the impulse responses reported in this paper should be interpreted as the cumulative effect to the variable of interest in period $(t+i)$ of a one percentage-point shock to real government spending in period t. We present the impulse response functions based on our methodology and compare them to the impulses obtained from a VAR.

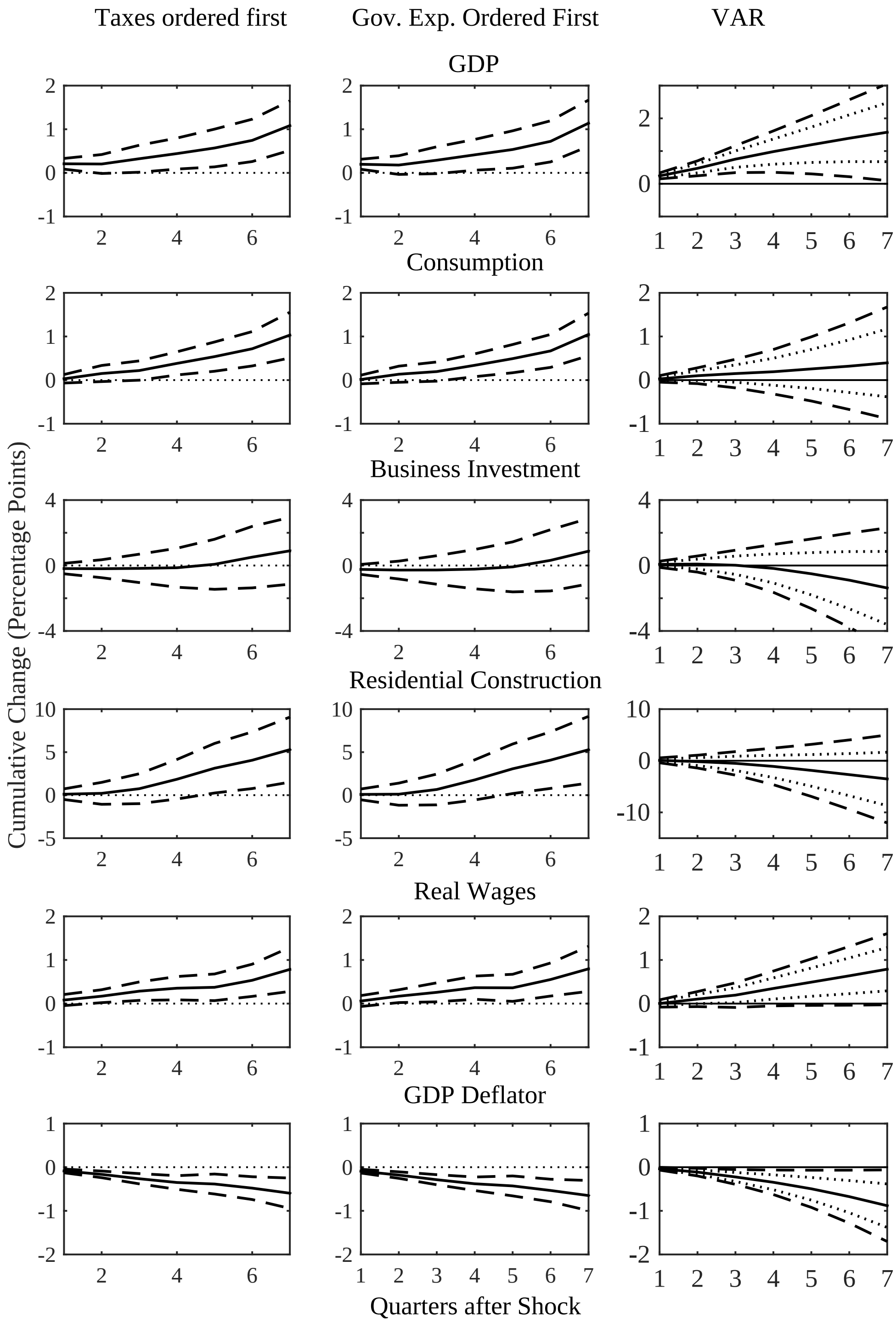
In all the figures in the rest of the paper, solid lines (—) represent the impulse responses, dashed lines (– –) represent 90% confidence intervals, and dotted lines (...) in the VAR represent 68% confidence intervals. Although the convention in the VAR literature has converged to report only one standard deviation bands (68% confidence intervals), our approach allows us to work with conventional measures of significance.

Figure 3 presents our baseline results. It depicts the impulse response functions of a one-percentage-point government spending shock to GDP, consumption, business fixed investment, residential investment, wages, and prices. The impulse responses in the first two columns are estimated using equations (12) above for the Greenbook forecasts over the 1978-2010 sample period, while the last column is based on the VAR methodology.

The column on the left is based on the projection of structural errors derived by ordering (cyclically-adjusted) taxes first, government spending second, and all other variables last. The

[14]We estimate a standard VAR with G, T, Y using current data, and estimated only once over the entire sample. This corresponds to $VAR1$ above.

Figure 3: Baseline Results, 1978-2010
Response to 1 Percentage Point Increase in G



ordering of the panel in the center is based on government spending being ordered first, cyclically-adjusted taxes second, and all other variables last. Recall, since we are only interested in the effects of shocks to government spending, the order of the variables after government spending is irrelevant to our analysis here.

The last column shows the impulse responses of the variables of interest obtained from a standard VAR with government spending ordered first, government receipts second, real GDP third, and one other additional variable of interest ordered last.[15] For real GDP the VAR includes only the first three variables. All variables are in logs. The small subset of variables is the standard set of variables that are used in the literature. The VARs are estimated over the post-1960 sample period.

As discussed in Section 2 above, the limitation to the number of horizons for which we can estimate impulse responses is based on data availability. The Greenbook results are therefore shown only for 7 quarters after a shock hits.[16] For ease of comparison to the Greebook results, we include only seven quarters of the implulse responses for the VARs as well.

Comparing the first two columns of Figure 3, we find that the effects of government spending are not sensitive to the ordering of the variables in the identification of the structural shocks. Our results support Blanchard and Perotti's (2002) claim that the ordering of government spending and taxes has a negligible effect of the impulse responses. For the rest of the paper, unless mentioned otherwise, we report results with government spending ordered first and cyclically-adjusted taxes second.

All three sets of results predict a statistically significant increase in the growth rate of GDP and a decrease in inflation. The VAR results are statistically significant only for these two variables. Our methodology implies that the impact response is a 0.2 percentage point change in output, which then accumulates to an increase of 1.4 percentage points after seven quarters. Figure

[15] Blanchard and Perotti (2002) find that the VAR results are largely unaffected bythe ordering of government spending and taxes. We therefore only present one set of VAR results.

[16] Although the Greenbook dataset contains 8-quarter ahead forecasts beginning in 1989, this forecast horizon is available only once a year. Due to the late initial date and the extremely small sample, we report impulse response functions only up to 7 quarters in this paper.

3 depicts an interesting effect of government spending on prices. Increased spending leads to a statistically significant decrease in the growth of the GDP deflator in all approaches. This "fiscal price-puzzle" has also been reported by Canova and Pappa (2007) and Mountford and Uhlig (2009).

All three approaches imply an increase in consumption and real wages. The results are statistically significant for the Greenbook forecasts, but not in the VAR. One of the main criticisms of the structural VAR approach to estimating the effects of fiscal policy is that the VAR does not account for changes in fiscal policy that were announced in the past. Ramey (2011b) finds that changes in defense and non-defense spending are anticipated by private agents a few quarters prior to the actual changes in spending, implying that VAR shocks are not true innovations. Our methodology accounts for any announcement effects. Presumably, the staff at the Federal Reserve accounts for announced changes to fiscal policy in their forecasts. Our results indicate, therefore, that the systematic differences between results using the narrative approach and VARs is not due to the timing of changes in fiscal policy.

The impact on business fixed investment is statistically insignificant in all three sets of results. For residential construction, however, Greenbook forecasts predict a strong and statistically significant increase five quarters after the increase in government spending. The VAR results are once again insignificant at all horizons.

Our estimates of the government spending multiplier are presented in table 4. Using Greenbook forecasts implies that, over the $1978 - 2010$ sample period, a government spending shock yields a fiscal multiplier of just under 1 on impact. The multiplier rises to 1.1 one year after the shock and $1.6 - 1.7$ seven quarters after the shock to government spending. This estimate lies just outside the $0.8 - 1.5$ range in Ramey's (2011b) meta-analysis of the literature as the most likely range for a temporary, deficit-financed spending multiplier.

Table 2: Fiscal Multiplier of G on GDP using Greenbook Data, 1978-2010

Number of Quarters ahead	1	2	3	4	5	6	7
Taxes Ordered First	0.92	0.86	1.00	1.16	1.28	1.45	1.67
G Ordered First	0.88	0.79	0.91	1.06	1.17	1.33	1.60

5 Extensions

In this section we exploit the flexibility of our methodology to discuss the results of various robustness exercises. We begin by comparing our approach to highlight the importance of controlling for anticipated changes in policy. Next, we discuss the impulse response functions and multipliers using shocks to federal government spending on defense. We then examine the importance of sample period choice on our results. Finally, we compare Greenbook forecast-driven results with those from other professional forecasters.

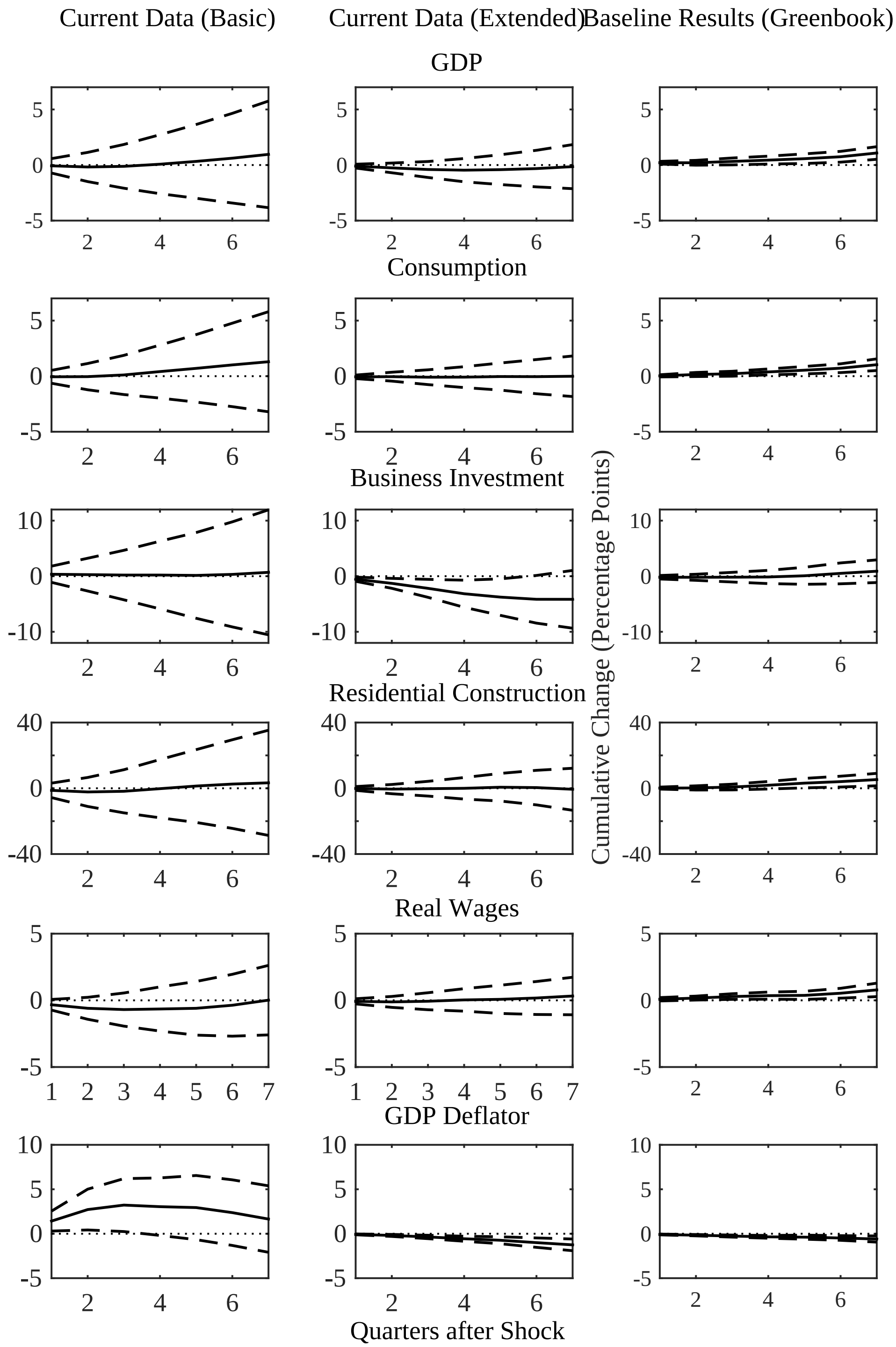
5.1 Importance of controlling for the information set

In this section, we use a standard local projection technique using current data to emphasize the importance of controlling for the information and highlight the advanatage of using our methodology.

In Section 2.3 we compared our method with that of Jordà (2005). Recall, the main difference between our approach and the standard Jordà local projection approach is that we project structural shocks on s-period ahead forecast errors $(Y_{t+s} - E_t Y_{t+s})$, whereas the standard Jordà method involves projecting shocks onto the s-period ahead value of the variable, say the log of Y_{t+s}. A well documented problem in the literature on estimating the effects of macroeconomic shocks is the limited information set of a VAR-based analysis. Jorda's local projection method is subject to the same problem. Researchers using local projection techniques typically include various additional control variables, which we can also include but by using forecasts in our estimation we have automatically included a very rich information set into our estimation.

Figure 4 compares the impulse response functions based on the two methodologies. The rightmost column contains our baseline results using our methodology with Greenbook forecast errors,

Figure 4: Using Greenbook Shocks and Current Data for Projections
Response to 1 Percentage Point Increase in G, 1978-2010



estimated using equation (12) individually for each horizon s, repeated here for convenience.

$$X_{t+s} - \mathbb{E}_{t-1}[X_{t+s}|I_{t-1}^{f}] = c + \beta^{s}\varepsilon_{t} + \theta_{t+s} \text{ for } s = 1, 2, ..., H,$$

where X_{t+s} is the growth rate of X between period t and $t+s$, $E_{t-1}[X_{t+s}|I_{t-1}^{f}]$ is the period $(t-1)$ Greenbook forecast of X_{t+s}, and β^{s} is the impulse response of X at horizon s to a one-percentage point innovation in ε_t. The impulse response at horizon s is the cumulative change in the growth rate of X between period t and $t+s$.

The middle and left columns contain the impulse responses based on standard local projection techniques. The impulses are obtained by estimating the equation below individually for each horizon s.

$$\sum_{j=1}^{s} X_{t+j} = c + b^{X,s} e_{1t}^{f} + control\ variables + \eta_{t+s} \text{ for each } s = 1, 2, ..., H.$$

where X_{t+j} is the log of the variables of interest in period $(t+j)$: real GDP, consumption, business fixed investment, residential construction, wages, and inflation, and $b^{X,s}$ is the impulse response of X at horizon s. The impulse response at horizon s is the cumulative change in the growth rate of X between period t and $t+s$. To highlight the advantage of our approach and the importance of control variables, the middle column includes only a linear time trend as a control variable. The right column includes a linear time trend, as well as two lags each of the shocks, government spending, receipts, GDP, and the own lags of the variable of interest.

As predicted, projecting current data rather than forecast errors on government spending shocks leads to results that are less precisely estimated with much wider error bands than our approach. Comparing the left column to the middle column for each variable implies that including appropriate controls for the information set leads to smaller confidence bands for the impulse response functions. Our methodology, however, leads to the smallest confidence bands.

Using current data, the results are statistically significant only for business investment and inflation. Once we control for the information set, both current data and our approach implies that inflation rates fall in response to an increase in government spending. In current research we are investigating the reasons for this result.

5.2 Defense Shocks

Much of the literature on fiscal shocks emphasizes the role of defense spending shocks on GDP, as defense spending is in many cases orthogonal to current macroeconomic conditions.[17] Shocks to defense spending are therefore credibly exogenous to the business cycle. The defense spending shocks are usually obtained via narrative approaches that involve reading newspaper and magazine articles to find announcements of changes in military spending. The main disadvantage of this narrative approach is that the results are driven by the military buildup of World War II and the Korean War. It is also not clear that military spending is representative of all government spending.

Unlike the VAR approach, our shocks control for any expected changes in policy. Nevertheless, we examine the effects of defense spending shocks for comparison with our baseline results. We project forecast errors of our variables of interest on government spending shocks instrumented by defense shocks, defined according to the same process as government spending shocks.[18] The sample period in this section is 1982-2010 since Greenbook forecasts for defense spending are available only beginning in 1982.

Figure 5 compares our baseline projections with those from instrumenting government spending with defense shocks. In each case, instrumenting with defense shocks does not alter the overall results. Relative to our baseline results, using defense spending as our measure of fiscal shocks leads to slightly dampened effects over the first six quarters after the shock, and the results are no longer statistically significant for most variables.

The fiscal multipliers at different horizons associated with using defense spending shocks are shown in Table 3 below. The multipliers are similar to the baseline case for the first six quarters after a shock, with somewhat of a spike in the seventh quarter.

[17] See Ramey and Shapiro (1998).

[18] When government spending is ordered first, the result shown, the forecast error is the structural shock. When government spending is ordered second, the structural shock is derived from equation 10.

Figure 5: Instrument Government Spending Shocks by Defense Spending Shocks
Response to 1 Percentage Point Increase in G, Greenbook Forecasts, 1982-2010

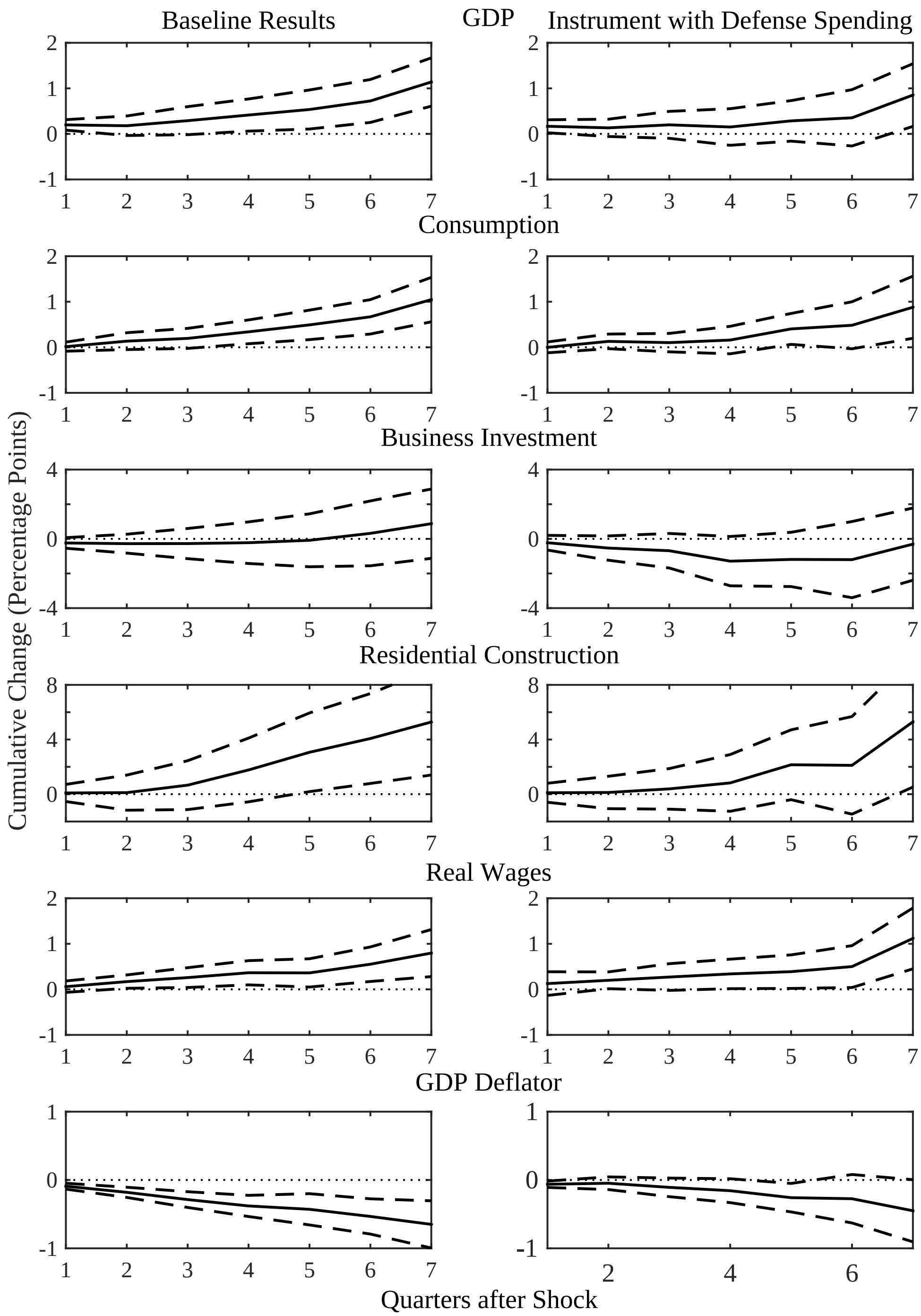


Table 3: Fiscal Multiplier when using Defense Shocks, 1982-2010

Number of Quarters ahead	1	2	3	4	5	6	7
Defense Shocks	0.82	0.82	0.99	1.00	1.16	1.32	1.93

5.3 Sub-samples

In this section we investigate whether our choice of sample periods influences the results. In the process, we present circumstantial evidence addressing the question of whether or not the efficacy of fiscal policy changed during the Great Moderation. The advantage of our approach, especially relative to a VAR-based analysis, is that due to the smaller number of variables we need to estimate, it is easy for us to apply our methodology to small samples. In particular, we examine the effects of a government spending shock over four different sample periods. The two terminal dates were chosen to represent either (i) the full sample available (2010) or (ii) to exclude the Great Recession completely (2005), respectively. For start dates, 1978 is the first year in our dataset, while 1990 represents the beginning of the Great Moderation.

Figure 6 compares the effects of government spending shocks on GDP, consumption, business fixed investment, residential investment, wages, and the GDP deflator over the four different sample periods. Qualitatively, the results over the different subsamples are very similar to the baseline results over the 1978-2010 sample, depicted in the leftmost column. GDP, consumption, and wages all increase while the GDP deflator decreases; and the results are statistically significant at most horizons. The effect on business fixed investment is statistically insignificant, while residential construction tends to increase after one year.

Quantitatively, the main difference in the results stems from the fact that the impulse responses are stronger as we exclude earlier data. The exclusion of the Great Recession from our sample period, however, lowers the estimated impulse responses in the full sample and slightly raises them in the sample beginning in 1990. These results are consistent with the relative size of the multipliers, shown in table 4, estimated from each sub-sample.

Figure 6: Subsample Stability, Greenbook Forecasts
Response to 1 Percentage Point Increase in G

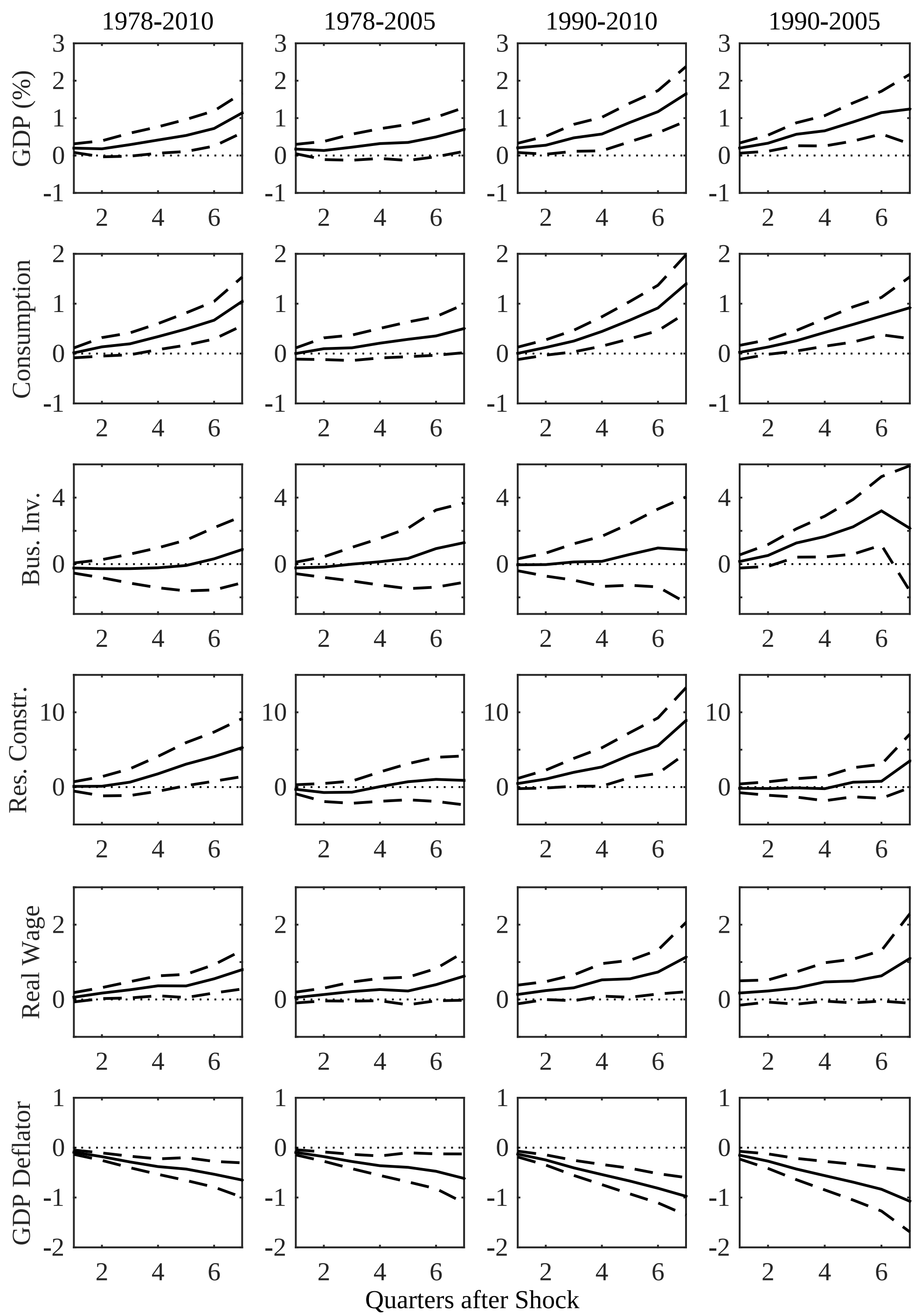


Table 4: Fiscal Multiplier of G on GDP, subsamples

$Number\ of\ Quarters\ ahead$	1	2	3	4	5	6	7
1978 – 2005	0.70	0.59	0.72	0.88	0.98	1.13	1.30
1978 – 2010	0.92	0.86	1.00	1.16	1.28	1.45	1.67
1990 – 2005	0.92	1.29	1.80	2.13	2.48	2.86	3.07
1990 – 2010	1.03	1.21	1.51	1.69	1.94	2.24	2.59

5.4 Comparison to Other Private Forecasts

One potential criticism of our approach is that the results are based on a potentially flawed model of the US economy. To investigate whether or not our results are solely due to our use of the Greenbook forecasts, we estimate our model using the RSQE and SPF forecasts. As discussed in Section 3, the Greenbook and RSQE forecasts are both based on large structural models of the US economy, while the Survey of Professional Forecasts are the consensus forecasts of 15-40 different professional forecasters. The number and composition of the forecasters included in the consensus SPF forecasts has changed over time.

Figure 7 compares the impulse response functions from each of these three datasets, over the 1983-2010 sample period.[19] Data limitations restrict the SPF impulse response to only four quarters after a shock to government spending.[20] Our baseline results are largely unchanged. In response to an increase in government spending, GDP, consumption, residential construction increase, while prices fall. For business fixed investment, GB results are not statistically significant, RSQE predicts a statistically significant fall, while SPF predicts a fall which is barely statistically significant. As a result, the multipliers implied by RSQE and SPF are lower than those implied by the GB forecasts.

[19]This is the sample over which all three datasets overlap. RSQE and SPF results are similar if we extend the sample to 2015.

[20]We do not include results for real wages in this figure since the SPF dataset does not include forecasts for wages. Both RSQE and GB predict that real wages increase in response to an increase in government spending.

Figure 7: Comparing Different Forecasters - Greenbook, RSQE, SPF (1983-2010)
Response to 1 Percentage Point Increase in G,

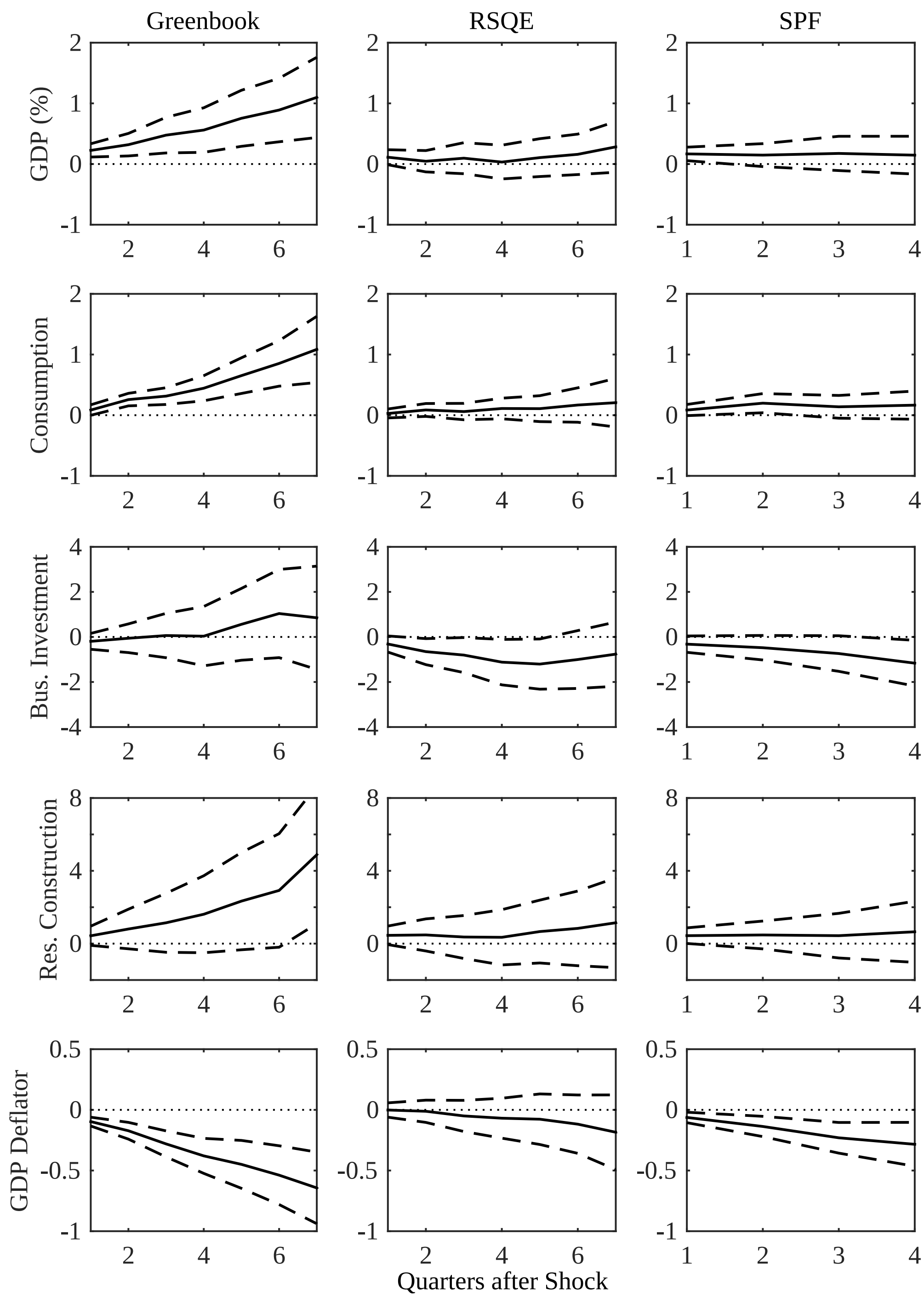


Table 5: Fiscal Multiplier of G on GDP, Forecasters, 1983-2010

Number of Quarters ahead	1	2	3	4	5	6	7
GB (T ordered first)	1.01	1.13	1.37	1.50	1.64	1.78	1.92
RSQE (T ordered first)	0.50	0.35	0.37	0.30	0.32	0.36	0.46
SPF (G ordered first)	0.75	0.71	0.74	0.70	-	-	-

6 Conclusion

In this paper we use the Federal Reserve's Greenbook one-step ahead forecasts to estimate a new measure of shocks to government spending, and longer horizon Greenbook forecasts to estimate the effect of changes in government spending on GDP, consumption, business fixed investment, residential construction, wages, and inflation. By using professional forecasts to estimate the impulse response functions we can very effectively control for the information set of policy-makers, which leads to estimates that are much more precisely estimated than VAR or standard local projection techniques to estimating the effects of government spending.

We find that there is a statistically significant increase in GDP, consumption, wages, and residential construction in response to an unanticipated increase in government spending. The inflation rate decreases in response to a fiscal expansion. The effects on business fixed investment are not statistically significant. Broadly speaking, our results lend support to the VAR-based literature. Using professional forecasts implies that we can control for anticipated fiscal policy changes. Our results indicate that the differences in the results between VAR-based and narrative-based studies are probably not being caused by the VAR mis-attributing anticipated government spending changes as shocks.

Our results imply that the government spending multiplier on impact is about 0.9 and it rises to 1.6 about two years after a government spending shock. We also find that government spending multipliers have increased in the post-1990 period, with the largest multipliers during the great moderation.

References

Auerbach, Alan, and Yuriy Gorodnichenko. 2012. “Measuring the Output Responses to Fiscal Policy.” *American Economic Journal: Economic Policy* 4 (2): 1-27.

Auerbach, Alan, and Yuriy Gorodnichenko. 2013. Fiscal Multipliers in Recession and Expansion Fiscal Policy After the Financial Crisis, eds. Alberto Alesina and Francesco Giavazzi, University of Chicago Press.

Blanchard, Olivier, and Roberto Perotti. 2002. “An Empirical Characterization of the Dynamic Effects of Changes in Government Spending and Taxes on Output.” *Quarterly Journal of Economics* 117 (4): 1329-1368.

Burnside, Craig, Martin Eichenbaum, and Jonas D.M. Fisher. 2004. “Fiscal shocks and their consequences.” *Journal of Economic Theory* 115 : 89-117.

Cogan, John F., Cwik, Tobias J., Taylor, John B, and Wieland, Volker. 2010. “New Keynesian versus old Keynesian government spending multipliers.” *Journal of Economic Dynamics and Control* 34 : 281-295.

Edelberg, Wendy, Martin Eichenbaum, and Jonas D.M. Fisher. 1999. “Understanding the Effects of a Shock to Government Purchases.” *Review of Economic Dynamics* 2 : 166-206.

Eichenbaum, Martin and Jonas D.M. Fisher. 2005. “Fiscal Policy in the Aftermath of 9/11.” *Journal of Money, Credit, and Banking* 37 (1) : 1-22.

Fatas, Antonio, and Ilian Mihov. 2001. “The effects of Fiscal Policy on Consumption and Employment: Theory and Evidence.” CEPR Discussion Papers 2760.

Galí, J., López-Salido, J.D., and Vallés, J. 2007. “Understanding the Effects of Government Spending on Consumption.” *Journal of the European Economic Association* 5 : 227-270.

Giorno, Claude, et al. 1995. “Estimating Potential Output, Output Gaps and Structural Budget Balances.” *OECD Economics Department Working Papers* No. 152.

Gordon, Robert J., and Krenn, Robert. 2010. “The End of the Great Depression 1939-41: Policy Contributions and Fiscal Multipliers.” *NBER Working Papers*. No. 16380

Jones, John Bailey. 2002. “Has fiscal policy helped stabilize the postwar U.S. economy?” *Journal of Monetary Economics* 49: 709-746.

Jordá, Oscar. 2005. “Estimation and Inference of Impulse Responses by Local Projections.” *American Economic Review* 95 (1): 161-182.

Mankiw, N. Gregory, and Matthew Weinzierl. 2011. “An Exploration of Optimal Stabilization Policy.” *Brookings Papers on Economic Activity*. Spring 2011: 209-249.

Mertens, Karel, and Morten O. Ravn. 2011. “Understanding the aggregate effects of anticipated and unanticipated tax policy shocks.” *Review of Economic Dynamics* 14: 27-54.

Monacelli, Tommaso, Roberto Perotti, and Antonella Trigari. 2010. “Unemployment fiscal multipliers.” *Journal of Monetary Economics* 57: 531-553.

Mountford, Andrew, and Harald Uhlig. 2009. “What are the Effects of Fiscal Policy Shocks?” *Journal of Applied Econometrics* 24: 960-992.

Perotti, Roberto. 1999. “Fiscal Policy in Good Times and Bad.” *Quarterly Journal of Economics* 114 (4): 1399-1436.

Perotti, Roberto. 2011. “Expectations and Fiscal Policy: An Empirical Investigation.” IGIER Working Paper 429.

Perotti, Roberto. 2007. “In Search of the Transmission Mechanism of Fiscal Policy.” *NBER Macroeconomics Annual* 22: 169-249.

Ramey, Valerie A. 2016. “Macroeconomic Shocks and their Propagation.” NBER Working Paper Series No. 21978.

Ramey, Valerie A. 2011a. “Can Government Purchases Stimulate the Economy?” *Journal of Economic Literature* 49 (3): 673-685.

Ramey, Valerie A. 2011b. “Identifying Government Spending Shocks: It’s All in the Timing.” *Quarterly Journal of Economics* 126 (1): 1-50.

Ramey, Valerie A., and Matthew D. Shapiro. 1998. “Costly capital reallocation and the effects of government spending.” *Carnegie-Rochester Conference Series on Public Policy* 48: 145-194.

Ramey, Valerie A. and Sarah Zubairy, 2014. “Government Spending Multipliers in Good Times and in Bad: Evidence from 20th Century Historical Data,” Working paper.

Romer, Christina D. and David H. Romer. 2010. “The Macroeconomic Effects of Tax Changes: Estimates Based on a New Measure of Fiscal Shocks.” *American Economic Review* 100(3): 763-801.

Sims, Christopher A. 1980. “Macroeconomics and Reality” *Econometrica* 48 (1): 1-48.

Taylor, John B. 2000. “Reassessing Discretionary Fiscal Policy.” *The Journal of Economic Perspectives* 14 (3): 21-36.

Thapar, Aditi. 2008. “Using private forecasts to estimate the effects of monetary policy.” *Journal of Monetary Economics* 55: 806-824.

Woodford, Michael. 2011. “Simple Analytics of the Government Expenditure Multiplier.” *American Economic Journal: Macroeconomics* 3 (1): 1-35.

Appendix

A Tax Elasticity of Output

The within-period elasticity of taxes with respect to output $\eta_{\tau,y}$ is estimated in a similar manner to Blanchard and Perotti (2002). $\eta_{\tau,y}$ is an average of various tax elasticities weighted by their respected shares of total taxes, $\sigma_{\tau_i,\tau}$, as follows:

$$\eta_{\tau,y} = \sum_i \eta_{\tau_i,B_i}\eta_{B_i,Y}\sigma_{\tau_i,\tau}. \tag{A.1}$$

In equation A.1 above, the elasticity of each tax is broken into two separate components: 1) the elasticity of tax i to its tax base B_i, and 2) the elasticity of the tax base to output. With the respect to relatively proportional taxes, such as payroll taxes, η_{τ_i,B_i} is closer to unity, while for more proportional taxes, such as the federal income tax, η_{τ_i,B_i} is a larger number.

In constructing these various elasticities, Blanchard and Perotti (2002) rely on estimates provided by Giorno et al (1995) and supply their own estimates based on a series of regressions of log changes of the tax base on leads and lags of log changes to output. The contemporaneous effect is treated as the within-period elasticity. We create a similar series of estimates below for each measure of taxation used in our analysis.

A.1 Greenbook

Using BEA data of personal income, output, and various tax receipts, supplemented by CPS measures of employment and average weekly earnings among the employed, we directly estimate elasticity measures for three federal taxes that make up more than 90 percent of gross federal receipts: the personal income tax (TPF), the corporate income tax (TCF), and social insurance taxes (TSIF).

A.1.1 Social Insurance Taxes

The relationship between social insurance taxes and output is characterized by Giorno et al (1995) as $T = t(w)w(E)E(Y)$, where $t(w)$ is the tax rate on wages, $w(E)$ is earnings, and $E(Y)$ is the employment level. With a little bit of rearrangement, the elasticity of federal social insurance taxes $\eta_{TSIF,y}$ can then be separated into three parts: 1) the elasticity of employment with respect to output $\eta_{E,y}$, 2) the elasticity of weekly average wages among the employed with respect to employment $\eta_{w,E}$, and 3) the elasticity of social insurance taxes with respect to wages $\eta_{wt,w}$.

$$\eta_{TSIF,y} = \eta_{E,y}\left[\eta_{w,E}\eta_{wt,w} + 1\right] \tag{A.2}$$

We use the quarterly average of seasonally-adjusted time privately employed production and non-supervisory employees as our measure of E. For wages w, we use the average weekly earnings of privately employed production and nonsupervisory employees. Regressing the log change of each variable on one lead and four lags of the log change in output and employment, respectively, we obtain within-period elasticity estimates of $\eta_{E,y} = .33$ and $\eta_{w,E} = .66$ when we restrict the sample to 1978-2010 (the period for which we have GB forecasts of federal receipts). These estimates are close to those of Blanchard and Perotti (2002), who find $\eta_{E,y} = .42$ and $\eta_{w,E} = .62$. When we change our sample to 1964-1997, the last year included in Blanchard and Perotti's analysis, our estimates are .36 and .68, respectively. Using different measures of earnings and wages slightly alters our results, but the product of the two elasticities is surprisingly robust to our choice of wage and employment pairing.

The elasticity of taxes with respect to earnings, $\eta_{wt,w}$, is taken to be 1.0 in Giorno et al and Blanchard and Perotti. We adopt instead the convention that the elasticity is equal to the share of covered earnings, which changes slowly over time. Thus, we use $\eta_{wt,w} = 0.85$. Combining the three measures, we get $\eta_{TSIF,y}$ slightly greater than 0.5.

A.1.2 Federal Personal Income Taxes

The elasticity of the federal personal income tax $\eta_{TPF,y}$ is calculated in a departure from Blanchard and Perotti. They utilize the elasticities calculated above and combine it with an estimate of $\eta_{wt,w}$ from Giorno et al, who find that in the United States the gross-earnings elasticity of income tax rises from 2.5 in 1978 to 3.9 in 1992. As this is an annual number, and the earnings base for the federal income tax is substantially different from earned income, we instead repeat the analysis above, but with personal income directly. We define *TPF*$(B(Y))$ to be the federal income tax *TPF* as a function of the income base B, which is a function of output Y. The elasticity $\eta_{TPF,y}$, then, is

$$\eta_{TPF,y} = \eta_{TPF,B}\eta_{B,y}. \tag{A.3}$$

We estimate $\eta_{TPF,B}$ using the same lead and lag formulation as for social insurance taxes, where B is aggregate personal income, and obtain an estimate of 2.34. The elasticity $\eta_{B,y}$ is 0.48, yielding $\eta_{TPF,y} = 1.12$.

A.1.3 Federal Corporate Income Taxes

Our corporate tax analysis is performed identically to the personal income tax. Given

$$\eta_{TCF,y} = \eta_{TCF,B}\eta_{B,y}, . \tag{A.4}$$

we estimate $\eta_{TCF,B} = 1.0$ and $\eta_{B,y} = 3.69$, yielding $\eta_{TCF,y} = 3.7$.

A.1.4 Other Federal Taxes and Aggregate Elasticity

Our measure of gross federal receipts in Federal Reserve Greenbook forecasts includes other federal taxes that are not included above, but which constitute on average almost 10 percent of revenues. We assume a within-period elasticity of 1.0 for these other taxes with respect to output.

The weighted average within-period elasticity of federal taxes with respect to GDP, $\eta_{\tau,y}$, is estimated to be about 1.18.

A.2 RSQE

The primary differences between our measure of receipts in RSQE forecasts and our Greenbook measure are 1) the presence of state and local revenues in RSQE forecasts, and 2) RSQE receipts are net receipts, so they include transfers such as social security payments. We define net receipts as follows:

$$NR_t = GFR_t + GSLR_t - GTRF_t - GTRSL_t, \tag{A.5}$$

where GFR_t = gross federal receipts (personal income taxes, corporate income taxes, indirect business taxes, social insurance taxes, household transfers to the federal government, and business transfers to the federal government);[21] $GSLR_t$ = gross state receipts (personal income taxes, corporate income taxes, indirect business taxes, social insurance taxes, and federal aid to states); $GTRF_t$ = federal transfers (government transfers to persons and federal aid to states); and $GTRSL_t$ = state and local transfers.

In addition, our RSQE forecasts range from 1983-2015 rather than 1978-2015, so some minor differences arise from this change of sample. The three federal taxes above, as well as state and local personal income taxes (TPSL), are calculated in the same manner as for the Greenbooks, and they yield the following estimated within-period elasticities as follows:

$$\eta_{TPF,y} = 0.45$$

$$\eta_{TSIF,y} = 0.40$$

$$\eta_{TCF,y} = 3.63$$

$$\eta_{TPSL,y} = 0.21$$

[21] Before 2004, RSQE did not itemize forecasts of household and business transfers to the federal government.

A.2.1 Indirect State and Local Business Taxes

Indirect business taxes mostly consist of sales taxes collected from retailers, but paid to businesses by consumers as a proportional tax on some subset of final goods. Blanchard and Perotti use an elasticity of 1.0 for this category, while noting that some goods are exempt from sales taxes. We estimate the elasticity using the BEA measure of indirect business taxes collected by state and local governments (TIBSL), and find $\eta_{TIBSL,y} = .61$.

A.2.2 Transfers and Aggregate Elasticity

Transfers enter into the average elasticity equation as a negative share, and their respective within-period elasticities span a large range. Social security benefits, which are a large portion of transfers, likely have little-to-no relationship with contemporaneous changes to output, while unemployment benefits react quite strongly (and inversely) to GDP. We use -0.2 as our elasticity of total transfers to GDP, following Blanchard and Perotti in using OECD estimates. Given net receipts (NR) is equal to gross receipts (GR) less transfers (TR),

$$\eta_{\tau,y} = \eta_{NR,y} = \eta_{GR,y}\sigma_{GR,NR} - \eta_{TR,y}\sigma_{TR,NR}, \tag{A.6}$$

where $\sigma_{GR,NR}$ is the share of gross receipts to net receipts and $\sigma_{TR,NR}$ is the share of transfers to net receipts.

The weighted average within-period elasticity of government net receipts with respect to GDP, $\eta_{\tau,y}$, is estimated to be about 1.75.

A.3 The Tax Elasticity Restriction

From section 2, the system of equations 9-11 combined with identifying restrictions $\alpha_{11} = 0$ and $\alpha_{22} = 0$ yields the following:

$$u_t^\tau = \alpha_{12}\varepsilon_t^y + \varepsilon_t^\tau \tag{A.7}$$

$$u_t^g = \alpha_{21}\varepsilon_t^\tau + \varepsilon_t^g \tag{A.8}$$

$$u_t^y = \alpha_{31}\varepsilon_t^\tau + \alpha_{32}\varepsilon_t^g + \varepsilon_t^y \tag{A.9}$$

The additional restriction $\eta_{\tau,y} = \eta$ allows us to estimate this system of equations uniquely. If forecast errors u_t are transformed to represent growth rate errors, then $\eta_{\tau,y} = \frac{du_t^\tau}{du_t^y} = \frac{du_t^\tau}{d\varepsilon_t^y} = \alpha_{12} = \eta$.