ESSAYS IN FISCAL POLICY

by

JEAN E. FALCONER

A DISSERTATION

Presented to the Department of Economics
and the Graduate School of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

June 2018

DISSERTATION APPROVAL PAGE

Student: Jean E. Falconer

Title: Essays in Fiscal Policy

This dissertation has been accepted and approved in partial fulfillment of the requirements for the Doctor of Philosophy degree in the Department of Economics by:

Jeremy Piger	Chair
Mark Thoma	Core Member
Bruce McGough	Core Member
David Guenther	Institutional Representative

and

Sara D. Hodges	Interim Vice Provost and Dean of the Graduate School

Original approval signatures are on file with the University of Oregon Graduate School.

Degree awarded June 2018

DISSERTATION ABSTRACT

Jean E. Falconer

Doctor of Philosophy

Department of Economics

June 2018

Title: Essays in Fiscal Policy

The subject of this dissertation is fiscal policy in the United States. In recent years the limitations of monetary policy have become more evident, generating greater interest in the use of fiscal policy as a stabilization tool. Despite considerable advances in the fiscal policy literature, many important questions about the effects and implementation of such policy remain unresolved. This motivates the present work, which explores both topics in the chapters that follow.

I begin in the second chapter by estimating Federal Reserve responses to changes in taxes and spending. Monetary responses are a critical determinant of fiscal policy effectiveness since central banks have the ability to offset many of the economic changes resulting from fiscal shocks. Using techniques commonly employed in the fiscal multiplier literature, my results indicate a willingness by monetary policymakers to alter policy directly in response to fiscal shocks in a way that either reinforces or counteracts the resulting effects.

In the third and fourth chapters I shift my focus to the conduct of fiscal policy. Specifically, I use Bayesian methods to estimate the response of federal discretionary policy to different macroeconomic variables. I allow for uncertainty about various characteristics of the underlying model which enables me to

determine, for example, which variables matter to policymakers; whether policy conduct has changed over time; and whether policy responses are state dependent. My results indicate, among other things, that policy responds countercyclically to changes in the labor market, but only during periods of weak economic activity.

CURRICULUM VITAE

NAME OF AUTHOR: Jean E. Falconer

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon, Eugene, OR
Middlebury College, Middlebury, VT

DEGREES AWARDED:

Doctor of Philosophy, Economics, 2018, University of Oregon
Master of Science, Economics, 2014, University of Oregon
Bachelor of Arts, International Politics and Economics, 2010, Middlebury College

AREAS OF SPECIAL INTEREST:

Fiscal and Monetary Policy
Time Series Econometrics
Bayesian Econometrics

GRANTS, AWARDS AND HONORS:

Graduate Employee Teaching Award, University of Oregon, 2017
Kleinsorge Summer Research Award, University of Oregon, 2017
Graduate Teaching Fellowship, University of Oregon, 2014-2018

ACKNOWLEDGEMENTS

I am grateful for the advice and guidance I have received from Jeremy Piger, Mark Thoma, Bruce McGough, and David Guenther. This dissertation also benefits from helpful comments courtesy of seminar participants at the University of Oregon's Macroeconomics Group. On a more personal note, I would like thank Elaine Lui, Bruce and Sharon Falconer, and Will Wheeler.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Table | Page

CHAPTER I

INTRODUCTION

For many years, empirical work on fiscal policy lagged far behind similar work on monetary policy. However, the creation of the Eurozone in the 1990s removed members' access to independent monetary policy while the zero lower bound episode that began in the United States in December of 2008 highlighted the limits of such policy. As a result, fiscal policy has received a great deal of attention in recent academic work.

The dominant theme of my dissertation is fiscal policy in the United States. In Chapter II I focus on Federal Reserve *responses* to fiscal policy, which I see as being a critical determinant of fiscal policy effectiveness. I then shift my focus directly to the conduct of fiscal policy. In Chapter III I examine which variables U.S. federal policy responds to as well as the sign and magnitude of policy responses. Chapter IV extends Chapter III by focusing on the possible state dependence of policy responses.

I begin in Chapter II by studying central bank responses to fiscal policy, investigating whether the Federal Reserve alters the federal funds rate directly in response to changes in taxes and spending. Surprisingly, interactions between fiscal and monetary policy have received little attention in empirical work. I estimate *direct* monetary policy responses to fiscal policy by identifying tax and spending shocks from a structural VAR and calculating two sets of impulse responses, one representing the actual federal funds rate response to fiscal policy shocks and the other representing the same response when the Federal Reserve is prevented from responding directly to fiscal policy. The difference between the two sets of impulse responses is the direct response. A nonzero direct response indicates that the Fed responds differently to changes in the macroeconomy when the source of those changes is fiscal policy.

My results suggest that the Federal Reserve responds directly to fiscal shocks, and that its response depends on the type of shock. Specifically, the FOMC *reinforces* the anticipated effects of spending shocks, for example decreasing the federal funds rate in response to expansionary spending shocks. In contrast, the Federal Reserve *impedes* tax shocks, for example increasing the federal funds rate in response to expansionary tax shocks. The direct monetary response to expansionary shocks is larger and more persistent than the response to contractionary shocks, as is the response to local spending relative to federal spending. My results also indicate that the direct response to tax shocks has declined in recent years while the response to spending is largely unchanged.

In Chapter III I focus on the conduct of federal discretionary fiscal policy and its response to the business cycle. Unlike monetary offset, there is an established literature on this topic. Nonetheless, findings vary widely, with different studies concluding that policy is procyclical, countercyclical, and acyclical, even within similar sets of countries and time periods. These conflicting results may be due in part to uncertainty about which covariates belong in the true model of fiscal policy conduct. I approach the subject of fiscal policy cyclicality from a Bayesian perspective, incorporating model uncertainty through the use of Bayesian model averaging (BMA). BMA proceeds by averaging estimates across a large number of models, weighting results from each according to how well the corresponding model fits the data. As a result, estimates are robust to a large number of specifications.

My results suggest a number of things about federal discretionary policy. First, policymakers are much more likely to respond to the change in the unemployment rate than output-based measures like the output gap. This is a striking result given that, to my knowledge, I am the first to consider employment-based measures in a cyclicality model. Second, overall policy is countercyclical, although countercyclical policy is limited to recessions; during expansions, in contrast, policymakers are unlikely to respond to economic conditions. Distinguishing between taxes and spending makes it clear that both are changed in response to the business cycle, although the magnitude of the tax response is larger than the magnitude of the spending response. Finally, I find no evidence of a structural break in business cycle responses, nor of substantive differences between intended policy and actual policy outcomes.

Finally, Chapter IV is motivated by the observation that the sign and magnitude of federal policy responses may differ depending on certain macroeconomic characteristics like the business cycle phase or level of debt. There is some evidence to suggest that this is the case. However, previous findings of state dependence in fiscal policy conduct rely on strong assumptions about the structure of the underlying model. I use Bayesian model comparison techniques to compare linear and threshold models that differ with respect to the included covariates and, for threshold models, the threshold variable and value that govern regime switching. I consider twenty-nine possible threshold variables related to the business cycle, debt, monetary policy, inflation, and political environment. My results provide strong evidence that discretionary fiscal policy is regime dependent. Among the threshold variables that I consider, the change in nonfarm payrolls and the change in the federal funds rate are the most probable. Threshold value posteriors for both of these threshold variables split observations into regimes consistent with periods of "normal" and "bad" economic performance that roughly coincide with recession dates. As in Chapter III, coefficient estimates indicate that policy is more countercyclical during the "bad" regime than during the "normal" regime.

CHAPTER II

DIRECT MONETARY RESPONSES TO FISCAL SHOCKS

1. Introduction

When central banks have the ability to change their policy interest rates, the scope for fiscal policy as a stabilization tool may be limited since the effects of such policy could be offset. That is, a tax or spending shock that causes inflation and output to deviate from their target levels may prompt the central bank to counteract these effects through changes in its policy interest rate, as long as it is capable of doing so. In December of 2008, the federal funds rate hit the zero lower bound, limiting not only the Federal Reserve's ability to stabilize the U.S. economy, but also its ability to offset changes in fiscal policy. The greater potential for nonzero fiscal multipliers coupled with the passage of stimulus packages like the American Recovery and Reinvestment Act led to greater interest in fiscal policy by academics and policymakers. Since then, research on fiscal policy has continued in earnest, although the macroeconomic effects of such policy are still hotly debated in the literature.

Despite greater interest in the use of fiscal policy as a stabilization tool, interactions between fiscal and monetary policy have received little attention in empirical work even though the effects of fiscal policy depend critically on how the central bank responds. This knowledge gap motivates the present paper, which considers Federal Reserve responses to fiscal policy in the United States. Specifically, I estimate *direct* responses of the federal funds rate to tax and spending shocks.

I define an indirect response as the change in the federal funds rate in response to the anticipated *effects* of a fiscal shock. The direct response, then, is defined as any additional response to the shock itself. A nonzero direct response indicates that the Fed responds differently to changes in the macroeconomy when the source of those changes is fiscal policy. The sign of the direct response indicates whether the Fed *reinforces* or *impedes* fiscal policy. Monetary policy will be said to reinforce fiscal policy if the direct response to an expansionary fiscal shock is expansionary and will be said to impede fiscal policy if the direct response is contractionary. In other words, if the federal funds rate rises by less (or falls by more) than is justified by the effects of an expansionary fiscal shock on the macroeconomy, monetary policy will be said to reinforce that fiscal shock.

Consider direct and indirect monetary responses in the context of a simple Taylor rule:

$$ffr_t = a_0 + a_1(\pi_t^e - \pi^*) + a_2(y_t^e - y_t^*) + a_3\tau_t + a_4 g_t + u_t$$

where ffr_t is the federal funds rate in period t, π_t^e and y_t^e are forecasted inflation and output in period t, π^* is the central bank's inflation target, y_t^* is potential output, and τ_t and g_t are taxes and spending. The indirect response to a fiscal shock is captured by a_1 and a_2: a shock to spending or taxes affects the inflation and output gaps, which then cause the federal funds rate to change. A direct response is represented by a_3 and a_4.

I use a vector autoregression (VAR) to estimate direct monetary responses to spending and tax shocks, which are calculated as the difference between two impulse responses. The first represents the actual evolution of the federal funds rate following a fiscal shock, while the other represents how it would evolve if it could respond only to the effects of the shock and not the shock itself. The inclusion of Greenbook forecasts in the VAR is critical because they represent the FOMC's information set. Controlling for these forecasts, any direct response of the federal funds rate can be interpreted as a response different from what the FOMC itself believes is justified by the forecasted effects of fiscal policy on the economy.

My results suggest that the Federal Reserve responds directly to fiscal shocks, and that its response depends on the type of shock. Specifically, it reinforces the anticipated effects of a spending shock but impedes the anticipated effects of a tax shock. I also find evidence of other asymmetries. The direct response to expansionary shocks is much larger and more persistent than the response to contractionary shocks, which disappear after about three years. The same is true of local spending shocks relative to federal spending shocks, but there is not a similar distinction between defense and nondefense shocks. Since the Great Moderation the direct response to tax shocks may have become smaller, but the response to spending shocks does not appear to have changed much.

The rest of the paper proceeds as follows. Section 2 summarizes related literature while Sections 3 and 4 discuss methodology and data. Section 5 presents results from the main specification before discussing robustness checks and asymmetry in responses to the subcomponents of the tax and spending variables. Section 6 concludes.

2. Related Literature

Until recently, empirical work on fiscal policy focused almost exclusively on output responses to fiscal shocks. Consequently the available evidence on central bank reactions to fiscal policy is relatively scarce and far from conclusive.

Monetary policy, when mentioned, is typically offered as a possible explanation of results rather than being of primary interest in and of itself. Indeed, many of the papers discussed below do not include the central bank policy interest rate in their analysis, which is arguably a better indicator of policy stance than other short-term interest rates. Consensus on the relationship between fiscal and monetary policy is further complicated by the proliferation of government spending and tax variables, which in some cases differ greatly from each other.

Despite these complications, there is some evidence to suggest that the Federal Reserve does not always try to offset fiscal policy. Romer and Romer (2016) provide anecdotal evidence that, at least in the past, this was the case. Excerpts from FOMC minutes over the period 1951-1992 suggest that policymakers were well aware of changes in taxes and Social Security benefits, and that they had traditional views on their probable economic effects. Nevertheless, in some instances Federal Reserve staff expressed a desire to refrain from or delay in responding to tax changes. The decision to offset policy changes depended on whether the aggregate demand effects of the policy were intentional or unintentional. For example, in February 1964 one member made the comment that "the stimulative effect of a tax cut, which was being counted on so heavily by the American people should not be offset by the System until such action was obviously necessary" (Minutes, 2/11/64, p. 48). Romer and Romer (2016) present other similar statements in their paper. In contrast, policymakers felt the need to counteract Social Security benefit increases immediately to prevent them from having unintended expansionary effects on the economy.

Early empirical work on central bank responses to fiscal policy comes from Muscatelli, Tirelli, and Trecroci (2002). Their focus, whether monetary and fiscal policy offset or accommodate each other's shocks, is similar to that of the present paper. The authors find that the federal funds rate decreases following a fiscal expansion, accommodative behavior that is in line with the narrative evidence presented above. This relationship holds both before and after 1979, a notable date in U.S. monetary history.

Although suggestive, the authors do not distinguish between direct and indirect federal funds rate responses, making it impossible to determine whether the federal funds rate falls due to the macroeconomic effects of an expansionary shock or due to the shock itself. Furthermore, the use of total budget deficits as the fiscal policy variable prevents one from distinguishing between automatic and discretionary fiscal policy. Finally, the authors do not distinguish between taxes and government spending, which is standard in more recent work.

Romer and Romer (2016) compare the macroeconomic effects of a narrative tax series they created with permanent increases in Social Security benefits. Following a tax cut, they observe a small (but significant) negative response of the nominal federal funds rate, which persists for about twelve quarters. Perotti (2004) obtains similar results for both nominal and real long-term interest rates, although these responses represent the average response across five OECD countries. Taken together, these results again indicate a willingness of central banks to lower the federal funds rate in support of expansionary fiscal policy.

In contrast, Guajardo, Leigh, and Pescatori (2014) find that central bank policy rates decline in response to tax increases. Since they use a linear model, this implies an *increase* following a tax cut and thus offsetting behavior. However, their results may not be directly comparable to those of Romer and Romer (2016) and Perotti (2004) since their data includes only contractionary changes in fiscal policy.

Additionally, their data covers seventeen OECD countries, many of which have not had independent central banks for some time.

As with taxes, the evidence on central bank responses to government spending is mixed. Some authors have found evidence of accommodative behavior by the Federal Reserve. In Ramey (2011), short-term nominal interest rates decrease following a positive shock to a defense news variable that measures the expected discounted value of changes in defense spending. Similarly, Perotti (2014) distinguishes between U.S. civilian and defense spending shocks and finds a negative interest rate response to a positive shock to either of these variables.

These results are consistent with an accommodative Federal Reserve although they should be met with caution. Ramey's defense variable has been criticized for having very little explanatory power when applied to data that excludes World War II and the Korean War, periods when things like patriotism, rationing, and price controls may have complicated the relationship between macroeconomic variables (Perotti, 2011). Indeed, Ramey herself suggests that the behavior of interest rates in her set-up may be driven by wartime central bank policies.

Perotti's earlier work using a single government spending variable and data from five OECD countries contradicts his 2014 findings (Perotti, 2004). In that paper, long-term nominal and real interest rate responses suggest offsetting behavior by central banks, increasing in response to a positive spending shock. Canova and Pappa (2011) observe a similar interest rate response in the U.S., although this response is not remotely significant. Nonetheless, these findings are more in line with the traditional assumption that central banks counteract the expansionary effects of government spending increases.

The extent to which the Federal Reserve offsets changes in fiscal policy may depend on the means by which policy changes. Perotti (2004) and Guajardo, Leigh, and Pescatori (2014) both consider asymmetric responses to spending and taxes, although their results differ, possibly due to the fact that the latter only consider contractionary policy changes. Perotti (2004) finds evidence of offsetting behavior following expansionary spending shocks but not expansionary tax shocks. Specifically, the real interest rate response is positive four and twelve quarters after a spending increase but negative after a tax cut. Guajardo, Leigh, and Pescatori (2014) detect offsetting behavior in response to a contractionary shock of either kind, although central bank policy rates fall more in response to spending-based shocks than tax-based shocks.

Perotti (2014) looks at asymmetric responses by the Federal Reserve to civilian and defense spending. His results indicate a negative response of short-term interest rates to a spending increase of either kind, but this negative response persists only for defense shocks. For a civilian spending shock, the interest rate becomes positive after about ten quarters. Perotti does not provide an explanation for his results, although his findings suggest that the Federal Reserve is less likely to offset spending related to military activity.

Romer and Romer (2016) find a difference in how the Federal Reserve responds to Social Security benefit increases relative to tax changes. They show that the federal funds rate increases dramatically following a Social Security benefit increase but falls slightly after a tax cut, despite the fact that both are predicted to have positive output effects. This fits with FOMC excerpts that suggest that members felt it prudent to respond immediately to changes in Social Security but not taxes.

Finally, Belinga and Ngouana (2015) find evidence that Federal Reserve responses to spending shocks are regime-dependent. In their set-up, the two regimes are "accommodative" and "non-accommodative" monetary policy, determined by deviations from an estimated policy reaction function. Unsurprisingly, they conclude that in response to a positive spending shock the federal funds rate declines in the accommodative regime but increases in the non-accommodative regime. This indicates that there may be times when the Federal Reserve wishes to counteract spending shocks and times when it does not.

In sum, whether, when, and to what extent the Federal Reserve offsets fiscal shocks is far from clear. Importantly, none of the papers discussed above distinguishes between direct and indirect responses to fiscal shocks, making it impossible to determine whether the federal funds rate response is driven solely by the economic consequences of a shock or if there is a direct response to the shock. However, there is sufficient anecdotal and empirical evidence to suggest that the subject warrants further investigation.

3. Methodology

As defined, a direct response to fiscal policy means that the Federal Reserve responds to spending and tax shocks differently from what is warranted by the predicted effects of such shocks on the economy. Consequently it can be thought of as the difference between two impulse responses. The first, an unrestricted impulse response, represents the actual evolution of the federal funds rate following a shock. If fiscal shocks affect macroeconomic variables of interest to the Federal Reserve, this impulse response does not allow us to determine whether the estimated changes in the federal funds rate are due to changes in GDP and inflation *caused* by fiscal shocks or whether there is an additional response to the shocks themselves. That is, the unrestricted impulse response is the sum of the direct and indirect responses to fiscal policy.

The other impulse response, a restricted impulse response, shows the progression of the federal funds rate when it is prevented from responding directly to fiscal shocks. This counterfactual illustrates the indirect monetary response to changes in spending and taxes that would occur if the Federal Reserve responded solely to the macroeconomic effects of fiscal policy. Once the restricted and unrestricted impulse responses are obtained, the direct monetary response can be calculated as the difference between the two.

The following section outlines the empirical strategy used to estimate the restricted and unrestricted impulse responses. Since my interest centers on Federal Reserve responses to fiscal policy (and not vice versa), the methods I use are motivated by common approaches in the fiscal policy literature. Specifically, I incorporate narrative tax shocks into a structural VAR identified from a recursive causal ordering.

3.1 Standard Fiscal Variables and the Narrative Approach

Blanchard and Perotti's (2002) seminal paper on the macroeconomic effects of fiscal policy established VARs as the paradigm for estimating the effects of such policy. The tax and spending variables they use in that paper also became the standard definitions used in subsequent fiscal policy research. However, in recent years the availability of alternative measures has reduced reliance on these potentially flawed variables. I will discuss each in turn.

In Blanchard and Perotti (2002), government spending is defined as the sum of state and federal government consumption expenditure and gross investment. Government consumption expenditure consists of spending on public goods and services like education and national defense, while gross investment consists of spending on infrastructure like highways[1]. The primary concern about government expenditure and gross investment is that it is anticipated (a concern that applies equally to taxes). Since legislated changes in spending are often implemented with a delay, economic agents may respond to the *announcement* of a spending shock that precedes any real change in spending. In fact, if the present discounted value of spending is the only thing that matters to economic agents (as is the case in the neoclassical model), the timing of actual changes in spending should be irrelevant. Ramey (2011) provides evidence that shocks to government expenditure and gross investment are anticipated, finding that they are Granger-caused by professional forecasts.

Taxes are commonly measured as tax revenues less transfer and interest payments[2]. Since transfer and interest payments fluctuate with output, this measure is often called cyclically-adjusted tax revenues, although such terminology is misleading. Crucially, it does not account for the fact that changes in output also affect the tax base and therefore tax revenues[3]. Nonetheless, it is argued that once revenues are cyclically-adjusted, any further change should reflect changes in discretionary tax policy.

---

[1] see http://www.bea.gov/national/ for more information.

[2] Transfer payments are counted as negative taxes rather than positive spending.

[3] To account for this, Blanchard and Perotti (2002) use external information to calculate the output elasticity of taxes, which they then impose as a restriction in their VAR. Caldara and Kamps (2008) find that the estimated output responses to taxes depend enormously on this elasticity, which is why I use a different approach.

The narrative approach developed largely in response to the concern that the traditional tax variable does not accurately reflect discretionary policy. It relies on using historical documents to construct measures of taxes and spending. Romer and Romer's (2010) tax series and Ramey's (2011) spending series are the primary series used in the fiscal literature, although others (Guajardo, Leigh, and Pescatori (2014), for example) have created similar series for other OECD countries.

Romer and Romer (2010) use historical documents to create a tax variable that measures legislated changes in tax liabilities, thus ensuring that it reflects actual changes in policy. The series is also constructed to be orthogonal to changes in GDP and other factors that "push growth away from normal." Specifically, they define a tax change as exogenous if the stated motivation for the change is higher long-run growth or to reduce an inherited budget deficit. They argue that these types of tax changes are exogenous because they are unlikely to be related to current or future economic conditions.

Romer and Romer did not consider possible anticipation effects in creating their tax series which, as mentioned, may yield incorrect results if responses to fiscal policy begin at the announcement of the policy rather than its implementation date. Mertens and Ravn (2012) find evidence of anticipation effects using Romer and Romer's tax measure. They note that excluding tax changes whose legislation and implementation did not occur within the same quarter reduces the estimated effects of tax changes on GDP by a considerable amount.

Ramey's (2011) defense news variable is constructed in a similar manner to the Romer and Romer series. It measures the present discounted value of changes in government spending motivated by military concerns. Importantly, it does not reflect the timing of actual changes in spending but rather announcements of such spending.

There are a number of arguments against the use of Ramey's defense news variable. As she herself discusses in her paper, the variable does a poor job of predicting actual changes in defense spending when World War II and the Korean War are excluded from the sample. Similarly, as suggested by Perotti (2014), it is unclear what a shock to a news variable even represents. Finally, Perotti (2014) demonstrates that two VARs with and without the defense news variable that are otherwise identical produce nearly identical results[4]. This suggests that although anticipatory effects may be problematic in theory, they do not appear to matter much in practice.

For estimation, I use Mertens and Ravn's version of Romer and Romer's tax series since my results change depending on which version is used, possibly because of anticipation effects. However, due to the objections raised above I do not use Ramey's defense news variable as my primary spending variable. Instead, I use the more common government expenditure and gross investment. Augmenting my baseline VAR with Ramey's defense news variable has no effect on the results. To estimate possible asymmetric responses to different types of fiscal shocks, I also divide taxes and spending into their subcomponents.

[4]Ramey (2011) estimated very different results using the same setup, but that was because she was not comparing the same shock for each. Rather, she estimated the effects of an actual spending shock in one VAR and the effects of a defense news shock in the other.

3.2 The VAR Approach

The reduced-form VAR can be written as

$$Y_t = \alpha + \pi(L)Y_{t-1} + u_t$$

where Y_t, α, and u_t are m-dimensional vectors of endogenous variables, constants, and reduced-form residuals, and $\pi(L)$ is a lag polynomial. The included variables are first-differenced since a Dickey-Fuller test indicates that spending has a unit root. I include eight lags because the errors are autocorrelated when I use fewer.

Since the reduced-form VAR residuals are likely correlated with each other, they do not have an economic interpretation. However, we can write down a structural VAR whose errors can be given such an interpretation:

$$AY_t = \mu + \phi(L)Y_{t-1} + \epsilon_t$$

Matrix A describes the contemporaneous relationships between the included endogenous variables while $\mu = A\alpha$ and $\phi(L) = A\pi(L)$. Written in this way, the relationship between reduced-form residuals u_t and structural shocks ϵ_t can be expressed as

$$u_t = A^{-1}\epsilon_t$$

The contemporaneous effect of structural shocks is determined by matrix A while their propagation is determined by the VAR reduced-form coefficients. Consequently identification of the system hinges on identifying the elements of A. Assuming that the structural shocks are uncorrelated with each other and scaling them to have unit variance (which makes variance-covariance matrix Σ_ϵ the identity matrix), it is easy to see that

$$\Sigma_u = QQ'$$

where $Q = A^{-1}$. This provides some identification restrictions but without further restrictions the model remains unidentified.

A number of different identification schemes have been proposed in the literature, including a recursive causal ordering of the variables, the use of external information to estimate specific parameters of A (this is the approach taken by Blanchard and Perotti (2002)), restrictions on the signs of impulse responses, partial identification of matrix A, and the use of narrative shocks that are constructed to be exogenous. I combine a recursive causal ordering with the use of narrative shocks.

A recursive causal ordering restricts matrix A to be lower-triangular, which implies assumptions about contemporaneous relationships between variables. Specifically, variables ordered higher in the system do not respond to variables ordered below them within a period. Once these restrictions are imposed, a unique matrix A can be obtained from a Cholesky decomposition of Σ_u.

In the main specification I order the variables as follows: taxes, government spending, GDP forecast, inflation forecast, and the federal funds rate[5]. This ordering implies that the tax variable does not respond contemporaneously to any other variable, and that government spending responds only to taxes within

[5]Although I include current values of GDP and inflation as a robustness check, I exclude them from the baseline specification since lags of forecasts should control for them. My results are largely unchanged when they are included.

a quarter. Ordering the fiscal variables at the top can be justified on the grounds that the tax series is constructed to be independent of other macroeconomic variables, while the legislative process is such that discretionary changes in government spending in response to economic conditions are implemented with a lag. The ordering of the other variables assumes that forecasts and the federal funds rate respond to fiscal policy contemporaneously but that forecasts do not respond contemporaneously to the federal funds rate. As explained in the following section, the federal funds rate data was constructed with this restriction in mind. It turns out that variable order matters little for my results, which are virtually identical as long as spending and taxes are ordered first in the system (with either taxes or spending as the first variable) and qualitatively the same when these variables are ordered last.

To construct the unrestricted impulse responses, I use impact matrix A and the reduced-form VAR coefficients $\pi(L)$ to simulate the effects of a one-standard deviation structural shock on the endogenous variables in the system. I construct the restricted impulse responses in a similar manner but prevent monetary policy from responding directly to fiscal policy by setting the relevant elements of A and $\pi(L)$ equal to zero. For example, in the baseline model elements (5,1) and (5,2) of matrix A measure the contemporaneous response of the federal funds rate to tax and spending shocks, while the same elements in each $\pi(L)$ determine how the federal funds rate responds to taxes and spending after the initial shock. Setting these elements equal to zero prior to simulating the effects of a shock removes the direct monetary response from the impulse response.

The use of forecasts is important to identifying a direct response to fiscal policy since delays in the effects of monetary policy mean that the Federal Reserve responds to both current and future economic conditions. Even more important, however, is the use of *Greenbook* forecasts. Since these forecasts are produced by Federal Reserve staff for the FOMC, they reflect policymakers' information about the probable effects of fiscal policy on the economy. Consequently a nonzero direct response estimated using Greenbook forecasts suggests that the FOMC knowingly responds directly to a fiscal shock. The use of alternative forecasts of GDP and inflation may yield similar but inaccurate conclusions. In that case, two possibilities may explain a nonzero direct response. First, the FOMC may correctly forecast the effects of a fiscal shock and deliberately respond differently than what is warranted by the shock. However, it could also be that the FOMC's information set is different from the information contained in the alternative forecasts. Consequently even if its goal is to merely offset the effects of a shock, it will appear to respond directly to it. The use of Greenbook forecasts reduces the likelihood of this second possibility[6].

4. Data

For most specifications I use quarterly data covering 1967q1-2007q4, although when forecasts are excluded I extend the sample to 1955q1 as a robustness check. Government spending, GDP, and the GDP deflator come from the Bureau of Economic Analysis NIPA tables, while the civilian noninstitutional population data used to transform variables comes from the Bureau of Labor Statistics. The modified Romer and Romer tax series was obtained from Karel Mertens's website[7]

[6] Even using Greenbook forecasts, a nonzero direct response may be the result of offsetting behavior by policymakers whose beliefs differ from Greenbook forecasts rather than a deliberate direct response to a shock. This is one of the possible explanations for a direct response offered in the introduction.

[7] https://mertens.economics.cornell.edu/research.htm

while the federal funds rate comes from the Board of Governors of the Federal Reserve System. Finally, Greenbook forecasts come from the Federal Reserve Bank of Philadelphia.

The sample dates were determined primarily by data availability[8] but they have a number of other advantages. First, they exclude notable episodes like World War II and the Korean War when massive increases in defense spending coincided with unusual circumstances (such as rationing and Federal Reserve regulations that discouraged the purchase of durable goods) that may obscure the "typical" relationship between macroeconomic variables [9]. Additionally, the sample excludes the Great Recession when the federal funds rate was at the zero lower bound, which forced the Federal Reserve to resort to unconventional monetary policy. Although the Federal Reserve response to fiscal policy during zero lower bound episodes is in itself an interesting area of research, including these years could, again, make it difficult to assess the typical monetary response to fiscal policy.

The variables are defined as follows: government spending is the sum of state and local defense and nondefense government consumption expenditure and gross investment. As discussed, this definition is common in the fiscal policy literature. Taxes are defined as the estimated revenue effects of legislated tax changes at the time of their implementation. GDP is used to forecast output and the GDP deflator is used for the inflation forecast. The quarterly federal funds rate was created from monthly data to ensure that it can respond to forecasts within a quarter, as implied by its ordering in the VAR. Specifically, it is the nominal federal

[8]Greenbook forecasts for GDP and inflation are available from 1967 onward while the Romer tax series is available until the end of 2007.

[9]See Ramey (2011) and Perotti (2011 and 2014) for further discussion.

funds rate in the month following the Greenbook publication closest to the middle of a given quarter[10].

As mentioned, all variables enter the VAR in first differences. This is done because the spending variable appears to have a unit root, and because the Romer and Romer tax variable is defined as the change in taxes rather than its level. Spending is expressed as the logged first difference of real per capita spending. The GDP forecast is expressed as a growth rate and inflation is the logged first difference of the GDP deflator. The federal funds rate is the first difference of the nominal rate, while the tax variable is expressed in nominal terms as a percentage of nominal GDP.

5. Results

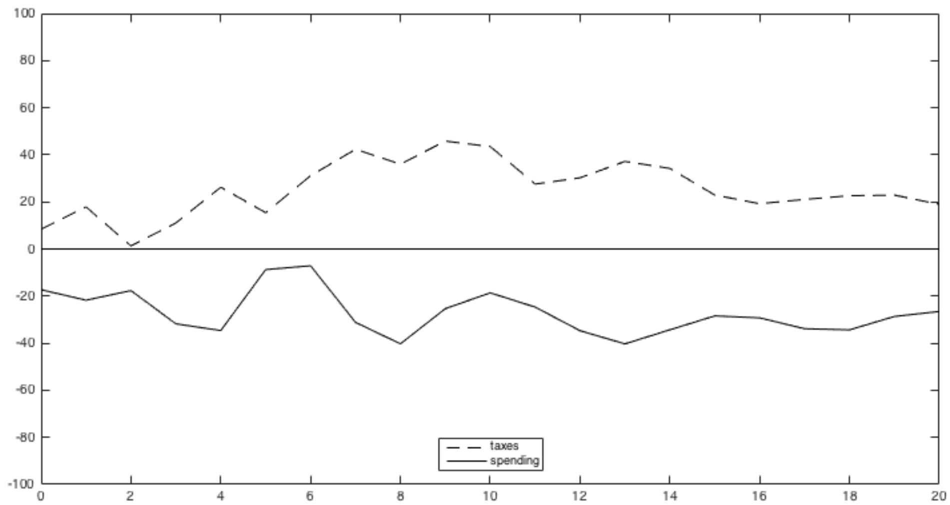
This section presents the direct response of the level of the federal funds rate (in basis points) to a one-standard deviation fiscal policy shock. With one noted exception, I consider only expansionary shocks (spending increases and tax cuts) to facilitate comparison between shock types and specifications. Remember that direct responses do not show how the federal funds rate actually changes following a shock but rather the *difference* between how it changes and how it would change if it could not respond directly to fiscal policy.

5.1 Baseline Model

The results from the baseline model, displayed in Figure 1, indicate nonzero direct responses to spending and taxes. They also suggest a difference in how the

[10] This is similar to the approach taken by Auerbach and Gorodnichenko (2012) to obtain quarterly Greenbook forecasts.

FIGURE 1. Direct Responses to Expansionary Spending and Tax Shocks



Notes: Responses are for the level of the federal funds rate and are expressed in basis points. The direct response is estimated using the baseline model.

Federal Reserve responds to each type of shock. An expansionary spending shock results in a negative direct response while a similar tax shock leads to a positive response. The peak response to spending (taxes) is about -40 (46) basis points and occurs after fourteen (twelve) quarters. The effects persist indefinitely, stabilizing at -27 (23) basis points after about twenty quarters. Note that the linear nature of the VAR implies a positive direct response to a contractionary spending shock and a negative direct response to a contractionary tax shock. Taken together, these results imply that the Federal Reserve reinforces the anticipated effects of spending, easing policy more than necessary in response to an expansionary shock and tightening following a contractionary shock. Conversely, it moves in opposition to the anticipated effects of taxes, lowering the federal funds rate more (or increasing it by less) than warranted by changes in macroeconomic variables following a contractionary tax shock and increasing it (or lowering it by less) after an expansionary shock.

5.2 Robustness

To check the robustness of the results above I estimate a number of variations of the baseline model. For some specifications I add current values of GDP and inflation alongside their forecasted values to see whether their exclusion from the baseline model really is innocuous. I also include current and forecasted values of the unemployment rate since there is some evidence that the Federal Reserve considers unemployment when making policy decisions (Check, 2016). In another specification I use only contemporaneous values of the variables in the baseline model and in yet another I use two-quarter ahead forecasts instead of one-quarter ahead. I also investigate whether accounting for the anticipation of spending shocks or failing to account for the anticipation of tax shocks affects results. To do so I augment the baseline VAR with Ramey's defense news variable in one specification and use the entire Romer and Romer tax series in another.

The results from the baseline model are quite robust to changes in specification. Table 1 summarizes these results. In nearly every case the estimated direct responses are qualitatively the same and quantitatively similar, particularly the responses to spending shocks. Notably, the direct response to a tax cut changes sign when the entire Romer and Romer tax series is used. The estimated response to spending in the VAR with Ramey's defense news variable is the largest of any specification but yields the same conclusions as the baseline model. Thus it appears that accounting for anticipated changes in fiscal policy is critical for taxes but less important for spending.

TABLE 1. Peak and Long-Run Direct Responses from Different Specifications

	Spending		Taxes	
	Peak	**Long-run**	**Peak**	**Long-run**
(1)	-40	-27	46	23
(2)	-38	-30	37	18
(3)	-34	-20	48	26
(4)	-45	-27	52	27
(5)	-37	-20	58	38
(6)	-34	-19	57	36
(7)	-52	-36	47	28
(8)	-44	-25	-24	-9

Notes: Responses are for the level of the federal funds rate and are expressed in basis points. (1) is the baseline model; (2) uses two-quarter ahead forecasts instead of one-quarter ahead; (3) uses contemporaneous values of GDP and inflation instead of forecasts; (4) augments the baseline model with the one-quarter ahead unemployment forecast; (5) includes both current and forecasted values of GDP and inflation; (6) includes both forecasted and current values of GDP, inflation, and unemployment; (7) augments the baseline model with Ramey's defense news variable; (8) uses Romer and Romer's entire tax series instead of just those legislated and implemented within the same quarter.

5.3 Asymmetry

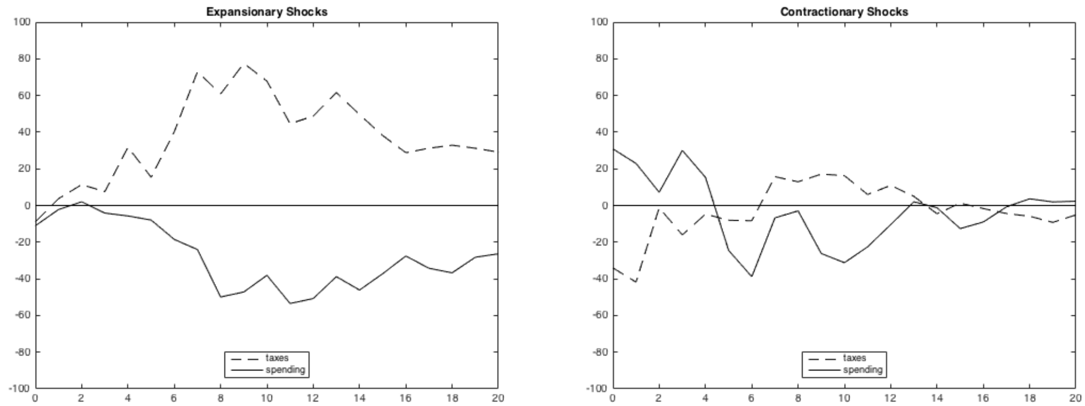
Positive and Negative Shocks

The baseline model does not allow for asymmetric responses to positive and negative fiscal shocks. Yet if Federal Reserve policymakers believe that expansionary and contractionary shocks have different effects on the economy, those beliefs may be reflected in policy. For instance, Guajardo, Leigh, and Pescatori's (2014) finding that central bank policy rates decline more following a contractionary spending shock than a contractionary tax shock could be explained by differences in direct policy responses. However, their setup does not allow one to distinguish between direct and indirect monetary responses, making it impossible to determine whether asymmetries reflect intentional actions by the Federal Reserve or rather differences in the macroeconomic effects of tax and spending shocks that feed back into the interest rate response.

To allow for this possibility, I create separate variables for positive and negative tax and spending shocks and include them all in the baseline VAR. I set the positive tax shock equal to the tax shock when the shock is greater than zero and set the negative tax shock equal to the tax shock when it is less than zero. I use the same strategy for changes in government spending and shock each of the variables to produce the results above. Instead of considering only expansionary shocks, I present results for the same type of shock as the variable considered (i.e. I consider a positive shock to the positive tax variable and a negative shock to the negative tax variable) for ease of interpretation.

Figure 2 displays the results, which show a clear difference in how the Federal Reserve responds to expansionary fiscal shocks relative to contractionary fiscal shocks. By and large, direct responses to expansionary shocks are much larger in magnitude than responses to contractionary shocks, which approach zero after about three years. In contrast, the direct response to expansionary shocks lasts indefinitely. The timing of the responses also differs between the shock types. The direct responses to expansionary shocks are close to zero for the first year while the largest responses to contractionary shocks occur during that time. Finally, unlike the responses to expansionary shocks, the responses to contractionary shocks switch signs after about a year before converging towards zero. The earlier observation that the Federal Reserve reinforces the effects of a spending shock but impedes the effects of a tax shock holds here as well.

FIGURE 2. Direct Responses to Expansionary and Contractionary Fiscal Shocks



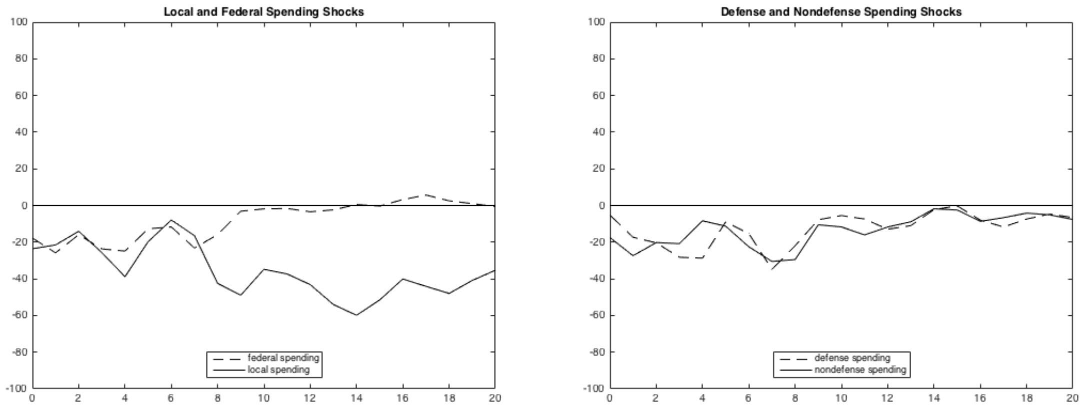
Notes: Responses are for the level of the federal funds rate and are expressed in basis points. The graph on the left illustrates the direct response to expansionary shocks (positive spending and negative taxes) while the graph on the right illustrates the direct response to contractionary shocks (negative spending and positive taxes).

Different Types of Spending

The government spending variable in the baseline model is comprised of state, local and federal spending, while federal spending is itself the sum of defense and nondefense spending. It is possible that different types of spending shocks prompt different Federal Reserve responses. For instance, Perotti's (2014) finding that defense spending is contractionary while nondefense spending is expansionary may be explained by asymmetric monetary responses. To see whether the type of spending matters for the monetary response I estimate two variations of the baseline model, one in which local and federal spending appear as two different variables and another that separates defense and nondefense spending in a similar manner.

The results, shown in Figure 3, indicate distinct direct responses to local and federal spending but not defense and nondefense spending. The direct responses to

FIGURE 3. Direct Responses to Different Types of Spending Shocks



Notes: Responses are for the level of the federal funds rate and are expressed in basis points. The graph on the left illustrates the direct response to expansionary local and federal spending shocks while the graph on the right illustrates the direct response to expansionary defense and nondefense spending shocks.

local and federal spending are nearly identical for two years but then diverge. The response to local spending is permanent, settling around -37 basis points, while the response to federal spending returns to zero after nine quarters. Direct responses to defense and nondefense spending, in contrast, largely move together, following the same trajectory as federal spending (which makes sense given that federal spending is the sum of defense and nondefense spending). Thus it seems as though monetary policymakers reinforce the effects of local spending more than federal spending and do not distinguish between defense and nondefense spending.

Motivation Behind Tax Change

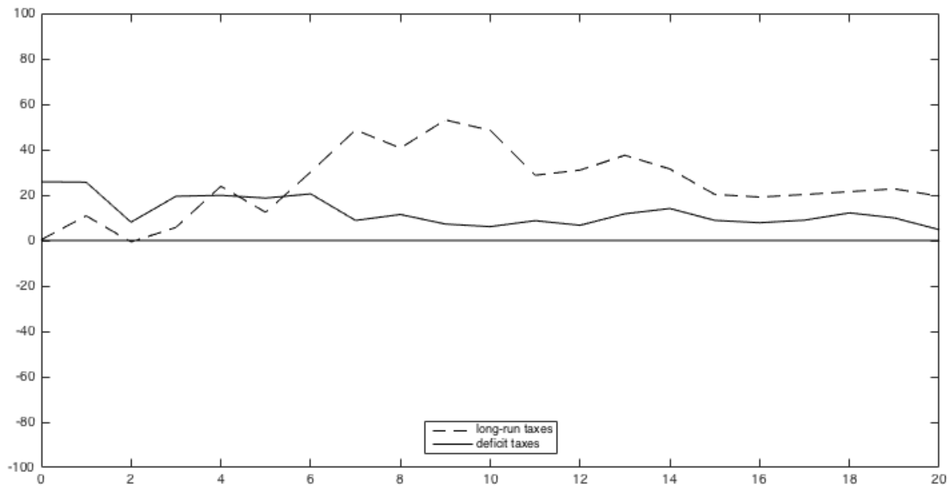
Romer and Romer (2010) classify the motivation behind each tax shock, making it possible to separately estimate direct responses to tax changes motivated by a desire to increase long-run growth and those motivated by a desire to reduce an inherited budget deficit. Figure 4 shows the results of this exercise. The direct monetary response to both kinds of tax shocks are positive, again indicating a direct response that opposes the predicted effects of an expansionary tax shock. However, the direct response to taxes intended to increase long-run growth is for the most part greater than the response to shocks intended to reduce inherited budget deficits. It is worth noting that the mean of the long-run tax shocks is negative while the mean of the deficit shocks is positive. I showed earlier that direct responses to negative tax shocks are larger than responses to positive shocks, so it could be the average signs of the shocks driving the results rather than their motivation.

5.4 Subsample Stability

To see whether the direct monetary response to fiscal policy changed following the Great Moderation, I estimated the baseline specification starting in 1984q1. As Figure 5 indicates, the direct responses to either type of expansionary fiscal shock is roughly zero in the year following the shock. After that, the response to each shock is identical in sign to the full sample, although the tax response is smaller and only temporary, approaching zero after four years. Interestingly, the direct response to spending is similar in magnitude to the baseline VAR and persists indefinitely. These results suggest that the FOMC has moderated its response to taxes in more recent years but continues to reinforce the anticipated effects of an expansionary spending shock, after a delay.

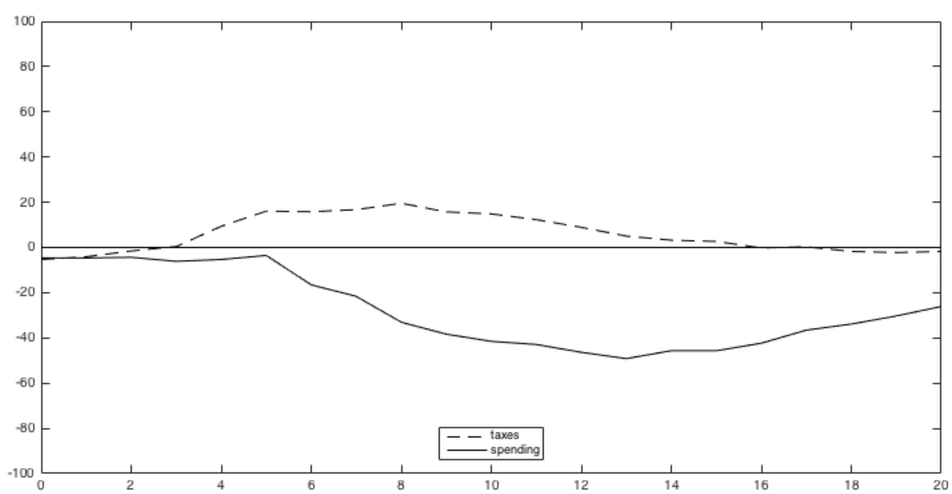
Greenbook forecasts for inflation and GDP are unavailable prior to 1967 so I am unable to estimate the baseline model over an extended sample as a robustness check. However, I estimated the baseline model using contemporaneous values of

FIGURE 4. Direct Responses to Different Types of Tax Shocks



Notes: Responses are for the level of the federal funds rate and are expressed in basis points. The solid line illustrates the direct response to a tax cut motivated by a desire to reduce an inherited budget deficit while the dashed line illustrates the direct response to a tax cut motivated by a desire to increase long-run growth.

FIGURE 5. Direct Responses After 1983



Notes: Responses are for the level of the federal funds rate and are expressed in basis points. Estimated responses to expansionary spending and tax shocks after 1983. The direct response is estimated using the baseline model and a sample that covers 1984q1-2007q4.

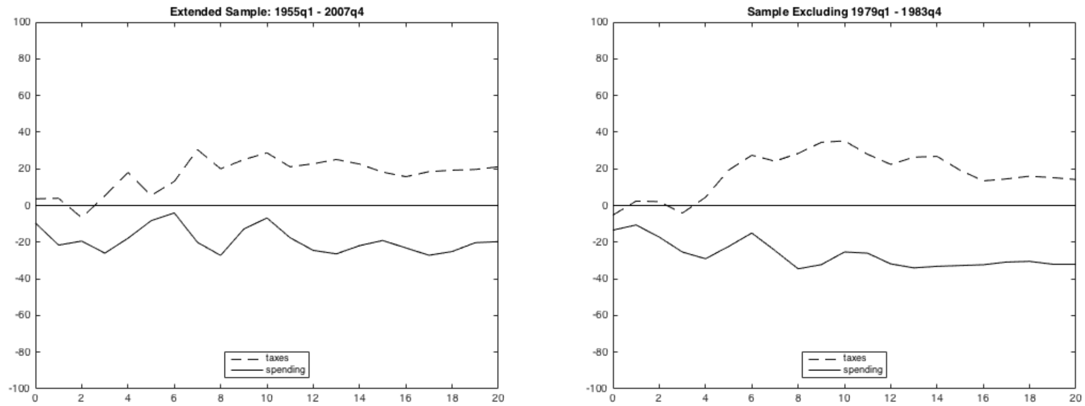
inflation and GDP on the baseline sample as well as an extended sample that goes back to 1955q1. I also removed Paul Volcker's first term as Federal Reserve chair from my baseline sample to see whether my results are driven by unusual monetary policy during that time. As Figure 6 demonstrates, the results from both of these samples are qualitatively similar to those from the baseline specification, albeit smaller in magnitude. Thus it appears my results are robust to changes in sample years.

6. Conclusion

This paper examines whether the Federal Reserve responds directly to fiscal shocks. Using a VAR framework, I estimate the difference between the actual federal funds rate impulse response following a fiscal shock and a counterfactual response in which the central bank is unable to respond directly to fiscal shocks. The inclusion of Greenbook forecasts is critical because they represent the FOMC's information set. Controlling for these forecasts, any direct response of the federal funds rate can be interpreted as a response different from what the FOMC itself believes is justified by the forecasted effects of fiscal policy on the economy.

I estimate a nonzero direct monetary response to fiscal policy, the direction and magnitude of which appears to depend on the type of fiscal shock. Expansionary (contractionary) spending shocks are accompanied by expansionary (contractionary) direct monetary responses, but the converse is true of tax shocks. In other words, the anticipated effects of spending shocks are reinforced while the anticipated effects of taxes are impeded. The estimated direct responses are robust to a number of variations in the baseline model.

FIGURE 6. Direct Responses Estimated Using Different Subsamples



Notes: Responses are for the level of the federal funds rate and are expressed in basis points. Estimated direct response of the federal funds rate to expansionary spending and tax shocks using different samples. The graph on the left illustrates the direct response from a model that uses only contemporaneous values of inflation and GDP as well as an extended sample (1955q1-2007q4). The graph on the right illustrates the direct response from the baseline sample (1955q1-2007q4) when Paul Volcker's first term as Federal Reserve chair is removed.

Further asymmetry can be seen when the baseline variables are disaggregated. Direct responses to expansionary shocks are larger and more persistent than responses to contractionary shocks, indicating that the Federal Reserve offsets the latter but not the former. The direct response to federal spending is small and short-lived while the response to local spending is negative and permanent. The FOMC does not appear to distinguish between defense spending and nondefense spending. There may be some differences in how the Federal Reserve responds to tax changes passed to increase long-run growth relative to changes made to reduce an inherited budget deficit, although this difference could be due to the sign of the shocks rather than their motivation. Finally, it appears that in recent years the direct monetary response to taxes has declined in magnitude and duration while the response to spending is largely unchanged (but delayed).

Any explanation of these results is speculative but there are a number of possibilities:

1. **Direct monetary responses are unintentional.** The FOMC may want to offset all shocks but fail to do so. This could occur if the divine coincidence does not hold or if certain types of shocks are prioritized over others. Similarly, the FOMC may pay greater attention to some shocks than to others.

2. **Direct monetary responses are intentional.** The FOMC may, as suggested by Romer and Romer (2016), be aware of all fiscal shocks and their predicted effects on the economy but choose not to offset them, either because they wish to support the policies of the legislative branch or because of political motivations.

3. **FOMC beliefs are not fully captured by Greenbook forecasts.** Under this scenario, it is impossible to determine whether direct responses are intentional or unintentional.

If direct responses are intentional, an interesting question is whether the decision to offset fiscal shocks depends on the initial state of the economy. For example, the FOMC may choose to offset an expansionary spending shock when output is above potential but not when it is below potential. If shock types depend on the state of the economy, asymmetric responses to these shocks may be driven by the state of the economy rather than the shock type itself. This could, for example, explain why direct responses to expansionary shocks are larger and more persistent than direct responses to contractionary shocks. Future research is needed to determine which explanations are most plausible.

Another interesting question is how Federal Reserve independence influences the sign and magnitude of direct monetary policy responses. One would expect a less independent Fed to engage in more accommodative policy at the behest of politicians interested in winning reelection. Indeed, there is anecdotal evidence to suggest that the FOMC occasionally succumbed to political pressure to lower the federal funds rate during the 1960s and 1970s.[11] To use the terminology introduced previously, one might expect the direct responses implemented by a less independent Fed to be lower (more expansionary) than the direct responses implemented by a more independent Fed.

It is difficult to determine whether Federal Reserve independence, which varied considerably over the course of my 40-year sample, helps explain the results obtained in this paper. Results using the post-1983 subsample indicate that the magnitude of the direct response to spending shocks is much smaller (and less expansionary) in the year following a shock than when results are estimated from the entire sample. Given that Fed independence increased after the 1970s, this is consistent with the idea presented above that direct responses should be smaller and less expansionary when the central bank enjoys greater independence. However, my finding that monetary policy impedes tax shocks is counterintuitive, especially since the direct response to tax shocks is smaller in magnitude (and more expansionary) using the post-1983 subsample. More work is needed to determine whether Federal Reserve independence influences the response of monetary policy to fiscal policy. For example, the estimation of time-varying coefficients (and therefore time-varying direct responses) could reveal differences in policy responses

[11] See "Federal Reserve Independence: The Never-Ending Story" by Mark Thoma at http://www.milkenreview.org/articles/federal-reserve-independence?IssueID=26.

corresponding to periods in which the Fed was more independent and periods in which it was less independent.

CHAPTER III

CYCLICALITY OF U.S. FISCAL POLICY

1. Introduction

Despite some well-known instances of countercyclical discretionary fiscal policy by the United States government, it is far from clear whether such policy is systematic. For instance, the onset of the Great Recession in late 2007 motivated the passage of two stimulus packages, the Economic Stimulus Act and the American Recovery and Reinvestment Act, which were clearly designed to counteract the large decline in output that had occurred. However, even as the economy experienced a tepid recovery and unemployment remained well above the natural rate, political focus shifted towards debt stabilization and a perceived need for fiscal austerity. In January 2013, with unemployment close to 8%, Congress passed the American Taxpayer Relief Act (ATRA) as a partial resolution to the U.S. fiscal cliff. ATRA led to a large increase in tax revenues, projected by the CBO to total 1.5 trillion dollars over the next five years.[1] Two months later, automatic cuts to discretionary spending went into effect as a result of the Budget Control Act of 2011.

Empirical work on fiscal policy has largely failed to yield a consensus on whether and how discretionary policy responds to the business cycle. Different studies have concluded that policy is procyclical, countercyclical, and acyclical, even across similar sets of countries and time periods.[2] Several papers have

[1]See "Estimates of the Budgetary Effects of H.R. 8, the American Taxpayer Relief Act, as passed by the Senate on January 1, 2013" at https://www.cbo.gov/sites/default/files/112th-congress-2011-2012/costestimate/american-taxpayer-relief-act0.pdf

[2]Table 1 in Golinelli and Momigliano (2009) provides a good summary of conflicting results.

investigated the source of these discrepancies, emphasizing the importance of data vintages and accounting for potentially autocorrelated errors.[3] Far less attention has been paid to covariate selection. Unlike the monetary policy literature, in which policy is typically modeled as following a Taylor rule, there appears to be little agreement about how to model the conduct of discretionary fiscal policy. As a result, models can differ substantially from paper to paper. All include some measure of output, but even then there are differences in the specific measure used as well as the time period in which it enters the model.

Table 2 provides a summary of covariates used in some well-known papers in the literature. It suggests that there is considerable uncertainty about the underlying model of fiscal policy conduct. It should be noted that, with the exception of Auerbach (2002 and 2003) and Cohen and Follette (2003), the literature has focused primarily on European fiscal policy. As a result, drawing conclusions about U.S. policy from these papers can be problematic. However, it seems likely that uncertainty about the correct model of fiscal policy extends to U.S. policy as well, particularly since it is less well-studied. Indeed, I find a number of instances in which failing to account for model uncertainty leads to flawed inferences about policy conduct.

Although counterintuitive, acyclical or procyclical policy may occur for a number of reasons. First, policymakers may think it unnecessary to respond to output if they believe the combined responses of monetary policy and non-discretionary fiscal policy can stabilize output without additional aid. Divisions between political parties may also make it difficult to pass legislation unless the situation is urgent, such as during a recession. In either case, we might expect

[3] See, again, Golinelli and Momigliano (2009) as well as Plodt and Reicher (2015).

TABLE 2. Variables Included in Previous Cyclicality Studies

Studies	Business Cycle Measure	Cyclical Period	CAD Lag	Debt	Election Dummy	Inflation	Other
Hallerberg & Strauch (2002)	Output Gap Change	y_t					Time/Country FE
Auerbach (2003)	Output Gap	y_{t-1}					Deficit Lag (Unadjusted) Divided Gov. Dummy President Party Dummy
Cohen & Follette (2003)	Output Gap	y_{t-1}	x				
Gali & Perotti (2003)	Output Gap	y_t y_{t-1}	x	x	x		
Lane (2003)	Output Growth	y_t					
Forni & Momigliano (2004)	Output Gap	$E_{t-1}y_t$ y_{t-1}	x	x			
Mink & de Haan (2006)	Output Gap Change	y_t			x		Pre-Election Dummy GDP Growth Forecast Error Inflation Forecast Error
Garcia et al. (2009)	Output Gap	y_t	x	x			
Egert (2010)	GDP Growth Output Gap	y_t		x	x	x	Housing Price Growth Stock Price Growth Gov. Consumption Population Growth Interest Payments Imports/Exports
Darby & Melitz (2011)	GDP Level	y_t y_{t-1}		x	x	x	
Fatas & Mihov (2012)	Output Growth Output Gap	y_t	x	x			
Benetrix & Lane (2013)	Deviations in GDP from Quadratic Trend	y_t	x	x			Current Account Balance Domestic Credit Growth
Plodt & Reicher (2015)	Output Gap Trend GDP	y_t		x	x		Output Gap Revision Output Gap Forecast Chief Executive Political Party Dummy
Bernoth et al. (2015)	Output Gap	y_t	x	x	x		

discretionary responses to be limited to particularly severe economic events or times in which other types of policy are ineffective (for example, at the zero lower bound). Another possibility is that policymakers have other goals they see as being more important than and incompatible with countercyclical policy. For example, policymakers may care more about reducing government debt than stabilizing output. As the U.S. experience in 2013 demonstrates, concerns about fiscal responsibility can result in contractionary policy even during times of weak economic growth.

Finally, acyclical or procyclical policy could be unintentional. Legislation takes time to implement, which may cause changes in policy to occur later in the business cycle than intended. For example, spending increases in the American Recovery and Reinvestment Act, passed in early 2009, peaked in the first quarter of 2010, six months after the official end of the Great Recession. Similarly, policy decisions may be based on faulty information. Revisions to output and employment variables can be substantial, which may cause large differences between intended policy and actual policy outcomes.[4] This issue has been studied in the literature, with many studies concluding that intended policy tends to be more countercyclical than actual policy.[5]

The purpose of this paper is to determine whether and how federal discretionary fiscal policy responds to the business cycle in the United States. Motivated by discrepancies in the existing literature, I approach the issue of model uncertainty from a Bayesian perspective, treating the set of covariates in the underlying model as an additional parameter to be estimated. Using Bayesian

[4]See Orphanides and van Norden (2002) and Aruoba (2008) for more about the magnitude of data revisions.

[5]See, for example, Forni and Momigliano (2004), Golinelli and Momigliano (2009), and Egert (2010).

techniques I calculate posterior probabilities for each of a large set of models, where each model is defined by the included covariates. These posterior probabilities indicate the probability that a particular model is the underlying model that generated the data. I then average coefficient posteriors across the entire set of models using posterior model probabilities as weights. This procedure, known as Bayesian model averaging (BMA), produces results that are not conditioned on any particular model and that, because they are weighted by posterior model probabilities, reflect uncertainty about the underlying model. It also produces inclusion probabilities that, in the current context, enable me to determine which variables matter to policymakers. In other words, BMA provides a formal statistical framework that explicitly accounts for the type of model uncertainty that appears to be widespread in this literature.

The set of covariates that I consider contains six measures of the business cycle as well as other control variables common in the literature. My inclusion of employment-based measures of the business cycle, the unemployment gap and the change in the unemployment rate, is novel. The number of models that I estimate is substantial, $2^{16} = 65{,}536$ for my initial results and $2^{33} = 8{,}589{,}934{,}592$ when I consider asymmetric responses to expansions and recessions. The computational requirements of estimating such a large number of models necessitates the use of Markov Chain Monte Carlo techniques developed by Madigan and York (1995).

My results suggest that federal discretionary policy in the United States is countercyclical: the portion of the deficit determined by discretionary policy actions increases in response to poor economic conditions. This response appears to be driven primarily by changes in taxes, although I find some evidence that spending exhibits a similar, albeit smaller, response. Distinguishing between expansions and recessions makes it clear that countercyclical responses are limited to recessions.

Indeed, during expansions policy has a very low estimated probability of responding to business cycle measures in either direction. Posterior probabilities indicate that policy is much more likely to respond to employment-based measures of the business cycle than output-based measures, particularly the change in the unemployment rate. This is a striking result given the ubiquity of output-based measures elsewhere in the literature. Also notable is my finding that policy is unlikely to respond to the level of publicly-held debt. A comparison of my results with those obtained from more traditional fiscal policy models suggests that the importance of debt may be overstated in models that do not account for model uncertainty.

In contrast with Auerbach (2002 and 2003) and Cohen and Follette (2003), I find little evidence of a shift in the responsiveness of policy to the business cycle during the course of my fifty-year sample. The posterior probability that a structural break occurred in my model coefficients is just .48%. Finally, and again in contrast with Cohen and Follette (2003), replacing ex post data with real-time data reveals that intended responses to the business cycle are very similar to the responses that actually occur. In the latter instance BMA proves to be useful in identifying the source of these different findings.

The rest of the paper proceeds as follows: Section 2 discusses traditional models of fiscal policy conduct and the implementation of Bayesian model averaging. Section 3 discusses data and Section 4 presents my results. Section 5 concludes.

2. Methodology

2.1 Traditional Models of Discretionary Policy Conduct

Models of discretionary fiscal policy conduct usually take the form

$$CAD_t = \alpha + \beta_1 CAD_{t-1} + \beta_2 output_t + \beta_3 debt_{t-1} + \beta_4 Z_t + e_t \tag{3.1}$$

where CAD_t is the so-called cyclically-adjusted deficit, *output* is usually the output gap or GDP growth, and Z_t represents other possible control variables.

The cyclically-adjusted deficit measures what the deficit would be if the economy were at full employment and is constructed to eliminate the effects of automatic stabilizers on the deficit.[6] As a result, it should reflect discretionary fiscal policy actions. I use cyclically-adjusted net federal government savings as my measure of the cyclically-adjusted deficit.[7] This measure, published quarterly by the Congressional Budget Office (CBO), is a translation of the U.S. federal budget into National Income and Product Accounts (NIPA) terms and therefore differs slightly from federal deficits and surpluses.[8] It has been used previously to study discretionary policy in the United States and is similar to measures constructed by the OECD and IMF to study European policy.[9] For simplicity I will refer to

[6]Automatic stabilizers are defined as automatic changes in government revenues and expenditures that occur in response to the business cycle.

[7]For ease of interpretation, I reversed the sign of this variable so that a positive value indicates a deficit and a negative value indicates a surplus.

[8]Differences result from coverage adjustments (certain transactions are included in one framework but not the other), timing differences (some transactions are recorded on an accrual basis in the NIPAs but a cash basis in the federal budget), and differences in the categorization of transactions (some transactions count as negative taxes in one framework and positive spending in another). For more information see "NIPA Translation of the Fiscal Year 2017 Federal Budget" at https://www.bea.gov/scb/pdf/2016/04%20April/0416_nipa_translation_of_the_2017_federal_budget.pdf.

[9]See Auerbach (2002,2003) and Cohen and Follette (2003) for its use to study U.S. policy and Golinelli and Momigliano (2009) for a list of papers that use it to study European policy.

cyclically-adjusted net federal government savings as the cyclically-adjusted deficit for the remainder of the paper.

Equation (1) implies that discretionary fiscal policy responds to output within a period. Since most of the cyclicality literature uses annual data, this is a plausible assumption. However, since I use quarterly data the possibility of policy lags is much more likely. It is well-known that fiscal policy is prone to a number of lags that may prevent it from responding immediately to economic conditions. For example, noisy data may prevent policymakers from recognizing that policy is needed, and the sometimes contentious nature of the policymaking process may delay responses even once the need to respond is established. Indeed, the assumption that discretionary policy does not respond to output within a quarter is used by much of the fiscal multiplier literature to identify fiscal shocks.[10] As a result, I make the same assumption in this paper, using the first lag of each business cycle measure in each model I estimate.

I use three different measures of the output gap, real GDP growth, the unemployment gap, and the change in the unemployment rate as business cycle measures. Control variables are similar to those found elsewhere in the literature. They include publicly-held debt, the federal funds rate, inflation, political dummy variables, three lags of the dependent variable, and a time trend. These variables are discussed in further detail in section 3.

2.2 Bayesian Model Averaging

I consider j=1, . . . , J linear regression models in which the cyclically-adjusted deficit is regressed on an intercept and a subset k_j of K possible explanatory

[10]See Blanchard and Perotti (2002) and Auerbach and Gorodnichenko (2012), among others.

variables. Formally, I estimate

$$CAD = \alpha\iota_T + X_j\beta_j + \epsilon \tag{3.2}$$

where CAD is a $T \times 1$ vector holding observations of the cyclically-adjusted deficit, ι_T is a $T \times 1$ vector of ones, $X_j \in X$ is a $T \times k_j$ matrix containing the regressors in model j, and ϵ is assumed to be $\mathrm{N}(0_T, h^{-1}\,\mathrm{I}_T)$. I assume that uncertainty about the underlying model of fiscal policy conduct extends only to which covariates belong in the model. As a result, each model is defined by the included regressors, X_j. As suggested in Fernandez, Ley and Steel (2001b), each of the variables in X is de-meaned to ensure that the intercept, α, has the same interpretation in each model.

The Bayesian approach to model comparison involves the estimation of posterior model probabilities, which indicate the probability that a given model is the true model. The posterior model probability for model M_j is calculated as

$$\Pr(M_j|Y) = \frac{p(Y|M_j)\Pr(M_j)}{\sum_{i=1}^{J} p(Y|M_i)\Pr(M_i)} \tag{3.3}$$

where $\Pr(M_j)$ is the prior for model j and $p(Y|M_j)$ is the marginal likelihood. The marginal likelihood is the expected value of the likelihood function where the expectation is taken with respect to the prior for the model's parameters. It can be interpreted as the average fit of a particular model over the prior parameter values.

Once posterior model probabilities are estimated, posterior distributions for objects of interest (e.g. slope coefficients) can be obtained. One option is to focus on results from the model with the highest posterior probability. However, this approach is problematic if there are multiple plausible models that produce different or even conflicting results. Another option is to average results across all possible models using posterior model probabilities as weights. This procedure is known as Bayesian model averaging (BMA). Using BMA, the posterior distribution for some object of interest, λ, is calculated as

$$p(\lambda|Y) = \sum_{j=1}^{J} p(\lambda|Y, M_j) \Pr(M_j|Y) \tag{3.4}$$

BMA is beneficial for a number of reasons. First, estimates incorporate results from many possible models which, as mentioned above, reduces the likelihood that inferences are driven by model choice. In fact, as equation (4) makes clear, estimates are not conditioned on any particular model. Additionally, BMA estimates reflect uncertainty about the underlying model since results from each model are weighted by the associated posterior model probability. Estimates based on a single model, in contrast, are calculated under the potentially implausible assumption that the model they come from has a 100% posterior probability. Finally, BMA generates inclusion probabilities that indicate the likelihood that a particular variable belongs in the underlying model. In the current context, inclusion probabilities are useful in determining which variables matter to policymakers.

When the total number of models is small, BMA can be implemented using the following steps:

1. Calculate posterior statistics (e.g. coefficient posterior means) for all models.
2. Calculate posterior model probabilities, as in equation (3). When models are given equal prior odds, $\Pr(M_j)$ and $\Pr(M_i)$ drop out of the equation.
3. Average posterior statistics using posterior model probabilities as weights, as in equation (4).

In practice, implementing steps 1-3 above is computationally burdensome and potentially infeasible when the number of possible models is large. In that case, one can sample from the model space using the Markov Chain Monte Carlo Model Composition (MC^3) algorithm of Madigan and York (1995). MC^3 produces random draws from a Markov chain whose stationary distribution is the distribution defined by the posterior model probabilities.

Starting with an initial model, M^{s-1}, MC^3 proceeds in the following steps:

1. Propose a new model, M^*. This can be done using a symmetric proposal distribution that adds or deletes a single variable from the previous model. The use of a symmetric proposal distribution simplifies the acceptance probability in step 2.

2. Calculate the acceptance probability as

$$\alpha(M^{s-1}, M^*) = min\left[\frac{p(Y|M^*)\Pr(M^*)}{p(Y|M^{s-1})\Pr(M^{s-1})}, 1\right] \tag{3.5}$$

   With equal prior model odds, (5) simplifies to the ratio of marginal likelihoods for each model.

3. Accept M^* as the new draw, M^s, with probability α and reject it with probability $(1-\alpha)$. If M^* is rejected, the algorithm remains at M^{s-1}, in which case $M^s = M^{s-1}$

4. Calculate posterior statistics of interest for M^s.

5. Return to step 1 and repeat until convergence.

Once a sufficient number of draws are obtained, $\Pr(M_j|Y)$ can be calculated as the fraction of total draws for which model M_j is selected.

Intuitively, MC^3 identifies a subset of models with relatively high posterior probability, which reduces the number of models under consideration. At each iteration a model is proposed and accepted based on how well it fits the data (as measured by its marginal likelihood). Consequently models with good explanatory power are drawn more frequently than those with poor explanatory power. In fact, models with very low posterior probability may not be drawn at all, in which case their estimated posterior probability will be zero. Since these models have such low posterior probability this approximation should have a negligible effect on estimated posterior probabilities for drawn models.

I implement BMA using MC^3 since doing so analytically would take a great deal of time at current computing speeds. For example, suppose a single model can be estimated in one-hundredth of a second. At that rate, estimating my largest specification, for which the total number of models is $2^{33} = 8{,}589{,}934{,}592$, would take nearly 500 days to complete. However, as a robustness check I also calculated analytical results for my non-asymmetry model, for which the total number of models is a more manageable $2^{16} = 65{,}536$. The analytical results are virtually identical to those obtained using MC^3, suggesting that my use of MC^3 is appropriate.

When I use MC^3, I assume that 250,000 draws are sufficient for the Markov chain to converge and use an additional 1,000,000 draws for inference. I also checked for convergence in a number of ways. First, I increased the number of burn-in draws to 1,000,000, which had no discernible effect on my results. Next, I initialized MC^3 using two very different models, one with no covariates and one with every possible covariate. Again, my results were unaffected by the initial model. Finally, I calculated posterior model probabilities analytically for the subset of models visited by the algorithm (posterior probabilities were set to 0 for models that were not drawn) and compared them with posterior probabilities obtained using MC^3 for the same set of models. The correlation between analytical and numerical results for visited models is greater than 0.99, as recommended in Fernandez, Ley, and Steel (2001b).

2.3 Priors

To implement BMA, prior density functions are required for all models and their parameters. I assume that each possible covariate enters the true model independently of all other covariates with probability θ, suggesting a model prior of the form

$$\Pr(M_j) = \theta^{k_j}(1-\theta)^{K-k_j} \tag{3.6}$$

Setting $\theta = 0.5$, a common choice in the BMA literature, implies equal prior probability across all possible models, so that

$$\Pr(M_j) = \frac{1}{J}, \quad j = 1, \cdots, J \tag{3.7}$$

Although (7) places equal prior weight on all models, the same is not true of model *size.* Instead, models with few or many regressors receive lower prior probability than models of moderate size. This is evident by noting that equation (6) implies a prior distribution for model size, W, of the form

$$W \sim \text{Binomial}(K, \theta) \tag{3.8}$$

This distribution is centered on $K\theta$, so the use of a uniform model prior ($\theta = 0.5$) means that models with $\frac{K}{2}$ variables receive the most prior probability while those with 1 or K variables receive the least.

As an alternative to (7), Ley and Steel (2009) suggest using a hierarchical prior, making θ random instead of fixing it at a particular value. Specifically,

$$\theta \sim \mathrm{Beta}(a, b) \tag{3.9}$$

Ley and Steel recommend setting hyperparameter $a = 1$ and using a prior mean model size, $m = \frac{a}{a+b}K$, to elicit hyperparameter b. Setting $b = 1$ results in a uniform prior for model size,

$$\Pr(W = w) = \frac{1}{K+1} \quad \text{for } w = 0, \cdots, K \tag{3.10}$$

while setting $b > 1$ places greater prior probability on smaller models than larger models. Once the prior for model size is obtained, the model prior is calculated as

$$\Pr(M_j) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \frac{\Gamma(a+k_j)\Gamma(b+K-k_j)}{\Gamma(a+b+K)} \tag{3.11}$$

For my main results I use (11) and set $b = 1$, although I demonstrate that these results are robust to the use of alternative model priors.

Turning to model parameters, I require a prior density function, $p(\alpha, \beta_j, h|M_j)$, for each set of parameters. I use the “benchmark” prior recommended in Fernandez, Ley, and Steel (2001b) for use when, as in the current paper, there is uncertainty about the covariates in a normal linear regression model. The prior involves the use of improper non-informative priors for parameters that appear in all models (α and h) and informative priors for those that do not (β_j). Specifically,

$$p(h) \propto h^{-1} \tag{3.12}$$

$$p(\alpha) \propto 1 \tag{3.13}$$

$$\beta_j | h \sim N(\underline{\beta}_j, h^{-1}(\underline{g} X_j' X_j)^{-1}) \tag{3.14}$$

where $(\underline{g} X_j' X_j)^{-1}$ is the g-prior of Zellner (1986). I set $\underline{\beta}_j = 0_{k_j}$ and $\underline{g} = 1/\max\{T, K^2\}$, again as recommended in Fernandez, Ley, and Steel (2001b). This prior is useful because it limits the choice of hyperparameters to one, $\underline{g}$, which is chosen in an automatic fashion. In addition to requiring little subjective information from the researcher, the authors find that it has little influence on posterior inference. Finally, it reduces the computational burden of implementing MC^3 since analytical results are available for $p(\beta_j|Y)$ and $\Pr(M_j|Y)$.

3. Data

I use quarterly data covering the period 1966q1-2016q3. Table 3 lists the source of each of the data series used to construct my final variables. I use three different measures of the output gap. One measure comes from the CBO and I estimate the other two myself. The first (which I call the trend break output gap) is estimated using a linear time trend and a break in the trend after 1973 to allow for a slowdown in GDP growth.[11] The other is estimated using the filter proposed in Hamilton (2017).[12]

[11]See Orphanides and van Norden (2002) for further explanation.

[12]The Hamilton filter is an alternative to the HP filter discussed in Hamilton (2017).

TABLE 3. Data Sources

Variable	**Source**
Cyclically-adjusted net federal government savings	BEA (via CBO)
Real GDP	FRED
Real Output Gap	CBO (via FRED)
Nominal Potential GDP	CBO (via FRED)
Unemployment Rate	FRED
Natural Rate of Unemployment	FRED
Publicly-Held Federal Debt	FRED
Federal Funds Rate	FRED
CPI	FRED
GDP Deflator	FRED
Presidential Election Dummy	The American Presidency Project
United Government Dummy	The American Presidency Project
Recession Dummy	NBER

Notes: BEA is the Bureau of Economic Analysis, CBO is the Congressional Budget Office, FRED is the Federal Reserve Economic Database, and NBER is the National Bureau of Economic Research.

The cyclically-adjusted deficit and debt are expressed as percentages of potential GDP, as measured by the CBO.[13] An augmented Dickey-Fuller test indicates that the cyclically-adjusted deficit is stationary so I include it and its lags in levels.[14] Both inflation variables are calculated as quarterly growth in the corresponding price indices and annualized. Real GDP growth is also calculated as quarterly growth and annualized. Finally, the "presidential election" dummy variable takes on a value of one in all quarters during an election year while the "united government" dummy variable takes on a value of one for quarters in which

[13]Results are unchanged when these variables are expressed as percentages of the other two potential output measures that I estimate.

[14]First-differencing the cyclically-adjusted deficit changes the estimated coefficients on the policy lags but otherwise has no effect on results.

the White House, Senate, and House of Representatives were all controlled by the same political party.

4. Results

4.1 Baseline Results

Inclusion probabilities, presented in Table 4, make it clear that discretionary policy responds to the business cycle: the posterior probability that at least one of the business cycle measures belongs in the underlying model of fiscal policy is 99.8%. Among the business cycle measures, the change in the unemployment rate receives far greater posterior probability, 99.1%, than any of the others. This result is striking since, to my knowledge, I am the first in this literature to consider employment-based measures of the business cycle. In contrast, GDP growth and the output gap, the measures most commonly employed in the literature, together receive just 22.7% posterior probability. Lastly, the unemployment gap receives 6.5% posterior probability, suggesting that policymakers care more about the direction of the unemployment rate than its level. This means, for example, that discretionary policy is less likely to respond to a high and stable unemployment rate than to a low unemployment rate that is increasing. The contraction in policy that occurred in 2013, when the unemployment rate remained elevated even after three years of steady decreases, is consistent with this finding.

Among the other variables, only the first two policy lags receive posterior probabilities greater than 50%. The high posterior probabilities received by the first and second lag, 100% and 94.8%, respectively, indicate that discretionary policy is persistent. Considering that most fiscal policy actions are determined as part of the budget negotiation process, this finding is intuitive. It seems likely that the starting

TABLE 4. BMA Inclusion Probabilities

Variable	**Probability**
CA Deficit First Lag	**100%**
CA Deficit Second Lag	**94.8%**
CA Deficit Third Lag	3.5%
Time Trend	3.3%
Output Gap (CBO)	10.4%
Output Gap (TB)	4.9%
Output Gap (Hamilton)	5.7%
GDP Growth	2.8%
Unemployment Gap	6.5%
Unemployment Change	**99.1%**
Debt	4.2%
Federal Funds Rate	4.8%
CPI Inflation	27.0%
GDP Deflator Inflation	6.6%
Presidential Election	3.6%
United Government	3.7%
Total Output Gap	19.9%
Total Inflation	29.3%

Notes: Results for variables with inclusion probabilities greater than 50% are bolded.

point for each year's budget is the budget from the previous year rather than a blank slate.

Finally, it is worth noting that the inclusion probability for the level of debt is just 4.2%. This is surprising given that Bohn (1998), Auerbach (2002 and 2003), and Cohen and Follette (2003) all find that U.S. fiscal policy responds to debt. Differences in findings may be driven in part by the use of a different dependent variable, in the case of Bohn (1998), or the use of a different debt measure, in the case of Auerbach (2002 and 2003) and Cohen and Follette (2003). However, they may also be due to the fact that these authors consider a single model instead of averaging results across many possible models as I do. As I demonstrate below, the use of a single model can result in significant estimated responses to this variable.

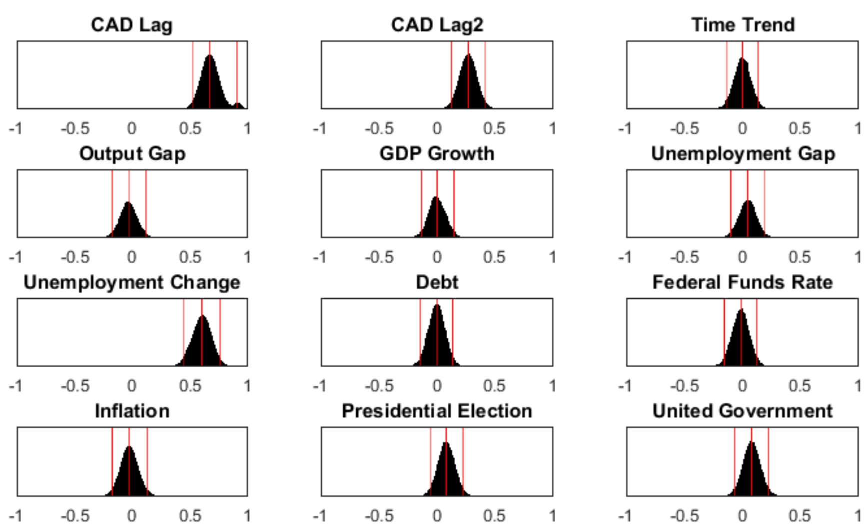
Table 5 and Figure 7 present information about coefficient posterior distributions, which measure short-run policy responses. Table 5 lists posterior means averaged across all possible models using posterior model probabilities as weights. For these results, I include values of zero that are assigned to coefficients whose corresponding variables are excluded from some models. In contrast, Figure 7 presents histograms for coefficient posterior distributions conditional on inclusion in the model. For these results I exclude values of zero for variables that do not appear in some models. In Table 5, posterior means are expressed in a number of different ways to ease interpretation. The first column lists posterior means as I estimate them, as a percentage of potential output. The second column converts the numbers in column one to billions of 2016q3 dollars to better convey their magnitude. Finally, the third column expresses the numbers in column two as a percentage of the 2016q3 cyclically-adjusted deficit to put these magnitudes in context.

TABLE 5. BMA Posterior Means

Variable	**Posterior Mean (% potential)**	**Posterior Mean (dollars)**	**Posterior Mean (% deficit)**
Intercept	**2.85**	**537**	**91.3**
CA Deficit First Lag	**0.68**	**128**	**21.8**
CA Deficit Second Lag	**0.25**	**48**	**8.1**
CA Deficit Third Lag	0.00	0	0.1
Time Trend	0.00	0	0.0
Total Output Gap	-0.01	-1	-0.2
GDP Growth	0.00	0	0.0
Unemployment Gap	0.00	1	0.1
Unemployment Change	**0.60**	**112**	**19.1**
Debt	0.00	0	0.0
Federal Funds Rate	0.00	0	0.0
Total Inflation	-0.01	-1	-0.2
Presidential Election	0.00	1	0.1
United Government	0.00	1	0.1

Notes: Column 1 includes posterior means measured as a percentage of potential output. Column 2 converts the numbers in column 1 to billions of 2016q3 dollars. Column 3 converts the numbers in column 2 to a percentage of the 2016q3 deficit.

FIGURE 7. Coefficient Posterior Distributions Conditional on Inclusion



Notes: For variables with inclusion probabilities less than 100% the large point mass at zero is ignored. Results for the third policy lag are omitted.

Together, Table 5 and Figure 7 indicate that discretionary policy is countercyclical: the coefficient posterior mean for the change in the unemployment rate is positive, conditional or unconditional on inclusion in the model. This means that an acceleration in the unemployment rate leads to larger deficits. The magnitude of this response is large. For example, assuming the unemployment rate has been stable, a one percentage point increase in this variable is predicted to increase the cyclically-adjusted deficit by about $111 billion dollars (using 2016q3 prices), which was nearly 19% of the total deficit in 2016q3.

It is informative to compare the results from my model with those that would be obtained from a more traditional cyclicality model like equation (1). To that end, Table 6 presents OLS estimates for individual models alongside my BMA estimates. For the OLS estimates, each model includes a single cylical variable as well as my non-cyclical covariates.[15] Bolded coefficients indicate a p-value of less than 0.1 (for columns 1-6) or an inclusion probability greater than 90% (for the BMA column).

Table 6 demonstrates that the importance of debt and inflation may be overstated in a single model that does not account for model uncertainty. This is particularly evident in columns (2) and (4), where coefficients for these variables have p-values of less than 0.1. I find, in contrast, that policymakers are unlikely to respond to either variable: the inclusion probability for CPI inflation is 27% while the inclusion probability for debt is just 4.2%. Similarly, the estimated responses to these variables are often larger in magnitude when a single model is used since BMA shrinks coefficients on variables with low estimated probability of being in the model towards zero.

[15] I exclude GDP deflator inflation because I do not want to include two inflation variables in an OLS regression and CPI inflation appears to explain the data better than GDP deflator inflation.

TABLE 6. Results from BMA and OLS

	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**BMA**
CA Deficit First Lag	**0.69**	**0.69**	**0.67**	**0.74**	**0.70**	**0.64**	**0.68**
CA Deficit Second Lag	**0.20**	**0.21**	**0.21**	**0.22**	**0.20**	**0.27**	**0.25**
CA Deficit Third Lag	0.00	-0.01	0.03	-0.01	-0.01	0.05	0.00
Time Trend	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Output Gap (CBO)	**-0.08**						-0.01
Output Gap (TB)		**-0.04**					0.00
Output Gap (Hamilton)			**-0.05**				0.00
GDP Growth				**-0.05**			0.00
Unemployment Gap					**0.09**		0.00
Unemployment Change						**0.65**	**0.60**
Debt	0.00	**-0.02**	-0.01	**-0.01**	-0.01	-0.01	0.00
Federal Funds Rate	-0.01	-0.02	0.00	-0.02	-0.01	-0.01	0.00
CPI Inflation	-0.03	**-0.04**	**-0.04**	**-0.04**	-0.03	-0.04	-0.01
Presidential Election	0.07	0.06	0.06	0.07	0.06	0.09	0.00
United Government	0.11	0.13	0.10	0.03	0.06	0.05	0.00

Notes: Each of columns (1)-(6) lists OLS results from a single model, where each model differs only by which business cycle measure it uses. The last column lists my BMA results. Coefficients with p-values less than 0.1 or inclusion probabilities greater than 90% are bolded.

In sum, my results indicate that discretionary policy in the United States is countercyclical. I find little evidence to support the idea that policymakers respond to output-based measures of the business cycle, an assumption made, to my knowledge, by all of the existing literature. Instead, policy is far more likely to respond to the change in the unemployment rate. Posterior probabilities also indicate that policy is persistent and unlikely to respond to the other covariates that I consider. Finally, accounting for model uncertainty greatly reduces the estimated influence of debt and inflation on policy outcomes.

4.2 Model Prior Robustness

The model prior employed in the previous section assigns equal prior probability to different model sizes but unequal prior probability to individual models. It is possible, then, that the results presented earlier are driven by the choice of model prior. To address this possibility, I examine the sensitivity of my results to the use of different model priors. Specifically, I consider a uniform prior across all models, achieved by setting $\theta = 0.5$ in equation (6), as well as a prior that places greater probability on smaller models than larger models, achieved by setting $b > 1$ in equation (11). In fact, when $b > 1$ the prior mode for model size is equal to zero. I considered a range of values for b but only present results for $b = 15$.[16]

Table 7 lists inclusion probabilities and averaged coefficient posterior means using three different priors: $b = 1$ (the original model prior), $\theta = 0.5$, and $b = 15$. Unsurprisingly, the average model size increases when greater prior probability is placed on larger models, and decreases when greater prior probability is placed on smaller models. This is due primarily to differences in inclusion probabilities for variables that are considered unlikely under the initial model prior. Inclusion probabilities for these variables are roughly twice as large when $\theta = 0.5$ and about half as large when $b = 15$. Since inclusion probabilities for these variables are so small under the initial model prior, even large proportional changes have little effect on overall conclusions. In contrast, inclusion probabilities are mostly unchanged for variables with high inclusion probabilities under the initial model prior. As a result, slope coefficient posterior means are also similar across model priors. It appears, then, that the basic conclusions from the previous section hold regardless of which model prior is used.

[16]Results for other values of b greater than 1 yield similar conclusions to when b is set to 15, which I chose because it implies a prior mean model size of one.

TABLE 7. Results for Alternate Model Priors

Variable	**Probability** ($b = 1$)	**Probability** ($\theta = 0.5$)	**Probability** ($b = 15$)
CA Deficit First Lag	**0.68**	**0.66**	**0.70**
	(100%)	**(100%)**	**(100%)**
CA Deficit Second Lag	**0.25**	**0.26**	**0.24**
	(94.8%)	**(98.0%)**	**(88.6%)**
Time Trend	0.00	0.00	0.00
	(3.3%)	(9.9%)	(1.2%)
Total Output Gap	-0.01	-0.01	0.00
	(19.9%)	(38.2%)	(10.7%)
GDP Growth	0.00	0.00	0.00
	(2.8%)	(6.3%)	(1.5%)
Unemployment Gap	0.00	0.01	0.00
	(6.55%)	(13.0%)	(3.3%)
Unemployment Change	**0.60**	**0.60**	**0.58**
	(99.1%)	**(99.85%)**	**(97.7%)**
Debt	0.00	0.00	0.00
	(4.2%)	(12.0%)	(1.4%)
Federal Funds Rate	0.00	0.00	0.00
	(4.8%)	(11.2%)	(2.2%)
Total Inflation	-0.01	**-0.01**	0.00
	(29.3%)	**(55.8%)**	(13.8%)
Presidential Election	0.00	0.01	0.00
	(3.6%)	(8.8%)	(1.4%)
United Government	0.00	0.01	0.00
	(3.7%)	(9.55%)	(1.4%)
Average Model Size	3.8	4.9	3.3

Notes: average model size excludes the constant that appears in every model. Coefficient estimates are on the same line as the corresponding variable name while inclusion probabilities are in parentheses. Variables with inclusion probabilities greater than 50% are bolded.

4.3 Business Cycle Asymmetry

It is possible that policy responds differently to economic conditions in expansions and recessions, for reasons outlined in the introduction. Cohen and Follette (2003) consider asymmetric responses to the business cycle in the United

States by estimating responses to positive and negative output gaps, which are included as two separate variables in their model. They find that responses to negative output gaps are significant and countercyclical while responses to positive output gaps are not significant, suggesting that discretionary policy is more likely to respond to poor economic conditions than to good economic conditions.[17]

I approach business cycle asymmetry in a slightly different manner. Instead of using positive and negative values of my business cycle measures to measure the strength of the economy, I use a lagged recession indicator that is equal to 1 if the economy was in a recession the previous period.[18] In contrast with Cohen and Follette (2003), I also allow responses to all covariates to differ between expansions and recessions instead of just the business cycle measures. To construct my asymmetry specification I interact the constant and each covariate with the lagged recession indicator, include the 17 interaction terms (which I call recessionary variables) alongside the uninteracted variables (which I call expansionary variables), and estimate the specification using BMA.

Table 8 presents inclusion probabilities and posterior means for the asymmetry specification alongside my initial results. Together, they make it clear that the discretionary response to business cycle measures estimated in the previous section is driven by large responses to the change in the unemployment rate during recessions. Inclusion probabilities for expansionary business cycle measures are all less than 5%, even for the change in the unemployment rate. Correspondingly, coefficient posterior means for these variables are small in magnitude although

[17]Because they define the output gap in the opposite way that I do, subtracting actual output from potential output, they actually estimate a significant and countercyclical response to *positive* output gaps, which they define as potential output exceeding actual output.

[18]Using a lagged recession indicator is consistent with my assumption that discretionary policy does not respond to economic conditions within a quarter.

TABLE 8. Asymmetry Specification Results

Variable	Baseline	Expansion	Recession
CA Deficit First Lag	**0.68**	**0.76**	0.73
	(100%)	**(100%)**	(15.1%)
CA Deficit Second Lag	**0.25**	**0.17**	0.15
	94.8%	**(72.3%)**	(11.2%)
Total Output Gap	-0.01	0.00	0.00
	(19.9%)	(1.9%)	(2.3%)
GDP Growth	0.00	0.00	0.00
	(2.8%)	(0.5%)	(0.6%)
Unemployment Gap	0.00	0.00	0.00
	(6.5%)	(0.6%)	(0.7%)
Unemployment Change	**0.60**	0.00	**1.16**
	(99.1%)	(0.7%)	**(99.8%)**
Debt	0.00	0.00	0.00
	(4.2%)	(0.5%)	(1.1%)
Federal Funds Rate	0.00	0.00	0.00
	(4.8%)	(1.1%)	(6.8%)
Total Inflation	-0.01	0.00	-0.01
	(29.3%)	(5.3%)	(11.9%)
Presidential Election	0.00	0.00	0.00
	(3.6%)	(0.7%)	(0.5%)
United Government	0.00	0.00	0.00
	(3.7%)	(0.6%)	(0.5%)

Notes: coefficient estimates are on the same line as the corresponding variable name while inclusion probabilities are in parentheses. Results are bolded for variables with inclusion probabilities greater than 50%. Results for the third deficit lag and time trend are omitted.

conditional on inclusion all but GDP growth are indicative of countercyclical policy. Thus it appears that during expansions policymakers feel little need to respond to economic conditions. The only variables to receive posterior probabilities greater than 5% are the first and second lags of the cyclically-adjusted deficit. The posterior means for these variables are relatively large at 0.76 and 0.17, indicating a great deal of persistence in policy.

During recessions, in contrast, policymakers have a very high probability of responding to the change in the unemployment rate. The inclusion probability for this variable increases from 0.7% during an expansion to 99.8% during a recession and its posterior mean, 1.16, is nearly twice as large as what I estimate for the non-asymmetry specification. It implies that, assuming a previously constant unemployment rate, a one percentage point increase in the unemployment rate during a recession causes a countercyclical response equal to about 220 billion dollars (in 2016q3 prices) or slightly less than 40% of the 2016q3 cyclically-adjusted deficit. This estimate appears to be plausible. For example, the change in the unemployment rate increased from 0.3% to 0.7% between 2008q3 and 2008q4 which, using my estimate, implies a 87.8 billion dollar increase in the cyclically-adjusted deficit (in 2016q3 dollars). The following quarter, the cyclically-adjusted deficit increased by 282 billion dollars.

4.4 Changes in Policy Over Time

A number of papers in the monetary policy literature have estimated changes in policy conduct over time, concluding that the Federal Reserve has responded more aggressively to changes in inflation since the mid-1980s.[19] Similarly, changes in political ideologies and priorities may have altered the conduct of fiscal policy over the past fifty years. Auerbach (2002 and 2003) and Cohen and Follette (2003) find evidence to support this, concluding that discretionary policy has responded more strongly to the output gap since 1993, a date chosen to coincide with the beginning of the Clinton administration. However, these authors do not test whether the change in policy responsiveness is significant, instead basing their conclusions on differences in coefficients estimated using different sample periods.

[19]See, for example, Taylor (1999b) and Stock and Watson (2002).

I consider changes in policy by searching for evidence of a structural break in the coefficients of my non-asymmetry specification, allowing for uncertainty across a number of dimensions.[20] Bayesian model comparison techniques outlined in section 2 can be used to compare “no break” models of the form

$$CAD_t = \alpha + X_t\beta + \epsilon_t \tag{3.15}$$

with structural break models of the form

$$CAD_t = \alpha + X_t\beta_1 + \gamma D_t + (X_t D_t)\beta_2 + \epsilon_t \tag{3.16}$$

where D_t takes on a value of zero before a particular break date and a value of one during and after the break date. In this setup there is uncertainty about the covariates in the model ($X_t \in X$), the existence of a structural break (whether equation (15) or (16) is the appropriate model type), and the date in which such a break may have occurred (D_t). I assume that if a variable appears in a particular structural break model, its coefficient is allowed to break. This reduces the model space, enabling me to compute analytical results.

As in earlier sections, inclusion probabilities can be obtained for individual covariates. Posterior probabilities can also be calculated for the existence and location of a structural break. Mathematically, the posterior probability that a structural break exists is

$$\Pr(\text{“break”}|Y) \propto p(Y|\text{“break”})\Pr(\text{“break”}) \tag{3.17}$$

[20] I do not consider the asymmetry model since the number of recession observations is greatly reduced when the sample is split in two. This means, for example, that a break date during the last decade of the sample would cause post-break policy to be identified solely off of the fiscal policy response to the Great Recession, an unusually severe and prolonged recession.

where

$$p(Y|\text{“break”}) = \sum_{X_t \in X} \sum_{D_t} p(Y|X_t, D_t, \text{“break”}) \Pr(X_t|D_t, \text{“break”}) \Pr(D_t|\text{“break”}) \tag{3.18}$$

is the marginal likelihood averaged across the set of structural break models using priors $\Pr(X_t|D_t, \text{“break”})$ and $\Pr(D_t|\text{“break”})$ as weights, and where $\Pr(\text{“break”})$ is the prior probability that a structural break exists. Posterior probabilities for different break dates are calculated in a similar manner,

$$\Pr(D_t|\text{“break”}, Y) \propto p(Y|D_t, \text{“break”}) \Pr(D_t|\text{“break”}) \tag{3.19}$$

where $\Pr(D_t|\text{“break”})$ is the prior probability that a particular break date is the true break date.

Since I have no prior beliefs about the existence and location of a structural break, the date range I test for a break as well as the prior probabilities I use in (17)-(19) reflect that. Following the recommendation in Andrews (1993), I use a 15% trimming value for a date range of (1973q2,2009q1). I set $\Pr(\text{“break”})$, the prior probability that a break exists, equal to 0.5 and $\Pr(D_t|\text{“break”})$, the prior probability that the break occurred at a particular date, equal to $\frac{1}{c}$ where $c = 144$ is the total number of dates under consideration. I use the same model prior, $\Pr(X_t|D_t, \text{“break”})$, that I use to obtain my baseline results.[21]

In contrast with previous studies, I find little evidence that a shift in policy conduct occurred sometime during the past fifty years. The posterior probability

[21] Results are unchanged when I use the alternate model priors discussed in section 4.2 instead.

that a break occurred, Pr(“break”$|Y$), is .48%. Similarly, the posterior odds ratio,

$$\frac{\text{Pr(“no break”}|Y)}{\text{Pr(“break”}|Y)} \tag{3.20}$$

indicates that the possibility that a break did not occur is 209 times more likely than the possibility that one did. This result is hard to overturn, requiring prior probabilities that overwhelmingly favor the existence of a break. As Table 9 demonstrates, Pr(“break”) must be set to more than 99% to produce a posterior probability greater than 50%.

4.5 Spending and Tax Responses

Thus far, policy has been measured as the portion of the deficit resulting from discretionary policy actions. While this measure has the advantage of summarizing overall policy, it does not distinguish between taxes and spending. Consequently it is useful to separately estimate tax and spending responses to determine whether Congress has relied more heavily on one fiscal lever than the other in responding to the business cycle. For the results below I replace the cyclically-adjusted deficit and its lags in equation (2) with the relevant fiscal measure.

My results, presented in Table 10, indicate that earlier results are driven primarily by taxes. Inclusion probabilities for the tax specification closely resemble those in section 4.1, suggesting that taxes have a high probability of responding to business cycle measures, particularly the change in the unemployment rate. The coefficient posterior mean for this variable is large and negative, indicating a countercyclical response. In contrast, the spending response to the business cycle appears to be much more modest. The probability that any one of the six business cycle measures belongs in the underlying spending model is 67.5% and, as Figure 8

TABLE 9. Structural Break Evidence

Pr("break")	Pr("break"$\vert Y$)
50%	0.48%
75%	1.42%
90%	4.13%
99%	32.1%
99.9%	82.7%

Notes: The first column displays different priors for the existence of a break date, Pr("break"), while the second column displays the corresponding posteriors, Pr("break"$|Y$).

makes clear, coefficient posterior distributions for these variables are centered on or near zero even conditional on being included in the model.

It is surprising that spending appears unlikely to respond to economic conditions given that unemployment benefit extensions and grants to states, both of which are included in cyclically-adjusted outlays, have been part of the federal response to many recent recessions.[22] It is therefore worth considering how results change when variations of the spending measure are used. As defined, cyclically-adjusted outlays include spending on defense and interest payments, neither of which are likely to be adjusted in response to the state of the economy. Instead, defense spending should depend primarily on U.S. military operations while interest payments are predetermined. Together, spending on defense and interest comprises nearly 40% of total cyclically-adjusted outlays over my fifty-year sample, which may make it harder to detect an empirical relationship between spending and business cycle measures even if one exists.

Cyclically-adjusted outlays also exclude two categories of spending commonly associated with fiscal stimulus legislation: gross government investment and capital

[22]Unemployment benefit extensions are not counted as automatic stabilizers because they require legislation to be enacted.

TABLE 10. BMA Results for Cyclically-Adjusted Tax Revenues and Outlays

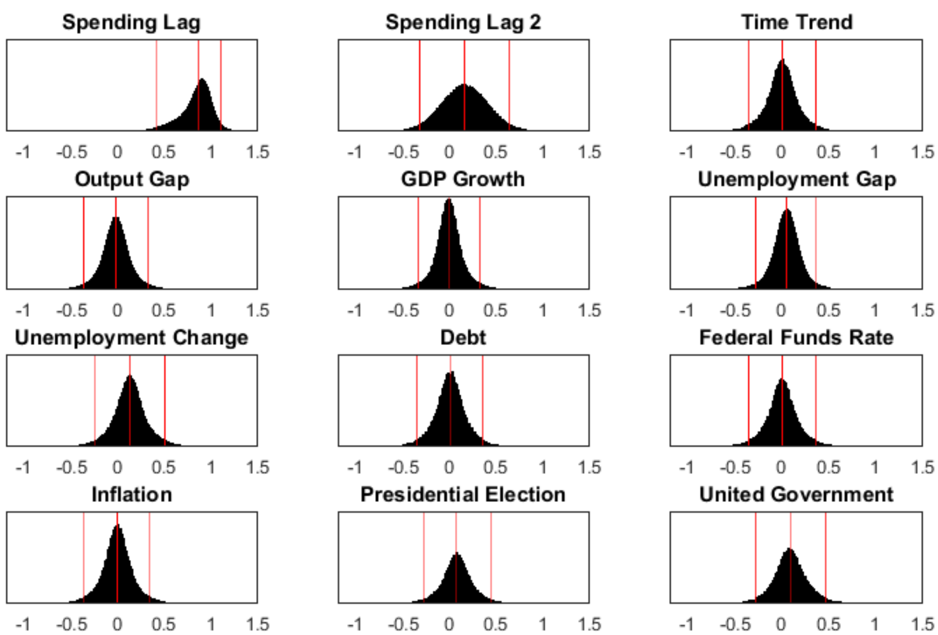
Variable	Revenues	Outlays
Dep. Var. First Lag	**0.62**	**0.84**
	(100%)	**(100%)**
Dep. Var. Second Lag	**0.26**	0.02
	(96.7%)	(10.1%)
Time Trend	0.00	0.00
	(2.2%)	(1.8%)
Total Output Gap	0.00	**-0.01**
	(9.1%)	**(55.2%)**
GDP Growth	0.00	0.00
	(2.5%)	(4.9%)
Unemployment Gap	0.00	0.00
	(3.4%)	(9.3%)
Unemployment Change	**-0.44**	0.01
	(99.5%)	(4.6%)
Debt	0.00	0.00
	(11.7%)	(1.8%)
Federal Funds Rate	0.00	0.00
	(12.8%)	(1.7%)
Total Inflation	0.00	0.00
	(13.3%)	(4.6%)
Presidential Election	0.00	0.00
	(1.8%)	(3.2%)
United Government	0.00	0.00
	(1.9%)	(3.6%)

Notes: coefficient estimates are on the same line as the corresponding variable name while inclusion probabilities are in parentheses. Results are bolded for variables with inclusion probabilities greater than 50%. Results for the third policy lag are omitted.

transfer payments. The former includes direct federal spending on structures such as schools and highways, while the latter includes grants to state and local governments for additional transportation infrastructure.[23]

[23] See "Concepts and Methods of the U.S. National Income and Product Accounts" at https://www.bea.gov/national/pdf/allchapters.pdf for further explanation.

FIGURE 8. Coefficient Posterior Distributions (Spending Specification)



Notes: for variables with inclusion probabilities less than 100% the large point mass at zero is ignored. Red lines indicate 5%, 50%, and 95% percentiles. Results for the third policy lag are omitted.

It is worth noting that government investment and capital transfer payments, which include spending on "shovel-ready projects," together accounted for just over 14% of the total expenditures included in the American Recovery and Reinvestment Act.[24] However, it is possible that they have played a larger role in government responses to other economic events over the past fifty years.

Column 1 of Table 11 presents inclusion probabilities and averaged coefficient estimates for the unadjusted spending variable alongside results for two variations of this measure. The first, whose corresponding results are listed in column 2, excludes spending on defense and interest payments. The second, whose results are listed in column 3, adds gross government investment and capital transfer payments to the measure in column 2. Table 11 shows that alternate definitions of discretionary spending provide stronger evidence that policymakers adjust spending in response to the business cycle. The inclusion probability for the change in the unemployment rate increases from 4.6% to 66.4% when spending on defense and interest payments is excluded, and increases further to 85.1% when gross investment and capital transfers are added. Similarly, the probability that any of the business cycle measures belongs in the model is about 95% for both alternate spending definitions. Coefficient posteriors suggest that spending responses are countercyclical. The coefficient posterior mean for the change in the unemployment rate is much larger in magnitude than for the unadjusted spending variable, albeit smaller than the response estimated for the tax specification. In sum, it appears that policymakers respond to the business cycle with a combination of tax and spending changes, although taxes tend to make up a larger portion of the response.

[24] See "Effect of the ARRA on Selected Federal Government Sector Transactions" at https://www.bea.gov/recovery/pdf/arra-table.pdf.

TABLE 11. BMA Results for Alternate Spending Variables

Variable	**(1)**	**(2)**	**(3)**
Outlays First Lag	**0.84**	**0.70**	**0.56**
	(100%)	**(100%)**	**(100%)**
Outlays Second Lag	0.02	0.04	**0.40**
	(10.1%)	(19.2%)	**(100%)**
Total Output Gap	**-0.01**	-0.01	0.00
	(55.2%)	(40.2%)	(13.5%)
GDP Growth	0.00	0.00	0.00
	(4.9%)	(3.7%)	(5.9%)
Unemployment Gap	0.00	0.00	0.00
	(9.3%)	(7.9%)	(2.6%)
Unemployment Change	0.01	**0.17**	**0.23**
	(4.6%)	**(66.4%)**	**(85.1%)**

Notes: column (1) lists results for the unadjusted spending variable. Column (2) removes spending on defense and interest payments from (1) and column (3) adds gross investment and capital transfers to (2). Coefficient estimates are on the same line as the corresponding variable name while inclusion probabilities are in parentheses. Results are bolded for variables with inclusion probabilities greater than 50%.

4.6 Intended Policy

Since Orphanides (2001), it has become common to assess monetary and fiscal policy conduct using both real-time and ex post data. Ex post data is useful for estimating actual policy outcomes but it may do a poor job of explaining intended policy if the information available to policymakers in real time differs substantially from fully revised data. In that case, intended policy responses can be estimated by replacing ex post data with real-time data. Using real-time data, studies of fiscal policy in European and OECD countries have often concluded that intended policy is more countercyclical than actual policy. Cohen and Follette (2003), in contrast, find that in the United States intended policy is less countercyclical than actual policy. They estimate a countercyclical response to the output gap using both ex post and real-time data, but find that the magnitude of the response is about half as large when real-time data is used.

In line with the existing literature, I estimate intended responses to the business cycle by replacing ex post data for business cycle measures with real-time data and re-estimating equation (2). Real-time data for each variable is available from the Archival Federal Reserve Economic Data (ALFRED) database. Because of data availability, I am forced to make a number of alterations to my dataset. First, real-time data for the CBO output gap and unemployment gap are available only for recent years so I exclude both as possible covariates. This reduces the total number of business cycle measures to four: two measures of the output gap, GDP growth, and the change in the unemployment rate.[25] Second, my real-time dataset begins in 1968q4 instead of 1966q1, again because of data availability.[26] Fortunately, results using ex post data for the smaller set of covariates and shorter sample, listed in column 1 of Table 12, are very similar to those presented in earlier sections, suggesting that such alterations are inconsequential.

My results indicate that there is little difference between intended and actual policy: estimated responses to business cycle measures, shown in column 2 of Table 12, are largely unchanged when real-time data is used. The estimated response to the change in the unemployment rate falls from 0.58 to 0.56, a difference of 3.5%, while the response to GDP growth is close to zero using both ex post and real-time data. The estimated response to the output gap changes by a large amount proportionally, from -0.01 to 0.00, but the magnitude of this coefficient is small

[25] Real-time data is available for potential output at an annual frequency from 1991 and biannual frequency from 1999 and is available for the natural rate of unemployment from 2011.

[26] Real-time data for GDP is available from 1965 but the estimation of the Hamilton output gap, which involves regressing y_{t+h} on $y_t,\cdots,y_{t-3}$ causes me to lose observations for the first three years.

TABLE 12. BMA Results for Real-Time Business Cycle Measures

Variable	Ex Post	Real Time
CA Deficit First Lag	**0.68**	**0.69**
	(100%)	**(100%)**
CA Deficit Second Lag	**0.25**	**0.25**
	(94.1%)	**(93.7%)**
Time Trend	0.00	0.00
	(4.7%)	(4.8%)
Total Output Gap	-0.01	0.00
	(24.6%)	(10.0%)
GDP Growth	0.00	0.00
	(3.9%)	(5.9%)
Unemployment Change	**0.58**	**0.56**
	(96.0%)	**(97.9%)**
Debt	0.00	0.00
	(6.4%)	(7.5%)
Federal Funds Rate	0.00	0.00
	(6.2%)	(5.8%)
Total Inflation	-0.01	-0.01
	(37.2%)	(44.1%)
Presidential Election	0.01	0.01
	(6.3%)	(6.3%)
United Government	0.00	0.00
	(5.1%)	(4.9%)

Notes: coefficient estimates are on the same line as the corresponding variable name while inclusion probabilities are in parentheses. Results are bolded for variables with inclusion probabilities greater than 50%.

enough relative to the coefficient on the change in the unemployment rate that it has little bearing on overall conclusions.

It is worth noting that, although small, the differences in results that I estimate using ex post and real-time data are in line with the finding by Cohen and Follette (2003) that intended policy is less countercyclical than actual policy. For both the change in the unemployment rate and the output gap, the sign of the coefficient posterior mean is unaffected when real-time data is used but the magnitude is smaller. It is possible that Cohen and Follette overestimate the magnitude of policy differences because they assume that policymakers respond to the output gap, a variable that I estimate to have a low probability of belonging in the underlying model. Re-estimating equation (2), first-differencing the dependent variable as in Cohen and Follette, confirms this. Conditional on inclusion in the model, the coefficient posterior mean for the output gap falls from -0.05 when ex post data is used to -0.02 when real-time data is used. This result is very similar to the values of -0.04 and -0.02 estimated by Cohen and Follette.

The output gap may produce larger differences in estimated policy responses because it is less accurate in real-time than the change in the unemployment rate. The information presented in Table 13 supports this idea. The first column, intended to convey average revision size for each business cycle measure, lists the average magnitude of revisions to each variable as a percentage of the average magnitude of the corresponding variable, measured using ex post data. The second column, intended to convey the frequency with which policymakers have had access to accurate real-time data, lists the percentage of observations for which the magnitude of a variable's revision is smaller than five percent of the variable's average magnitude, again measured using ex post data. Table 13 indicates that real-time data for the change in the unemployment rate has been more accurate than real-time data for the other business cycle measures. Policymakers have had access to accurate real-time data for this variable more often than for other business cycle measures, and when real-time data has been inaccurate, revisions have tended to be smaller in magnitude than revisions to other variables.

In sum, the use of the output gap may yield inaccurate conclusions about differences between actual and intended policy. Posterior probabilities indicate that policymakers are much less likely to respond to the output gap, a variable prone

TABLE 13. Business Cycle Measure Revision Summary

Variable	Average Revision Size	RT Accuracy Frequency
Output Gap (TB)	89.6%	1.6%
Output Gap (Hamilton)	43.5%	6.3%
GDP Growth	42.8%	6.4%
Unemployment Change	35.6%	24.5%

Notes: “Average Revision Size” is calculated as the average revision magnitude as a percentage of the average ex post variable magnitude while “RT Accuracy Frequency” is calculated as the percentage of observations for which the revision magnitude is less than five percent of the ex post variable magnitude.

to relatively large and frequent inaccuracies in real time, than to the change in the unemployment rate, a variable that is more accurate in real time. Accounting for model uncertainty and in contrast with Cohen and Follette (2003), I find little difference between intended policy and actual policy outcomes in the United States.

5. Conclusion

The conduct of discretionary fiscal policy has been the subject of a large number of papers. Nevertheless, these papers have failed to yield a consensus on even the most basic aspects of policy, such as whether and how it responds to the business cycle. Conflicting results may stem in part from model uncertainty, particularly uncertainty about which covariates belong in the underlying model of fiscal policy. Motivated by discrepancies in the existing literature, I estimate the response of U.S. policy to different business cycle measures using a Bayesian approach that explicitly incorporates model uncertainty. My results indicate that policy responds to business cycle measures, particularly the change in the unemployment rate, in a countercyclical manner. These countercyclical responses are driven by large responses to recessions. During expansions, in contrast, policy shows little indication of responding to economic conditions. Policymakers appear to rely more heavily on tax cuts than spending increases during recessions, although I find evidence that both change in response to the change in the unemployment rate. Finally, I find no evidence of a structural break in my model coefficients, nor of substantive differences between intended policy and actual policy outcomes.

Although my focus in this paper is on responses to the business cycle, my finding that debt has little effect on discretionary fiscal policy relates to a large and growing literature on fiscal and monetary policy switching. In that literature there exist two fiscal policy regimes, active and passive, where the regimes are defined by whether or not fiscal policy stabilizes debt. My results suggest that if fiscal policy can be described as operating under two regimes, those regimes may be better defined by whether the economy is in a recession or an expansion, with the terms "active" and "passive" referring to whether or not fiscal policy responds to business cycle measure.

The conduct of federal discretionary fiscal policy is an important topic in and of itself. However, any analysis of fiscal policy is incomplete as long as it excludes non-discretionary and subnational policy. Both types of policy operate under very different conditions from federal discretionary policy and as such likely respond to the economy in very different ways. Applying the techniques employed in this paper to either type of policy is an obvious avenue for future research.

CHAPTER IV

REGIME SWITCHING IN U.S. FISCAL POLICY

1. Introduction

The Great Recession prompted large-scale fiscal policy initiatives throughout the world. In the years since, policymakers and academics alike have vigorously debated the efficacy of such policy, focusing primarily on the sign and magnitude of spending and tax multipliers. A key question that has emerged is whether such multipliers are regime dependent. There is some evidence to suggest that they are, depending on the amount of slack in the economy, the stance of monetary policy, and other macroeconomic characteristics.[1]

A related question is whether the *implementation* of fiscal policy, particularly discretionary fiscal policy, is regime dependent. This topic has received some attention in empirical work. For example, a number of authors have determined that the cyclicality of discretionary policy differs according to the state of the business cycle.[2] Others have found that policy responses depend on the level of outstanding public debt.[3] While these findings are suggestive, they often rely on strong assumptions about the underlying model. For instance, many of the aforementioned authors conclude that policy responses vary over the business cycle because estimated responses to positive and negative values of the output gap differ. The implicit assumption in these papers is that policy regimes are

[1]See, for example, Auerbach and Gorodnichenko (2012), Fazzari, Morley, and Panovska (2015), Ramey and Zubairy (2018), Christiano, Eichenbaum, and Rebelo (2011), and Ilzetzki, Mendoza, and Vegh (2013).

[2]See Falconer (2018), Cimadomo (2007), Forni and Momigliano (2004), and Cohen and Follette (2003).

[3]See Combes, Minea, and Sow (2017) and Egert (2012).

determined by the value of the output gap, where the specific value that initiates a regime switch is zero. More generally, the variables and corresponding values that define regimes are frequently imposed rather than estimated in this literature. Additionally, since the tendency is to focus on switching in the response to a particular variable, responses to other variables are often assumed to be invariant across regimes. In sum, the models used previously to examine state dependence in discretionary fiscal policy have allowed for relatively little uncertainty about the structure of the underlying model governing such dependence.

In this paper I search for evidence of regime dependence in federal discretionary fiscal policy conduct in the United States. My interest is primarily in whether policy is better modeled as a linear or regime-switching process and, if the latter, which variables and values characterize regime switching. I am also interested, to a lesser extent, in which variables policy responds to as well as the responses to these variables across regimes. It should be noted that this was the focus of Falconer (2018), which found that policy is more countercyclical during recessions than during expansions. My aim, then, is not to retread the same ground but rather to determine whether the conclusions of that paper are robust to the use of more plausible regime definitions.

This paper not only extends the existing literature on fiscal policy conduct, but also complements existing work on regime dependent fiscal multipliers. If the effectiveness of fiscal policy varies according to the value of some variable, it is important to determine whether policy conduct exhibits a similar pattern. Indeed, if policy is more effective at certain times than at others, policymakers may want to modify existing policy to take advantage of that fact, if they do not do so already.

I evaluate regime dependence in discretionary policy by using Bayesian model comparison techniques to compare a large number of linear and regime-switching models motivated by the existing literature. Doing so enables me to determine which models are most likely to have generated the data and, correspondingly, which model characteristics are most plausible. I assume that regime switching is limited to the underlying model's coefficients, which can be represented with a threshold model. I allow for uncertainty about the underlying model's covariates and, conditional on that model being a threshold model, its threshold variable and threshold value. As a result, the models that I estimate differ across each of those dimensions. I consider twenty-nine possible threshold variables related to the business cycle, debt scenario, monetary policy, inflation, and the political environment. After obtaining measures of uncertainty about the characteristics of the underlying model I incorporate that uncertainty into parameter estimates using Bayesian model averaging.

My results provide strong evidence that discretionary fiscal policy is regime dependent. Posterior probabilities indicate that the set of threshold models is at least five times more likely than the set of linear models. This result is hard to overturn, requiring prior probabilities that greatly favor the set of linear models. Among the threshold variables that I consider, the change in nonfarm payrolls and change in the federal funds rate are the most likely drivers of regime change. Notably, the sign of the output gap has a very low probability of defining regimes. The same is true of the five debt variables that I consider.

Threshold value posteriors support the idea that the change in nonfarm payrolls and the change in the federal funds rate identify similar regimes. Extreme negative values have the highest posterior probability for both variables, splitting observations into periods of "normal" and "bad" macroeconomic conditions that roughly coincide with recession dates. The number of "bad" episodes that occurred in the United States between 1966 and 2016 is smaller when the change in nonfarm payrolls is used as the threshold variable than when the change in the federal funds rate is used, but the duration of these episodes is typically longer.

Finally, inclusion probabilities and coefficient posterior means largely corroborate the findings in Falconer (2018): policy is persistent and has a high probability of responding to changes in the labor market. Responses to the labor market are mostly limited to periods of weak economic performance, however; at other times policy is slightly countercyclical or acyclical. Fiscal policymakers may also follow the Federal Reserve's lead during the "bad" regime, implementing expansionary fiscal policy in the quarter following expansionary monetary policy. During the "normal" regime, in contrast, there is no evidence of comovement between fiscal and monetary policy. In sum, my results support the idea that fiscal policymakers drastically alter policy when the economy is doing poorly, responding more strongly (and countercyclically) to the business cycle. The decision to alter policy appears to depend either directly on labor market developments or on the actions of the Federal Reserve in response to these and other macroeconomic developments.

The paper proceeds as follows: Section 2 discusses the linear and threshold models I consider as well as Bayesian model comparison and Bayesian model averaging. Section 3 describes data sources and variable definitions and Section 4 presents results. Section 5 concludes.

2. Methodology

2.1 Linear Models

The starting point for my analysis is the set of models in Falconer (2018).[4] The results of that paper indicate that there is substantial uncertainty about the covariates in the underlying fiscal policy model. Consequently I consider j=1, . . . , J linear regression models, each of which differs with respect to the included covariates. Formally, I estimate

$$CAD = \alpha\iota_T + X_j\beta_j + \epsilon \tag{4.1}$$

where CAD is a $T \times 1$ vector holding observations of the cyclically-adjusted deficit, ι_T is a $T \times 1$ vector of ones, $X_j \in X$ is a $T \times k_j$ matrix containing the regressors in model j, and ϵ is assumed to be N(0_T,h^{-1} I_T). As suggested in Fernandez, Ley and Steel (2001b), each of the variables in X is de-meaned to ensure that the intercept, α, has the same interpretation in each model.

The cyclically-adjusted deficit measures what the deficit would be if the economy were at full employment. It is constructed to eliminate the effects of automatic stabilizers on the deficit and should, as a result, reflect discretionary fiscal policy actions.[5] The cyclically-adjusted deficit is used throughout the literature as a measure of discretionary fiscal policy.[6] I use cyclically-adjusted net federal government savings, published for the United States by the Congressional

---

[4] Greater detail can be found in that paper.

[5] Automatic stabilizers are defined as automatic changes in government revenues and expenditures that occur in response to the business cycle.

[6] See Auerbach (2002) and (2003) as well as Cohen and Follette (2003) and Falconer (2018) for its use in studying U.S. discretionary fiscal policy.

Budget Office (CBO), as my discretionary fiscal policy measure and for simplicity refer to it as the cyclically-adjusted deficit for the remainder of the paper.[7]

The variables in X are motivated by the existing literature as well as the results of Falconer (2018). It includes five business cycle measures: the CBO output gap, real GDP growth, the change in nonfarm payrolls, the CBO unemployment gap, and the change in the unemployment rate (called U-3 by the Bureau of Labor Statistics). I assume that discretionary fiscal policy is unable to respond to the business cycle within a quarter, an assumption used throughout the fiscal multiplier literature.[8] As a result, I include the first lag of each business cycle measure in all models. The other variables in X are: the level of publicly-held federal debt, the nominal federal funds rate, CPI inflation, three lags of the cyclically-adjusted deficit, and two political indicator variables. The first political variable is equal to one during presidential election years and the other is equal to one during quarters in which Congress and the presidency were controlled by the same party. Section 3 provides greater detail about variable definitions.

2.2 Threshold Models

I model regime dependence in discretionary fiscal policy conduct using threshold models of the form

$$CAD_t = \alpha + X_t'\beta + \delta D_t + (X_t D_t)'\theta + \epsilon_t \tag{4.2}$$

$$D_t = 1 \text{ if } Z_{t-1} \geq \gamma \tag{4.3}$$

[7]For ease of interpretation, I reversed the sign of this variable so that a positive value indicates a deficit and a negative value indicates a surplus.

[8]Blanchard and Perotti (2002) and Auerbach and Gorodnichenko (2012) are two prominent examples.

where CAD and X are, as before, the cyclically-adjusted deficit and the model covariates, Z is the threshold variable, and γ is the threshold value. Note that, conditional on a particular threshold variable and value, the threshold model above can be written as a linear regression model using regime indicator variable D. As a result, estimation proceeds in the same manner as for linear models.

As equations (2) and (3) make clear, I assume there are two regimes, where the current regime depends on the value of the threshold variable in the previous period. In other words, the split between regimes is determined by a particular value, γ, of the threshold variable, Z, in the previous quarter. I also assume a discrete transition between regimes, although as discussed below I allow for uncertainty about the threshold value that initiates a regime switch. Finally, I assume that if a covariate is included in a particular model, the response to that covariate is allowed to differ across regimes. Consequently each included covariate appears twice in a given threshold model: once by itself and once interacted with regime indicator variable D. This assumption reduces the model space and, correspondingly, the computational burden of implementing Bayesian model comparison and Bayesian model averaging.

In addition to allowing for uncertainty about the included covariates, for threshold models I also allow for uncertainty about the threshold variables and values that govern regime switching. As a result, threshold models differ with respect to each of these model characteristics. Table 14 provides a list of the threshold variables and range of threshold values that I consider. All threshold variables are variations of possible covariates, or closely related in concept. They enable me to consider potential regime dependence in fiscal policy responses driven by the business cycle phase, debt scenario, monetary policy regime, inflation rate,

TABLE 14. Threshold Variables and Values Considered for Model Comparison

Threshold Variable	Variable Type	Number of Values	Threshold Value Range
Recession	Indicator	1	1
Output Gap	Level	146	[-3.29,1.03]
	Change	146	[-0.66,0.68]
	Level Sign	1	1
	Change Sign	1	1
Real GDP Growth	Level	146	[0.25,6.03]
	Sign	1	1
Nonfarm Payrolls	Change	146	[-0.13,0.90]
	Change Sign	1	1
Unemployment Gap	Level	109	[-0.92,2.3]
	Sign	1	1
Unemployment Rate	Change	7	[-0.3,0.3]
	Change Sign	1	1
Publicly-Held Debt	Level	146	[32.63,77.77]
	Change	146	[-0.58,1.16]
	Change Sign	1	1
Debt Gap	Level	146	[0.001,4.18]
Debt Ceiling Increase	Indicator	1	1
Federal Funds Rate	Level	128	[0.36,9.32]
	Change	77	[-0.72,0.65]
	Change Sign	1	1
CPI Inflation	Level	146	[1.61,6.88]
	Change	146	[-1.67,1.87]
	Change Sign	1	1
Presidential Election	Indicator	1	1
President (Democrat)	Indicator	1	1
United (Either Party)	Indicator	1	1
United (Democrat)	Indicator	1	1
United (Republican)	Indicator	1	1

Notes: "Number of Values" refers to the number of threshold values I consider for a particular threshold variable. For indicator variables the only threshold value considered is 1. Variable definitions can be found in Section 3.

and political landscape. Note that the "debt gap" refers to the difference between the debt ceiling and the current level of debt.

2.3 Bayesian Model Comparison

Bayesian model comparison involves the calculation of posterior model probabilities that indicate the probability that a particular model is the underlying model that generated the data. Posterior probabilities for linear models are calculated as

$$Pr(M_j|L,Y) = \frac{p(Y|M_j,L)Pr(M_j|L)}{\sum_{i=1}^{J} p(Y|M_i,L)Pr(M_i|L)} \tag{4.4}$$

where $p(Y|M_j,L)$ is a model's marginal likelihood and $Pr(M_j|L)$ is its prior. Conditional on a particular threshold variable, Z_r, and threshold value, γ, posterior probabilities for threshold models are calculated as

$$Pr(M_j|\gamma,Z_r,NL,Y) \propto p(Y|M_j,\gamma,Z_r,NL)Pr(M_j|\gamma,Z_r,NL) \tag{4.5}$$

where $p(Y|M_j,\gamma,Z_r,NL)$ is a model's marginal likelihood and $Pr(M_j|\gamma,Z_r,NL)$ is its prior. Note that L stands for linear and NL stands for nonlinear.

A model's marginal likelihood, defined as the average value of the likelihood function over the model's parameter prior, measures how well the model explains the data. It is similar in concept to the maximized value of the likelihood function often used by frequentists for the purpose of model comparison. However, instead of measuring model performance based on a single set of parameter values, the marginal likelihood measures model performance over a range of parameter values, where the range of values is determined by the parameter prior. As is well known in the Bayesian literature, larger models are automatically penalized in the marginal likelihood, meaning that, *ceteris paribus*, smaller models have higher posterior probabilities than larger models.[9]

A model's prior represents the researcher's beliefs about the model before seeing the data. If each model has the same prior probability, the posterior model probabilities in (4) and (5) depend solely on each model's marginal likelihood. Conversely, if two models have the same marginal likelihood, the model with the higher prior probability with also have the higher posterior probability. As discussed below, I present results for two different model priors, each of which is relatively agnostic about the underlying model.

Closely related to posterior model probabilities are inclusion probabilities, which indicate the probability that a particular covariate is in the underlying model. In the current context, inclusion probabilities enable me to determine which variables matter to fiscal policymakers. Inclusion probabilities are easily derived from posterior model probabilities by summing posterior probabilities across models that include each covariate.

Marginal likelihoods can be used to obtain posterior probabilities for other objects of interest. For example, in order to determine whether the underlying model is linear or not, I calculate posterior probabilities for the *set* of threshold models as

$$Pr(NL|Y) \propto p(Y|NL)Pr(NL) \tag{4.6}$$

where $p(Y|NL)$ is the marginal likelihood for the set of threshold models and $Pr(NL)$ is the prior probability that the underlying model is a threshold model. Just as the marginal likelihood for a single model averages the likelihood function over the parameter prior, the marginal likelihood for the set of threshold models

[9] See Koop (2003) for further explanation.

averages marginal likelihoods over model, threshold variable, and threshold value priors. Specifically,

$$p(Y|NL) = \sum_r \int_\gamma \sum_j p(Y|M_j, \gamma, Z_r, NL) Pr(M_j|\gamma, Z_r, NL) p(\gamma|Z_r, NL) d\gamma Pr(Z_r|NL) \tag{4.7}$$

Although γ is a continuous parameter, its effect on the likelihood function is the same as if it were discrete. This is due to the fact that the split between regimes remains the same between observed values of a particular threshold variable. To use the example in Koop and Potter (1997), if the threshold value is chosen as the 25th smallest observation of the threshold variable, the split between regimes is constant for alternate threshold values that fall between the 25th and 26th smallest observations. Consequently, marginal likelihoods only need to be calculated for the values of Z_r that are actually observed. In order to integrate marginal likelihoods over the prior for γ, marginal likelihoods calculated for each observed value need to be multiplied by the width of the interval in which the split of the data remains constant and divided by the width of the interval that receives nonzero prior probability in $p(\gamma|Z_r, NL)$.

Posterior probabilities for threshold variables and threshold values are calculated as

$$Pr(Z_r|NL, Y) \propto p(Y|Z_r, NL) Pr(Z_r|NL) \tag{4.8}$$

$$Pr(\gamma|Z_r, NL, Y) \propto p(Y|\gamma, Z_r, NL) p(\gamma|Z_r, NL) \tag{4.9}$$

As in (7), the marginal likelihoods in (8) and (9) are calculated by averaging marginal likelihoods for individual models across models and, for $p(Y|Z_r, NL)$, averaging across different values of γ, using the associated prior probabilities as weights.

2.4 Bayesian Model Averaging

In addition to posterior probabilities that highlight characteristics of the underlying model, I am interested in the sign and magnitude of policy responses to the included covariates. My results indicate that there is uncertainty about the underlying model of fiscal policy, and that this uncertainty extends to a number of different model characteristics. Consequently, focusing on results from a single model, even if that model has the highest posterior probability, may be unwise. Doing so is especially problematic if multiple plausible models yield different or even conflicting results.

When there are multiple plausible models, Bayesian model averaging (BMA) can be used to incorporate uncertainty about those models into parameter estimates. As its name suggests, BMA proceeds by averaging results across alternative models using posterior model probabilities as weights. For example, BMA estimates for linear models are calculated as

$$p(\lambda|L,Y) = \sum_{j=1}^{J} p(\lambda|M_j,L,Y)Pr(M_j|L,Y) \tag{4.10}$$

Note that the BMA estimates in (10) are not conditioned on a particular model, which reduces the likelihood that inferences are driven by model choice. They also reflect uncertainty about the underlying model since each set of results are weighted by the model's posterior probability. This potentially important source of uncertainty is ignored when estimates come from a single model.

BMA estimates for threshold models are calculated in the same way as for linear models, although averaging occurs across models with different covariates and threshold values. I do not average results across models with different threshold variables because they may capture different types of regime switching.

Instead, I present results for the threshold variables with the highest posterior probabilities.

The total number of linear models is equal to the total number of covariate combinations: $2^{13} = 8,192$. For a given threshold variable, the total number of threshold models is equal to $2^{13} \times V$ where V is the total number of threshold values under consideration. Using the information in Table 14, this yields a total of 13,524,992 threshold models. Although these numbers are large, modern computing speeds combined with the priors I use (discussed below) enable me to calculate results analytically.

2.5 Priors

In order to calculate the posterior quantities in the previous section, I require prior probabilities for each linear and threshold model, $Pr(M_j|L)$ and $Pr(M_j|\gamma, Z_r, NL)$; prior density functions for each set of model parameters, $p(\alpha, h, \beta_j|M_j, L)$ and $p(\alpha, h, \beta_j|M_j, \gamma, Z_r, NL)$; prior probabilities for each threshold variable, $Pr(Z_r|NL)$; prior density functions for threshold values, $p(\gamma|Z_r, NL)$; and prior probabilities for the set of linear and threshold models, $Pr(L)$ and $Pr(NL)$. Each threshold model uses the same model prior as the corresponding linear model.[10] Parameter priors are obtained in the same way for linear and threshold models, although each threshold model has k_j more parameters than its linear counterpart. All other priors are associated with a single model type (linear or threshold).

[10]Each included covariate appears twice in a given threshold model: once by itself and once interacted with the regime indicator variable, D. Consequently, conditional on a particular threshold variable and threshold value the number of nonlinear models is equal to the number of linear models, and the number of nonlinear model sizes is equal to the number of linear model sizes. This enables me to use the same model prior for linear and nonlinear models.

I present results for two different model priors. The first assigns equal prior probability to each model while the second assigns equal prior probability to each model size.[11] For both priors I assume that a particular covariate enters the true model independently of all other covariates with probability θ, implying a model prior of the form:

$$\Pr(M_j) = \theta^{k_j}(1-\theta)^{K-k_j} \tag{4.11}$$

For the first model prior I set $\theta = 0.5$, so that

$$\Pr(M_j) = \frac{1}{J}, \quad j = 1, \cdots, J \tag{4.12}$$

Although the prior in (12) is uniform across models, it favors moderately-sized models because they outnumber small and large models. Indeed, the Binomial distribution for model size implied by (12) peaks at K/2.

The second prior that I consider, suggested in Ley and Steel (2009), assigns a prior to θ rather than fixing it at a particular value. Specifically,

$$\theta \sim \text{Beta}(a, b) \tag{4.13}$$

I set $a = 1$ and $b = 1$, which yields a uniform prior for model size. Unlike the other model prior, this prior favors small and large models over moderately-sized models.

Since the model space is relatively large, I use the automatic procedure outlined in Fernandez, Ley, and Steel (2001b), hereafter FLS, to set priors for model parameters. FLS recommend using this procedure when estimating groups of linear regression models that differ only with respect to the included covariates.

[11] Falconer (2018) provides greater detail about these priors than what is presented here.

For each model, improper non-informative priors are used for parameters common to all models (α and h) and informative priors are used for those that are not (β_j). Specifically,

$$p(h|M_j) \propto h^{-1} \tag{4.14}$$

$$p(\alpha|M_j) \propto 1 \tag{4.15}$$

$$\beta_j|h, M_j \sim N(\underline{\beta}_j, h^{-1}(\underline{g}X_j'X_j)^{-1}) \tag{4.16}$$

where $(\underline{g}X_j'X_j)^{-1}$ is the g-prior of Zellner (1986). I set $\underline{\beta}_j = 0_{k_j}$ and $\underline{g} = 1/\max\{T, K^2\}$, again as recommended in FLS.

Note that the "benchmark" prior in (14)-(16) contains a single hyperparameter, $\underline{g}$, whose value is chosen based on prespecified criteria. The FLS procedure is useful because it requires little subjective information from the researcher and, as the authors demonstrate, generates priors that have little influence on posterior quantities. Simulations indicate that these priors also produce accurate posterior model probabilities.

I assign equal prior probabilities to the set of linear and threshold models, so that

$$Pr(L) = Pr(NL) = \frac{1}{2} \tag{4.17}$$

I do the same for each threshold variable:

$$Pr(Z_r|NL) = \frac{1}{R}, \quad r = 1, \cdots, R \tag{4.18}$$

For threshold values, I use a uniform prior over an interval that ensures that at least 14% of observations are in each regime. In other words, when the threshold variable observations are ordered from smallest to largest, the prior is uniform across the interval that contains the center 72% of observations. Values outside of this interval receive no prior probability. By preventing extreme values from being considered as threshold values, this prior reduces the likelihood that results are driven by outliers. I chose a value of 14% (rather than 15% as recommended in, for example, Koop and Potter (1997)) because 14% of the observations in my sample occur during a recession. Using a larger value to determine the prior interval eliminates values of the considered threshold variables that typically occur during or near recessions, which my results suggest drive regime switching in fiscal policy conduct.

3. Data

I use quarterly data covering the period 1966q1-2016q3. Table 15 lists the source of each of the data series used to construct my final variables. The cyclically-adjusted deficit, debt, and debt gap are expressed as percentages of potential GDP (as estimated by the CBO). An augmented Dickey-Fuller test indicates that the cyclically-adjusted deficit is stationary so I include it and its lags in levels.[12] The change in nonfarm payrolls is calculated as a quarterly growth rate and the change in the unemployment rate is calculated as a quarterly change (in percentage points). Inflation is calculated as quarterly growth in the Consumer Price Index and annualized. Similarly, real GDP growth is calculated as quarterly growth and annualized. Finally, the political indicator variables take on a value of one during presidential election years, quarters in which the president was a Democrat, quarters in which Congress and the presidency were controlled by

[12] First-differencing the cyclically-adjusted deficit changes the estimated coefficients on the policy lags but otherwise has no effect on results.

TABLE 15. Data Sources

Variable	Source
Cyclically-adjusted net federal government savings	Congressional Budget Office
Nominal Potential GDP	Congressional Budget Office
Nominal GDP	Bureau of Economic Analysis
Nonfarm Payrolls	Bureau of Labor Statistics
Unemployment Rate	Bureau of Labor Statistics
Natural Rate of Unemployment	Congressional Budget Office
Recession Indicator	National Bureau of Economic Research
Publicly-Held Debt	U.S. Treasury Department
Debt Ceiling	White House
Federal Funds Rate	Federal Reserve Board of Governors
Consumer Price Index	Bureau of Labor Statistics
Presidential Election Indicator	The American Presidency Project
Democratic President	The American Presidency Project
Republican President	The American Presidency Project
United Democrat Indicator	The American Presidency Project
United Republican Indicator	The American Presidency Project

the Democratic party, and quarters in which Congress and the presidency were controlled by the Republican party.[13]

4. Results

4.1 Evidence for Regime Switching

Table 16 presents posterior probabilities for the set of threshold models using different model type prior probabilities and model priors. When the linear and threshold models are equally likely under the prior, the posterior probability for the threshold models is 86% and 100%, respectively, for the different model priors. As the table makes clear, the set of threshold models has to receive substantially

[13]I consider a particular party to be in control of the House of Representatives and Senate if at least 50% of the seats in each chamber of Congress are held by that party.

TABLE 16. Posterior Probabilities for the Set of Threshold Models

	Model Prior=1	**Model Prior=2**
$Pr(NL) = 0.5$	85.59%	100%
$Pr(NL) = 0.25$	66.45%	100%
$Pr(NL) = 0.14$	50%	99.98%
$Pr(NL) = 0.0001$	0.06%	79.43%

Notes: "Model Prior=1" assigns equal prior probability to different model sizes while "Model Prior=2" assigns equal prior probability to each model.

lower prior probability before posterior probabilities favor the set of linear models. Note, too, that each threshold model has twice as many covariates as its linear counterpart. These models are penalized for their larger size in both the marginal likelihood and the model prior that assigns equal prior probability to different model sizes ("Model Prior=1").[14] Given that, *ceteris paribus*, linear models have higher posterior probabilities than threshold models, the posterior probabilities in Table 16 provide relatively strong evidence that the underlying model exhibits regime switching.

4.2 Variables Characterizing Regime Switching

Table 17 lists posterior probabilities for the threshold variables in Z. Among the twenty-nine candidates, the change in nonfarm payrolls and the change in the federal funds rate are the most likely drivers of regime change. Without further information it is difficult to determine whether both variables define similar regimes but this could occur if the Federal Reserve responds directly (and

[14] "Model Prior=1" actually favors both small and large models over moderately-sized models. However, since the threshold models with the highest marginal likelihoods contain only a subset of the possible covariates, these moderately-sized models are penalized for their larger size.

TABLE 17. Posterior Probabilities for Different Threshold Variables

Threshold Variable		**Model Prior=1**	**Model Prior=2**
Recession	Indicator	0.34%	0.08%
Output Gap	Level	0%	0%
	Change	0%	0%
	Level Sign	0%	0%
	Change Sign	0%	0%
Real GDP Growth	Level	0%	0%
	Sign	0.01%	0%
Nonfarm Payrolls	Change	**57.75%**	**51.60%**
	Change Sign	**3.89%**	**5.01%**
Unemployment Gap	Level	0%	0%
	Sign	0%	0%
Unemployment Rate	Change	0.71%	0.52%
	Change Sign	0.04%	0.05%
Publicly-Held Debt	Level	0%	0%
	Change	0%	0%
	Change Sign	0%	0%
Debt Gap	Level	1.17%	0.56%
Debt Ceiling Increase	Indicator	0%	0%
Federal Funds Rate	Level	0%	0%
	Change	**36.01%**	**42.15%**
	Change Sign	0%	0%
CPI Inflation	Level	0%	0%
	Change	0.02%	0.01%
	Change Sign	0.02%	0%
Presidential Election	Indicator	0%	0%
President (Democrat)	Indicator	0.04%	0%
United (Either Party)	Indicator	0%	0%
United (Democrat)	Indicator	0%	0%
United (Republican)	Indicator	0%	0%

Notes: Bolded values indicate a posterior probability greater than 3%. "Model Prior=1" assigns equal prior probability to different model sizes while "Model Prior=2" assigns equal prior probability to each model.

contemporaneously) to the change in nonfarm payrolls, or to changing economic conditions that are also reflected in employment changes. For example, if monetary policymakers decrease the federal funds rate in response to a decrease in nonfarm payrolls, fiscal policymakers may interpret the central bank's actions as a signal that the economy is worsening and respond by drastically altering their own policy. Alternatively, fiscal policymakers may respond directly to the change in nonfarm payrolls, in which case the seeming importance of the change in the federal funds rate is merely due to the fact that it responds to the same variable as fiscal policymakers. The evidence presented in the following sections suggests that both variables do, in fact, define similar regimes.

The fact that the change in the unemployment rate has a low probability of triggering regime switches is surprising given that it is highly correlated with the change in nonfarm payrolls. However, policymakers may view the change in nonfarm payrolls as a more reliable indicator of labor market conditions, particularly during expansions. While the signs of both variables are nearly always in agreement during periods of weak economic growth (meaning that a decrease in nonfarm payrolls is nearly always accompanied by an increase in the unemployment rate), the same is not true during expansions. Instead, the unemployment rate occasionally increases during periods in which employment gains are consistently positive. Policymakers may hesitate to enact large changes in policy following a single quarter of unemployment increases, preferring instead to focus on the direction of employment growth. Returning to Table 17, the fact that the sign of the change in nonfarm payrolls has a higher posterior probability than the sign of the change in the unemployment rate provides further evidence in support of this view.

It is notable that the signs of the output gap and GDP growth receive virtually no posterior probability given that these variables have often been used in the literature to characterize regime switching.[15] Similarly, the recession indicator variable used in Falconer (2018) also does a poor job of defining different fiscal policy regimes. Finally, fiscal policy does not appear to change noticeably in response to the current debt situation, a finding at odds with the "active" and "passive" fiscal policy regimes as defined in the fiscal and monetary policy switching literature.

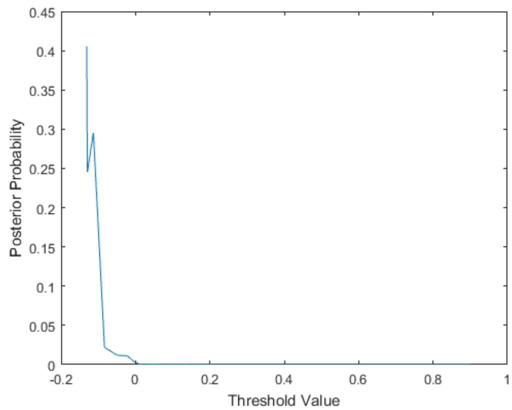
4.3 Threshold Values

Figure 9 presents posterior distributions for the threshold value when the change in nonfarm payrolls and the change in the federal funds rate are used as threshold variables. Extreme negative values receive the highest posterior probability for both variables. Altering the threshold value priors, $p(\gamma|Z_r, NL)$, to allow regimes to contain fewer than 14% of observations again yields the result that the most negative values are determined to be the most probable. While this means that the prior, in effect, determines the specific values that receive the highest posterior probability, the conclusions resulting from different priors are the same. Specifically, since decreases in both nonfarm payrolls and the federal funds rate typically occur during periods of weak economic growth, the two regimes defined by these threshold values are consistent with a "normal" and "bad" economy.

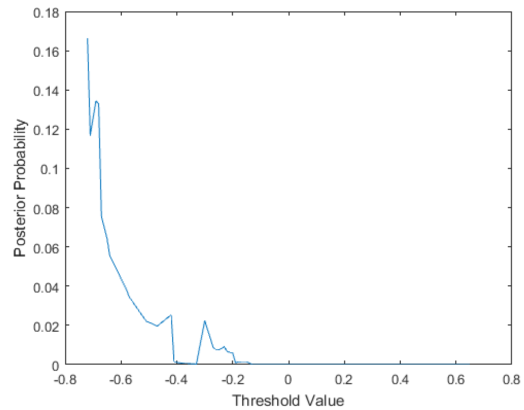
Figure 10 illustrates the probability that the economy was in the "bad" regime over the entire sample, accounting for uncertainty about the threshold value. To calculate these probabilities I create, for each possible threshold value, a $T \times 1$

[15] See, for example, Cimadomo (2007), Forni and Momigliano (2004), Cohen and Follette (2003), and Egert (2010).

FIGURE 9. Threshold Value Posterior Distributions



Change in Nonfarm Payrolls



Change in the Federal Funds Rate

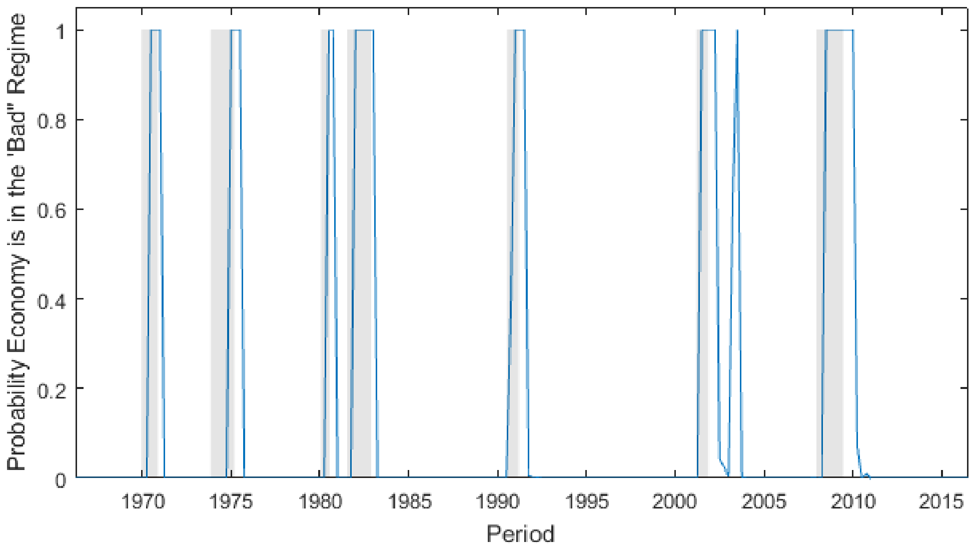
Notes: The change in nonfarm payrolls is defined as a quarterly growth rate. The considered values range from -0.13% to 0.90%. The change in the federal funds rate is measured in percentage points. The considered values range from -0.72 to 0.65.

vector indicating the quarters in which the economy was in the "bad" regime. In other words, entry $t = 1, \cdots, T$ is equal to 1 if the value of the threshold variable was smaller than the threshold value in period $t-1$. I then average across vectors using threshold value posterior probabilities as weights.[16]

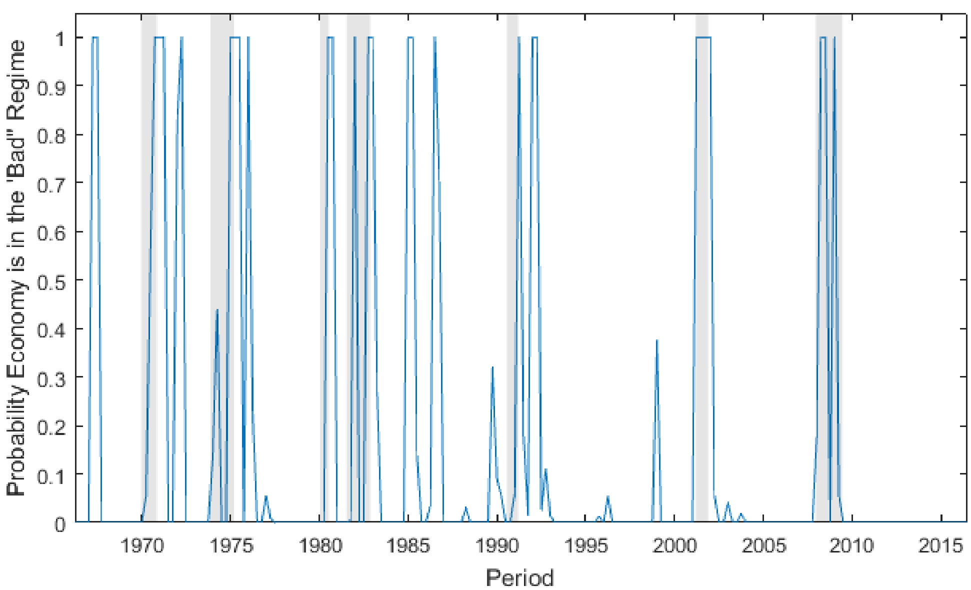
This figure indicates that both threshold variables define similar regimes, which appear to be closely related to recession dates. However, there are some apparent differences between the two pictures. First, the number of "bad" episodes is smaller when the change in nonfarm payrolls is used as the threshold variable, but the duration of each event is typically longer. These events also correspond more closely to recession dates than those identified using the change in the federal funds rate. Second, there is less uncertainty about whether the economy was in a particular regime using the change in nonfarm payrolls: the probability of being

[16]Note that this vector is simply $\iota_T - D$ where D is the $T \times 1$ indicator variable used in each threshold model.

FIGURE 10. Probability of Being in the “Bad” Regime



Change in Nonfarm Payrolls



Change in the Federal Funds Rate

Notes: The probabilities above account for uncertainty about the threshold value.

in the “bad” regime is nearly always 100% or less than 1%. In contrast, when the change in the federal funds rate is used there are a number of quarters in which the probability of being in the “bad” regime is between 5% and 85%. These differences reflect the fact that there is less uncertainty about the threshold value for the change in nonfarm payrolls. Finally, the economy spends slightly less time in the “bad” regime when the change in nonfarm payrolls is used as the threshold value. This difference is small, however: the probability that the economy was in the “bad” regime exceeds 50% for 14.4% of observations for the change in nonfarm payrolls compared to 15.3% of observations for the change in the federal funds rate.

4.4 Inclusion Probabilities and Posterior Means

Table 18 lists inclusion probabilities for possible covariates in both model types (linear and threshold). These probabilities indicate the likelihood that a particular covariate is in the underlying model and, therefore, whether that variable appears to be important to policymakers. Note that the probabilities in Table 18 are averaged across all possible threshold variables and threshold values using posterior probabilities for these model characteristics as weights.

Inclusion probabilities for threshold models closely resemble those for linear models. They suggest that policy is persistent and responsive to changes in the labor market: the probability that either the change in the unemployment rate or the change in nonfarm payrolls (or both) are in the underlying model exceeds 95% for both model types and model priors. Posterior probability is shared more equally between these two variables in the threshold case due to the fact that the inclusion probability for the change in nonfarm payrolls is greater than 99% when this variable is used as the threshold variable, while the inclusion probability for the change in the unemployment rate is less than 1%. This causes the change in

TABLE 18. Inclusion Probabilities for Threshold and Linear Models

	Threshold		Linear	
Variable	**Prior=1**	**Prior=2**	**Prior=1**	**Prior=2**
CA Deficit First Lag	**100%**	**100%**	**100%**	**100%**
CA Deficit Second Lag	**70.76%**	**87.84%**	**94.63%**	**97.52%**
Output Gap	0.45%	0.87%	12.23%	19.94%
Real GDP Growth	0.40%	0.53%	3.94%	7.07%
Payrolls Change	**39.57%**	**39.80%**	24.94%	23.51%
Unemployment Gap	0.23%	0.57%	7.86%	13.17%
Unemployment Change	**55.70%**	**60.12%**	**79.58%**	**84.78%**
Debt	0.32%	0.85%	8.08%	15.99%
Federal Funds Rate	0.93%	2.45%	6.39%	11.28%
Inflation	1.10%	3.04%	25.94%	**41.81%**
Presidential Election	0.17%	0.46%	5.09%	9.36%
United Government	0.16%	0.44%	5.25%	9.63%
Unemployment or Payrolls Change	**95.09%**	**99.4%**	**99.57%**	**99.82%**

Notes: Inclusion probabilities for threshold models are averaged across threshold variables and threshold values using posterior probabilities $Pr(\gamma|Z_r, NL, Y)$ and $Pr(Z_r|NL, Y)$ as weights.

nonfarm payrolls to be weighted more heavily in the averaging process because it has a relatively high probability of being the threshold variable.

As in Falconer (2018), policy appears to be unresponsive to the other covariates that I consider. In general, posterior probabilities for these variables are smaller for the threshold models than for the linear models because additional variables are penalized more in this model type. However, both model types produce similar conclusions about the relative importance of each covariate to policymakers.

Table 19 presents coefficient posterior means for the "bad" and "normal" regimes using the two most likely threshold variables. Posteriors are averaged across all possible covariate combinations and threshold values using joint posterior probabilities $Pr(M_j, \gamma|Z_r, NL, Y)$ as weights. Consequently coefficients on variables

TABLE 19. Posterior Means Averaged Across Models and Thresholds

	Payrolls Change		**FFR Change**	
	"Bad" Regime	**"Normal" Regime**	**"Bad" Regime**	**"Normal" Regime**
CA Deficit First Lag	**0.32**	**0.95**	**0.12**	**0.85**
CA Deficit Second Lag	0.08	0.03	**0.71**	**0.03**
Output Gap	0.01	0.00	0.00	0.00
Real GDP Growth	0.05	-0.01	0.00	0.00
Payrolls Change	**-2.11**	**-0.08**	**-0.15**	**-0.04**
Unemployment Gap	0.00	0.00	0.00	0.00
Unemployment Change	0.03	0.00	**0.80**	**0.15**
Debt	0.00	0.00	0.00	0.00
Federal Funds Rate	-0.02	0.00	**-0.08**	**0.00**
Inflation	0.00	0.00	0.00	0.00
Presidential Election	0.00	0.00	0.00	0.00
United Government	0.00	0.00	-0.05	0.01

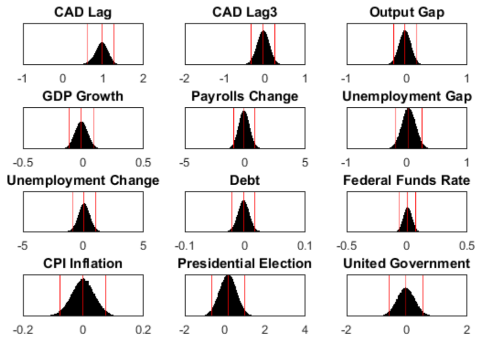
Notes: Posterior means are averaged across all possible covariate combinations and threshold values using $Pr(M_j|Z_r, NL, Y)$ and $p(\gamma|Z_r, NL, Y)$ as weights. Variables with inclusion probabilities greater than 30% are bolded.

with low inclusion probabilities are shrunk towards zero. Posterior distributions for coefficients conditional on the associated variable being in the model are located in Figure 11. These graphs ignore the large point mass at zero for covariates with inclusion probabilities below 100%.

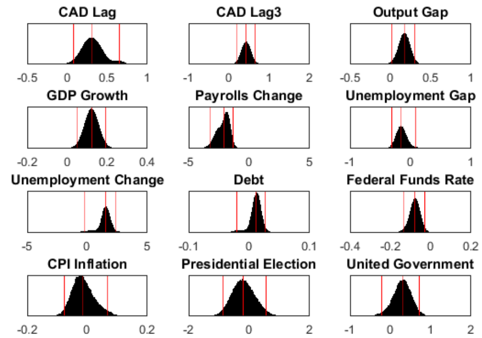
Both sets of results are largely consistent with the findings in Falconer (2018). They indicate that during the "normal" regime discretionary fiscal policy is very persistent and slightly countercyclical or acyclical, whereas during the "bad" regime policy is less persistent and strongly countercyclical.[17] Specifically, an acceleration

[17]The posterior mean for the coefficient on the first lag of the dependent variable suggests that this variable is nonstationary during the "normal" regime. As mentioned in section 3, first-differencing the dependent variable has very little effect on overall results, with the obvious exception that coefficients and inclusion probabilities for the lags of the dependent variable change.

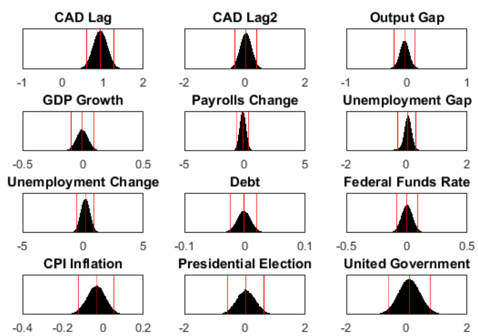
FIGURE 11. Coefficient Posterior Distributions Conditional on Inclusion



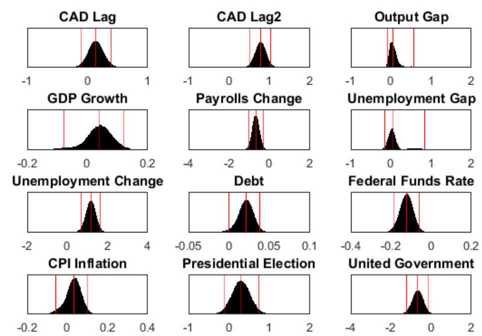
"Normal" Regime, Z_r: Payrolls Change



"Bad" Regime, Z_r: Payrolls Change



"Normal" Regime, Z_r: FFR Change



"Bad" Regime, Z_r: FFR Change

Notes: The posterior distributions above ignore the large point mass that occurs at zero for variables with inclusion probabilities less than 100%. The left and right red lines designate the 95% central posterior interval while the middle red line is the posterior median.

in unemployment increases and payroll decreases leads to expansionary policy (an increase in the cyclically-adjusted deficit).

Since each business cycle measure is expressed in different units, it is difficult to compare the magnitude of responses to these variables across threshold variables using the coefficients in Table 19. To facilitate that comparison, Table 20 presents, for each threshold variable and regime, posterior means for the implied response

TABLE 20. Business Cycle Responses For Different Threshold Variables

	Payrolls Change	**FFR Change**
"Normal" Regime	-$13.06	-$13.83
"Bad" Regime	$181.35	$72.23

Notes: Responses are expressed in billions of 2016q3 dollars. For the "normal" regime I calculate the total response to one-standard deviation increases in the output gap, GDP growth, and the change in nonfarm payrolls and one-standard deviation decreases in the unemployment gap and unemployment change. For the "bad" regime I consider changes of the opposite sign.

(in billions of 2016q3 dollars) of the cyclically-adjusted deficit to one-standard deviation changes in each business cycle measure.[18] For the "bad" regime I consider the total response to one-standard deviation decreases in the output gap, GDP growth, and nonfarm payrolls change, and one-standard deviation increases in the unemployment gap and unemployment change. I consider changes of the opposite sign for the "normal" regime. Table 20 indicates that although business cycle responses are of a similar magnitude during the "normal" regime defined by each threshold variable, the same is not true of responses during the "bad" regime. Instead, the response is nearly 2.5 times larger when the change in nonfarm payrolls is used as the threshold variable than when the change in the federal funds rate is used.

Turning to results for the other variables, Table 19 and Figure 11 indicate that fiscal policy mimics monetary policy during the "bad" regime but not during the "normal" regime. During the "bad" regime fiscal policymakers implement expansionary policy in the quarter following expansionary monetary policy. It is possible that in doing so fiscal policymakers are responding directly to monetary

[18]In order to convert estimated coefficients into 2016q3 dollars I multiply each coefficient, which measures the response of the cyclically-adjusted deficit as a percentage of potential GDP, by the value of potential GDP in 2016q3.

policy. However, another possibility is that fiscal and monetary policymakers have similar ideas about how to respond to a weak economy but fiscal policymakers are unable to respond as quickly as monetary policymakers.

Finally, for both threshold variables the coefficient on the "United Government" variable is negative during the "bad" regime, suggesting that policy responds less to a weak economy when Congress and the presidency are controlled by the same party. It should be noted that this coefficient is estimated off of only nine observations and appears to be driven primarily by 1980q3-q4; deleting these observations results in an averaged coefficient very close to zero.[19] The instability of this estimate combined with the fact that the "United Government" variable has a very low inclusion probability suggests that one should take care when interpreting this coefficient.

5. Conclusion

In recent years regime-dependent fiscal multipliers have received a great deal of attention in empirical research. Regime dependence in fiscal policy conduct has received far less attention, although there is evidence to suggest that discretionary policy responses vary according to the state of the business cycle and debt levels. These findings rely on strong assumptions about the structure of the underlying model, however; the variables and corresponding values that induce regime switches are often imposed rather than estimated. I extend the existing literature by using Bayesian model comparison techniques to compare linear and threshold models that differ with respect to the included covariates and, for threshold models, the threshold variable and value that initiate a regime switch. Posterior probabilities

[19]The cyclically-adjusted deficit actually increased in both 1980q3 and 1980q4 but its growth was outpaced by the growth in potential GDP. Since the cyclically-adjusted deficit is expressed as a percentage of potential GDP in all the models I consider its value fell.

obtained from this process enable me to determine which model type (linear or threshold), covariate combination, threshold variable, and threshold values are most plausible. I consider twenty-nine possible threshold variables related to the business cycle, debt, monetary policy, inflation, and political environment.

My results indicate that fiscal policy is regime dependent: the set of threshold models is at least five times more probable than the set of linear models. Among the different threshold variables, the change in nonfarm payrolls and the change in the federal funds rate are the most likely drivers of regime change. Notably, the sign of the output gap and the five debt variables that I consider have a very low probability of determining regimes. Threshold value posteriors favor extreme negative values when either the change in nonfarm payrolls or the change in the federal funds rate are used as the threshold variable, splitting observations into periods of "normal" and "bad" economic performance that roughly coincide with recession dates. While there are a lot of similarities between the regimes defined by either threshold variable, the number of "bad" episodes that occurred in the United States between 1966 and 2016 is smaller using the change in nonfarm payrolls, but the duration of each episode is longer. Finally, policy is strongly countercyclical during the "bad" regime while at other times it is slightly countercyclical or acyclical.

My findings suggest that discretionary fiscal policymakers, knowingly or not, may benefit from the fact that large policy responses are limited primarily to periods of weak economic activity when such policy is likely to be most effective. In the future, it would be useful to explore this idea further by determining whether the same variables that govern regime switching in policy conduct determine regime switching in fiscal policy effectiveness. If not, fiscal policymakers may benefit from modifying policy responses in a way that more closely corresponds to periods of high policy effectiveness. Although there is some overlap between the threshold variables considered in this paper and a paper by Fazzari et al (2015) on regime-dependent multipliers, they do not consider either the change in nonfarm payrolls or the change in the federal funds rate, suggesting that this topic warrants further research.

REFERENCES CITED

Andrews, D. (1993). Tests for Parameter Instability and Structural Change With Unknown Change Point. *Econometrica*, *16*(4), 821-856.

Aruoba, S. (2008). Data Revisions are Not Well-Behaved *Journal of Money, Credit, and Banking*, *40*(2-3), 319-340.

Auerbach, A. (2002). Is There a Role for Discretionary Fiscal Policy? *NBER Working Paper*.

Auerbach, A. (2003). Fiscal Policy, Past and Present. *NBER Working Paper*.

Auerbach, A. & Gorodnichenko, Y. (2012). Measuring the Output Responses to Fiscal Policy. *AEJ: Economic Policy*, *4*(2), 1-27.

Belinga, V. & Lonkeng Ngouana, C. (2015). (Not) Dancing Together: Monetary Policy Stance and the Government Spending Multiplier. *IMF Working Paper*.

Benetrix, A. & Lane, P. (2013). Fiscal Cyclicality and EMU. *Journal of International Money and Finance*, *34*, 164-176.

Bernanke, B. & Boivin, J. (2003). Monetary Policy in a Data-Rich Environment. *Journal of Monetary Economics*, *50*, 525-546.

Bernoth, K., Hallett, A., & Lewis, J. (2015). The Cyclicality of Automatic and Discretionary Fiscal Policy: What Can Real-Time Data Tell Us? *Macroeconomic Dynamics*, *19*, 221-243.

Blanchard, O. & Perotti, R. (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.

Bohn, H. (1998). The Behavior of U.S. Public Debt and Deficits. *Quarterly Journal of Economics*, *113*(3), 949-963.

Caldara, D. & Kamps, C. (2008). What are the Effects of Fiscal Policy Shocks? A VAR-Based Comparative Analysis. *European Central Bank Working Paper* 877.

Canova, F. & Pappa, E. (2011). Fiscal Policy, Pricing Frictions and Monetary Accommodation. *Economic Policy*, *26*(68), 555-598.

Check, A. (2016). The Taylor Rule or a Tailor-Made Rule? *Working Paper*.

Christiano, L., Eichenbaum, M., & Rebelo, S. (2011). When is the Government Spending Multiplier Large? *Journal of Political Economy*, *119*(1), 78-121.

Cimadomo, J. (2007). Fiscal Policy in Real Time. *CEPII Working Paper* 2007-10.

Cohen, D. & Follette, G. (2003). Forecasting Exogenous Fiscal Variables in the United States. *Federal Reserve Working Paper*.

Combes, J., Minea, A., & Sow, M. (2017). Is Fiscal Policy Always Counter- (Pro-) Cyclical? The Role of Public Debt and Fiscal Rules. *Economic Modelling*, *65*, 138-146.

Darby, J. & Melitz, J. (2011). Joint Estimates of Automatic and Discretionary Fiscal Policy: the OECD 1981-2003. *CEPII Working Paper*.

Egert, B. (2014). Fiscal Policy Reactions to the Cycle in the OECD: Pro- or Countercyclical? *Mondes en developpement*, *167*, 35-52.

Falconer, J. (2018). Cyclicality of U.S. Discretionary Fiscal Policy. *Working Paper*.

Fatas, A. & Mihov, I. (2012). Fiscal Policy as a Stabilization Tool. *B.E. Journal of Macroeconomics*, *12*(3), 1-68.

Fazzari, S., Morley, J., & Panovska, I. (2015). State-Dependent Effects of Fiscal Policy. *Studies in Nonlinear Dynamics and Econometrics*, *19*(3), 285-315.

Fernandez, C, Ley, E. & Steel, M. (2001b). Benchmark Priors for Bayesian Model Averaging. *Journal of Econometrics*, *100*, 381-427.

Forni, L. & Momigliano, S. (2004). Cyclical Sensitivity of Fiscal Policies Based on Real-Time Data. *Applied Economics Quarterly*, *50*(3), 299-326.

Gali, J. & Perotti, R. (2003). Fiscal Policy and Monetary Integration in Europe. *Economic Policy*, *18*(37), 533-572.

Garcia, A., Arroyo, M., Minguez, R., & Uxo, J. (2009). Estimation of a Fiscal Policy Rule for EMU Countries (1984-2005). *Applied Economics*, *41*, 869-884.

Golinelli, R. & Momigliano, S. (2009). The Cyclical Reaction of Fiscal Policies in the Euro Area: The Role of Modelling Choices and Data Vintages. *The Journal of Applied Public Economics*, *30*(1), 39-72.

Guajardo, J., Leigh, D., & Pescatori, A. (2014). Expansionary Austerity: New International Evidence. *Journal of the European Economic Association*, *12*(4), 949-68.

Hallerberg, M. & Strauch, R. (2002). On the Cyclicality of Public Finances in Europe *Empirica*, *29*, 183-207.

Hamilton, J. (2017). Why You Should Never Use the Hodrick-Prescott Filter. *Working Paper*, University of California San Diego.

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

Koop, G. (2003). Bayesian Econometrics. *John Wiley and Sons Ltd.*, England.

Koop, G. & Potter, S. (2000). Nonlinearity, Structural Breaks or Outliers in Economic Time Series? *Nonlinear econometric modeling in time series analysis*, Cambridge University Press.

Lane, P. (2003). The Cyclical Behavior of Fiscal Policy: Evidence from the OECD. *Journal of Public Economics*, *87*, 2661-2675.

Ley, E. & Steel, M. (2009). On the Effect of Prior Assumptions in Bayesian Model Averaging with Applications to Growth Regression. *Journal of Applied Econometrics*, *24*(4), 651-674.

Madigan, D. & York, J. (1995). Bayesian Graphical Models for Discrete Data. *International Statistics Review*, *63*(2), 215-232.

Mertens, K. & Ravn, M. (2014). A Reconciliation of SVAR and Narrative Estimates of Tax Multipliers. *Journal of Monetary Economics*, *68*, S1-S19.

Mink, M. & de Haan, J. (2006). Are there Political Budget Cycles in the Euro Area? *European Union Politics*, *7*(2), 191-211.

Muscatelli, A., Tirelli, P., & Trecroci, C. (2002). Monetary and Fiscal Policy Interactions over the Cycle: Some Empirical Evidence. *University of Glasgow Working Paper.*

National Economic Accounts. *http://www.bea.gov/national/.* Bureau of Economic Accounts.

Orphanides, A. (2001). Monetary Policy Rules Based on Real-Time Data. *American Economic Review*, *91*(4), 964-985.

Orphanides, A. & van Norden, S. (2002). 'The Unreliability of Output-Gap Estimates in Real Time. *The Review of Economics and Statistics*, *84*(4), 569-583.

Perotti, R. (2004). Estimating the Effects of Fiscal Policy in OECD Countries. *European Central Bank Working Paper* 168.

Perotti, R. (2011). Expectations and Fiscal Policy: An Empirical Investigation. *IGIER Working Paper 429.*

Perotti, R. (2014). Defense Government Spending is Contractionary, Civilian Government Spending is Expansionary. *NBER Working Paper* 20179.

Plodt, M. & Reicher, C. (2015). Estimating Fiscal Policy Reaction Functions: the Role of Model Specification. *Journal of Macroeconomics*, *46*, 113-128.

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

Ramey, V. & Zubairy, S. (2018). Government Spending Multipliers in Good Times and Bad: Evidence from US Historical Data. *Journal of Political Economy*, *126*(2), 850-901.

Romer, C. & Romer, D. (2010). The Macroeconomic Effects of Tax Changes: Estimates Based on a New Measure of Fiscal Shocks. *American Economic Review*, *100*, 763-801.

Romer, C. & Romer, D. (2016). Transfer Payments and the Macroeconomy: The Effects of Social Security Benefit Increases, 1952-1991. *American Economic Journal: Macroeconomics*, *8*(4), 1-42.

Stock, J. & Watson, M. (2002). Has the Business Cycle Changed and Why? *NBER Macroeconomics Annual*, *17*, 159-218.

Taylor, J. (1999). A Historical Analysis of Monetary Policy Rules. In *Monetary Policy Rules*, Chicago: University of Chicago Press for NBER, 319-40.

Wu, J. & Xia, F. (2016). Measuring the Macroeconomic Impact of Monetary Policy at the Zero Lower Bound. *Chicago Booth Research Paper* No. 13-77.

Zellner, A. (1986). On Assessing Prior Distributions and Bayesian Regression Analysis with G-prior Distributions. *Bayesian Inference and Decision Techniques: Essays in Honor of Bruno de Finetti*, (eds. P. K. Goel and A. Zellner), 233-243.