Spring 4-14-2013

Essays on Nonlinearities and Structural Breaks in the Relationships between Macroeconomic Variables

Irina B. Panovska
Washington University in St. Louis

WASHINGTON UNIVERSITY IN ST. LOUIS

Department of Economics

Dissertation Examination Committee:
Werner Ploberger, Chair
James Morley, Co-chair
Steven Fazzari
Michael McCracken
Raul Santaeulalia-Llopis
Itai Sened
Stephen Williamson

Essays on Nonlinearities and Structural Breaks in the Relationships between Macroeconomic Variables

by

Irina Blagoja Panovska

A dissertation presented to the
Graduate School of Arts and Sciences
of Washington University in
partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

May 2013

St. Louis, Missouri

Contents

List of Figures

List of Tables

ACKNOWLEDGEMENT

I would like to express my gratitude to my advisers, Werner Ploberger and James Morley, for their guidance, help, and support. I am grateful to many people in the Economics Department for their help. In particular, I would like to thank Steven Fazzari and Raul Santaeulalia-Llopis for comments and suggestions on earlier drafts of the chapters of the dissertation. I would also like to thank Luiggi Donayre, Yunjong Eo, and Tara Sinclair for their comments and suggestions on my work, and for the peer mentorship they have provided over the years. I gratefully acknowledge the financial support of the Mr. and Mrs. Spencer T. Olin fellowship, the Weidenbaum Center, and the Institute for New Economic Thinking.

I would like to thank my parents, Blagoja Panovski and Vesna Grivcheva, for encouraging my academic curiosity and for their love and support. I am also indebted to Herman and Ann Hill for their academic guidance, support, and friendship. Finally, I would like to thank Andrew Snell for his love and infinite patience.

For my parents and for Andrew

ABSTRACT OF THE DISSERTATION

Essays on Nonlinearities and Structural Breaks in the Relationships between Macroeconomic Variables

by

Irina Blagoja Panovska

Doctor of Philosophy in Economics

Washington University in St. Louis, 2013

Professor Werner Ploberger, Co-Chair

Professor James Morley, Co-Chair

The dissertation is set out in three chapters, focusing on the structural changes that led to jobless recoveries in following the past three recessions, the asymmetric effects of fiscal stimulus over the business cycle, and the asymmetric effects of fiscal cuts and stimulus respectively. In the first chapter, I examine the three leading theoretical explanations for the recent jobless recoveries using a correlated unobserved components model of aggregate data for output, sales, employment, and hours. The main finding is that employment now respond to demand shocks in a way that is consistent with just-in-time utilization of labor resources. The second and the third chapter focus on asymmetric responses to fiscal policy. In the second chapter, I investigate the effects of government spending on U.S. economic activity using a threshold version of a structural vector autoregressive model. The empirical findings support state-dependent effects of fiscal policy. In particular, the effects of a government spending shock on output are significantly larger and more persistent when the economy has a high degree of underutilized resources than when the economy is close to capacity. The third chapter examines whether there are sign and size asymmetries in the responses of output, output components, and employment to fiscal policy. When the economy is not constrained, a large fiscal stimulus is more effective at increasing employment and output, and cuts have larger effects than increases.

Introduction

The following three chapters focus on the structural changes that led to the jobless recoveries following the past three recessions, the asymmetric effects of fiscal stimulus over the business cycle, and the asymmetric effects of fiscal cuts and stimulus respectively. In each chapter, empirical time series methods are used both to extend and challenge the existing empirical literature and to evaluate the importance of different channels implied suggested by theoretical models for the cause of the structural breaks and nonlinearities.

In the first chapter, I examine the relative importance of the channels suggested by theory for the changes in the behavior of employment during each recovery. There are three leading explanations for the recent jobless recoveries: increasing importance of permanent shocks; long expansions creating greater need for restructuring; and structural changes that cause a shift towards adjusting labor inputs more at the intensive margin (hours) rather than the extensive margin (employment). This chapter considers these competing explanations using a correlated unobserved components model of aggregate data for output, sales, employment, and hours. The main finding is that there is little empirical support for the first two hypotheses; however, hours and employment now respond to demand shocks in a way that is more consistent with just-in-time utilization of labor resources.

The second and the third chapter focus on asymmetric responses to fiscal policy and are based on joint work with Steven Fazzari and James Morley (Fazzari, Morley, and Panovska, 2013a, and Fazzari, Morley, and Panovska, 2013b). In the chapters that are based on joint projects, I developed the empirical model to test the theoretical implications, collected the data,wrote and ran the code necessary to conduct estimation, and reported and interpreted the results under the supervision of my co-authors.

In the second chapter, I investigate the effects of government spending on U.S. economic activity using a threshold version of a structural vector autoregressive model, and examine if they vary over the business cycle. The empirical findings support state-dependent effects of fiscal policy. In particular, the effects of a government spending shock on output are significantly larger and more persistent when the economy has a high degree of underutilized resources than when the economy is close to capacity. This evidence is consistent with an underlying structure of the economy in which insufficient aggregate demand often constrains the level of economic activity.

In the third chapter, I build on the second chapter, and examine if there is evidence in favor of sign and size asymmetry within a state. In order to evaluate the degree of size and sign asymmetry, I use formal impulse response comparison. The main questions addressed in the second and third chapter are closely related, but distinct, especially when it comes to designing fiscal policy. The main question addressed in the second chapter is related to whether there is any evidence in favor of implementing a stimulus or austerity measures at all, and if the impact of these stimulus measures varies depending on the level of capacity utilization. If there is evidence in favor of state-dependence indicating that spending multipliers, on average, are larger during periods of slack, then one could make a case for using fiscal stimulus to increase output in times of recession or, in general, in times when the economy is below full capacity. The third chapter focuses on size and sign asymmetry conditional on the existence of state-dependent responses. If there is evidence in favor of size and sign asymmetry within a state, this can help us determine the optimal size and timing for a stimulus and the optimal time and size for an austerity measure. For example, if there is evidence that a larger stimulus is more effective, dollar-for-dollar, than a small stimulus in times of recession, then it would be optimal to implement a single large stimulus package rather sequence of smaller measures. Similarly, if there is evidence that austerity measures have much smaller negative impact on output during robust recoveries than during weak recoveries, then there would be an optimal time to implement austerity measures.

The main findings of the third chapter are that there is strong evidence in favor of sign and size asymmetry. When the economy is not constrained, a large fiscal stimulus is more effective at increasing employment and output. On the other hand, when the economy is close to capacity, a large stimulus has smaller effects on output than a smaller stimulus. Consumption exhibits the same asymmetric responses as output. A cut crowds in investment more than an expansion crowds out investment in the constrained state, but a large cut has smaller effects than a small cut. Investment is crowded in in the unconstrained state, but the response is small, and there is no significant evidence of sign or size asymmetry. When the economy is very close to capacity, fiscal cuts have larger effects on payroll employment than fiscal expansions. The effects on private employment are smaller in magnitude, but qualitatively similar to the effects of cuts and expansions on overall employment.

1. Chapter 1: What Explains the Recent Jobless Recoveries?

1.1. Introduction

The past three recoveries in the United States have been markedly different from most postwar recoveries prior to 1990 where payroll employment returned to its pre-recession level just a few months after the trough in output. The typical recovery from earlier recessions was characterized by fast job creation that quickly offset the job losses resulting from the recession. By contrast, employment growth has been sluggish or negative for months and even years after the NBER-determined trough in the past three recoveries.Because of this slow, delayed employment growth, many economists refer to the past three recoveries as "jobless". If there is a common explanation for this sluggish employment growth in recent recoveries, then it might be possible to undertake policies to mitigate the effects of prolonged joblessness.

While there is consensus in the literature regarding the stylized facts and the change in the dynamic behavior of employment, there is no consensus on the cause of jobless recoveries. In the current literature, there are three leading hypotheses: increasing importance of permanent movements relative to cyclical movements; organizational restructuring causing accumulation of inefficiencies during expansions and large restructuring during recessions; and innovations in labor demand leading to more flexible hiring practices and just-in-time use of labor resources. Supporting the first hypothesis, Groshen and Potter (2003) argue that most of the slow growth in employment can be attributed to structural changes that are related to permanent shifts across industries. Supporting the second hypothesis, Berger (2011) relates the jobless recoveries to the length of expansions preceding the recessions. During expansions, firms want to fill vacancies quickly and they therefore choose to hire employees that are not perfectly matched, thus causing employment to rise above its trend level, creating a significant overhang at the end of a long expansion. A jobless recovery is simply a return to the long-run trend level. The third hypothesis on just-in-time use of labor, first proposed by Schreft and Singh (2003), relates the changes in the labor market after the 1990s to the switch to just-in-time management of workers. Since firms now have access to more flexible labor inputs, it became cheaper to absorb the initial increase in demand during the first stages of a recovery through overtime hours and flexible labor inputs, rather than through increasing full-time employment. In addition, productivity-driven changes, compositional changes in labor supply, and particularly adverse demand shocks during recessions and recoveries are also frequently mentioned as possible explanations for the jobless recoveries.

The main goal of this chapter is to examine the extent to which changes in the relationship between employment, output, sales, and hours can explain the change in the behavior of employment. I construct an empirical model that nests the three leading theoretical explanations for jobless recoveries. The model provides strong evidence in favor of switching to just-in-time utilization of labor resources. A large part of the change in the behavior of employment since the mid 1980s can be explained by changes in the adjustment cost of employment relative to hours. Meanwhile, the changes in the behavior of employment and hours per employee do not arise from a change in the persistence of sales or other shocks to output.

A lot of the movements in employment can still be explained by cyclical fluctuations. In particular, recessions and the periods immediately following recessions are characterized by movements in output and employment that are driven by large transitory shocks. As discussed in detail below, in contrast with the first two hypotheses, the lackluster recoveries are not primarily caused by shallow and short recessions; instead, they are caused by a change in how firms respond to demand shocks. Counterfactual analysis also suggests that firms wait for signs of a robust recovery in sales before they hire new employees, and simply increase the number of hours or utilize more flexible labor resources in the meantime. This is consistent with the predictions of the just-in-time hypothesis. Results obtained using sectoral data are very similar to the results obtained using aggregate data, and also provide evidence in favor of a switch towards just-in-time management of workers.

The rest of this chapter is organized as follows. Section 1.2 provides some background about the jobless recoveries. It includes stylized facts about the dynamics of real macroeconomic aggregates and labor market variables and the theoretical motivation for the unobserved components model. The empirical model is presented in Section 1.3. The results obtained using aggregate data are presented in Section 1.4. The results of the counterfactual analysis and the variance decomposition of employment are also presented in Section 1.4. In Section 1.5, I investigate the causes of the changing behavior of employment by comparing the results obtained using aggregate data to the results obtained using disaggregated data, and I perform robustness checks. Section 1.6 concludes.

1.2. Stylized Facts

There is a general consensus in the jobless recovery literature that the changes in the behavior of employment coincided with the start of the Great Moderation. Figure 1 and Table 1 present some stylized facts about the behavior of payroll employment and of macroeconomic aggregates before 1984 and after 1984.

Figure 1 illustrates how different the last three recoveries have been from the typical postwar recovery prior to 1984. The graph shows the difference in the behavior of payroll employment for each of the past three recoveries and for the average of the recoveries prior to 1984 in the three years following the trough.[1] The jobless nature of the past three recoveries is quite clear from the graph. Following the past three recession, the return to the peak level of employment was slow and sluggish, and employment continued to decline after the NBER trough, as shown by the downward-sloping paths of employment.

Table 1 shows that there were changes not just in the behavior of employment, but also in the behavior of output, sales, and hours. These changes in the behavior of both real variables and labor market variables appear to effect the economy over the entire business cycle, and not only during recoveries, as evidenced by the first two rows of the

[1] The two NBER recessions in the early 1980s were treated as a single recession. Treating them as separate recessions does not affect the results significantly.

table. The past 28 years were a period of lower volatility and lower growth rates, both in recoveries and in normal times. The last two columns of the table illustrate the stark differences between the pre-jobless recoveries period and the period after the mid 1980s. Recoveries before 1984 were characterized by fast growth in output and sales, a slightly delayed but large growth in employment and hours, and a decline in hours per current employee that was mostly due to the rapid increase in the number of employees.[2] The third and the fourth row of Table 1 show that most of the increase in aggregate hours during the recovery phase in the pre-jobless recoveries period came from increases in payroll employment. By contrast, both aggregate hours and employment declined, but average hours per current employee increased in the past three recoveries. In addition to changes in the behavior of the individual series, there were also changes in the relationship between the series. The average correlation between employment growth and growth in hours per employee dropped from 0.72 to 0.28, and the correlation between output per hour and employment growth dropped from 0.18 to -0.27.

1.3. Theoretical Background

As discussed in the introduction, the theoretical literature about jobless recoveries can roughly be divided into three main categories. The first strand of the literature emphasizes the increasing importance of secular trends relative to cyclical movements in employment. The second strand attributes the jobless recoveries to the long expansions and shallow recessions, creating employment overhangs either at the beginning or at the end of the recessions. The third strand of the literature highlights the fact that labor markets became more flexible since the mid 1980s, and relates these changes to a model in which firms use a just-in-time approach when it comes to utilizing labor inputs. Alternative explanations outside the three main strands include productivity driven recoveries, a shift towards self-employment (postulating that a lot of the slow growth in employment

[2] Aggregate hours are measured using the updated version of Francis and Ramey's, 2009 BLS data. Using alternative measures of hours leads to very similar results.

following a recession occurs because the official employment numbers do not account for self-employment), and joblessness that can be attributed entirely to sluggish recoveries in output.

Theoretical models that focus on the importance of secular movements include the sectoral shift hypothesis and the job polarization hypothesis; both hypotheses focus on the fact that reallocation, either across sectors or across occupations, can create temporary inefficiencies leading to slow employment growth. The sectoral shift hypothesis suggests that jobless recoveries occur due to shifts from one industry to another that occur primarily during recessions. Groshen and Potter (2003) use the JOLTS database to examine permanent and temporary layoffs across industries during the 2001 recession and recovery, and during recessions and recoveries prior to 2001. They look at temporary versus permanent layoffs, interpreting temporary layoffs as the cyclical component of disemployment, and permanent layoffs as the trend component of disemployment. Most of the new jobs created after the 2001 recession were not due to rehiring workers who were temporarily laid off, but due to new hires. Based on this definition of trend and cycle in (dis)employment, they conclude that the slow job creation during the 2001 recovery was due to structural changes, and that cyclical movements in employment became less important relative to permanent movements since the mid-1980s. Groshen and Potter offer sectoral shifts as one of the most plausible explanations for these observed movements in industry-level employment.

Similarly, Jaimovich and Siu (2012), look at secular trends across occupations, and relate the jobless recoveries to the increasing job polarization and the shift towards non-routine based occupations in the United States. Jaimovich and Siu note that a lot of the losses in the middle of the occupational distribution occur during recessions. They attribute the jobless recoveries to this hollowing out of the middle of occupational distribution, which in their model is caused by an increase in the importance of large and negative permanent shocks. If jobless recoveries are primarily caused by trend movements, movements in aggregate employment during recessions and recoveries will be caused primarily by permanent shocks; furthermore, compared to the pre-jobless recoveries period, permanent movements should be more important for explaining the overall dynamics of employment relative to cyclical movements.

The second strand of the literature relates the jobless recoveries to changes in the persistence and depth of the business cycle, leading to employment overhang. Koenders and Rogerson (2005) and Berger (2011) relate the jobless recoveries to the long expansions preceding the recessions. During expansions, firms can afford to hire employees that are not perfectly matched, thus causing employment to rise above its trend level, creating a significant overhang at the end of a long expansion. In this framework, firms use recessions to get rid of the least productive employees, and delay hiring until the restructuring is completed. A jobless recovery is simply a return to the long-run trend level. Bachmann (2011) uses a DSGE model in which jobless recoveries also occur due to employment overhang, but in this case the overhang occurs at the end of a recession and is caused by short and shallow recessions. He builds a theoretical model in order to analyze the responses of employment to cyclical movements in output during the 1991 and 2001 recoveries. In this type of DSGE model, firms respond to permanent movements by adjusting on the extensive margin, and to transitory movements by adjusting on the intensive margin. In this framework the jobless recoveries during the 1991 and 2001 recovery occurred because the recessions appear to be caused by shocks that are less persistent than the shocks causing the previous recessions. Rather than laying off employees, firms responded by adjusting hours, creating a large employment overhang at the end of the recession. The slow job growth during the recovery phase is caused by this overhang at the end of the recession. It is important to note that the overhang in this model occurs because the 1991 and 2001 recessions were caused by movements that were more shallow and less persistent than the movements that caused the recessions before 1984. If the jobless recoveries are caused by employment overhangs, then there should be evidence of employment being above its long-run trend level either at the start or at the end of the recessions. Furthermore, if the change in employment is due to shallow and less persistent business cycles, there should be statistically significant evidence in favor of a change in cyclicality of output.

The just-in-time hypothesis emerged after the 2001 recession, and it postulates that the

Great Moderation was linked to a switch to just-in-time in inventory management and use of labor inputs. This hypothesis predicts that firms will adjust along the intensive margin because the relative cost of adjusting along the extensive margin has increased, and the relative cost of adjusting along the intensive margin has decreased. Extensive summary of studies that focus on the 2001 recession is provided in Schreft and Singh (2003). In a follow-up paper Hodgson et al. (2005) examine the behavior of hours, overtime, and temporary employment, and conclude that the slow growth of employment during the 2001 recovery can be explained by switching to just-in-time employment. Since firms have access to flexible labor inputs like temporary employees, and are able to increase both overtime hours and aggregate hours much more cheaply than they can increase employment, they can postpone hiring permanent employees until demand is robust enough.

Motivated by the changes in the sample statistics, and by the competing theories, I build an unobserved components model in order to formally examine if the changes in the relationship between real economic activity and employment are statistically and economically significant, and if they are consistent with the theoretical models used to explain the jobless recoveries.

In principle, one could build and solve a theoretical DSGE model that has adjustments costs along both margins and that incorporates both permanent and cyclical movements in output. Allowing for a break in the relative adjustment costs and a break in the persistence of the cyclical movements would thus nest all three main hypotheses.[3] Unfortunately, as discussed both in Bachmann (2011) and Ramey and Vine (2006) full structural analysis is quite challenging for models of this type, even for the simple model that has constant adjustment costs along one margin only, and does not allow any breaks in the parameters. Instead, the empirical model used in this chapter nests the three competing hypotheses, and the reduced-form parameters can be related to the parameters from the

[3]The just-in-time hypothesis corresponds to changes in the adjustment parameters. The structural reorganization hypothesis corresponds to lower persistence and amplitude of cyclical movements, especially during recessions, and to employment rising significantly above trend during expansion. The sectoral shift hypothesis and the job polarization hypothesis corresponds to increasing importance of permanent movements and adjustments to permanent movements, in particular during recessions.

structural models.

1.4. Model

In order to examine if there are changes in the relationship between output, sales, and labor market variables before and after 1984, I adapt the approach used by Morley and Singh (2012), and consider an unobserved components (UC) model where the innovations to the permanent component and the transitory innovations are allowed to be correlated, but the underlying structural shocks driving the model are orthogonal. An unobserved components model of this type is particularly convenient for analyzing the behavior of employment, not only because it nests the three competing theories, but also because it does not entail making the restrictive assumption that all permanent layoffs are due to structural changes and are made with no intention of ever rehiring workers to fill that position again.

The empirical model is given by the following equations

$$y_t = \tau_t + c_{yt} \tag{1}$$

$$s_t = \tau_t + c_{st} \tag{2}$$

$$e_t = \zeta_t + c_{et} \tag{3}$$

$$h_t = \mu_h + c_{ht} \tag{4}$$

where y_t is a measure of real output, s_t is a measure of real final sales, e_t is a measure of employment, and h_t is a measure of hours per employee. Output and sales have the same stochastic trend

$$\tau_t = \mu_1 + \tau_{t-1} + \eta_t, \tag{5}$$

reflecting the fact that output and sales are cointegrated, as discussed in detail in the Appendix. When looking at y and s in levels, the series $y - s$ is equal to change in inventories (by definition). Since the model here uses log levels, the difference d can be interpreted as a close proxy for the change in inventories.

The stochastic trend in employment is the sum of two components: τ_t and κ_t, where

$$\zeta_t = \tau_t + \kappa_t \tag{6}$$

and

$$\kappa_t = \mu_2 + \kappa_{t-1} + \nu_t. \tag{7}$$

The cyclical components are assumed to be stationary, and their dynamics can be described by the following equations:

$$\Phi_1(L)(y_t - \tau_t) = \lambda_{y\tau}\eta_t + \lambda_{ys}\epsilon_s + \epsilon_y. \tag{8}$$

$$\Phi_2(L)(s_t - \tau_t) = \lambda_{s\tau}\eta_t + \epsilon_s \tag{9}$$

$$\Phi_3(L)(e_t - \tau_t - \kappa_t) = \lambda_{e\tau}\eta_t + \lambda_{ey}\epsilon_y + \lambda_{es}\epsilon_s + \epsilon_e \tag{10}$$

$$\Phi_4(L)(h_t - \mu_h) = \lambda_{h\tau}\eta_t + \lambda_{hy}\epsilon_y + \lambda_{hs}\epsilon_s + \lambda_{he}\epsilon_e + \epsilon_h \tag{11}$$

In this framework, firms can respond to output shocks by adjusting inventories, employment, and by adjusting hours. The first stochastic component, τ_t, can be interpreted as the productivity trend that affects long run output and therefore it affects long run employment. The second trend component reflects demographic, preference, and exogenous shocks to the labor share that do not affect the cyclical components of output, sales, hours, or employment directly. This specification is motivated by the results of Francis and Ramey (2009) and Kahn and Rich (2009), who show that low-frequency movements in hours and employment can be explained by two separate factors- productivity and demand shocks and demographic or taste shifts that are not perfectly correlated with productivity shocks. The shock ν can also be interpreted as a preference shift as in Rios-Rull and Santaeulalia-Llopis (2010). These shocks are allowed to be correlated with shocks to the productivity trend, but they do not affect the cyclical components of any of the series directly within a period. Not restricting the correlation between the two trends to be zero allows demographic changes like aging population or changes in the labor force to affect output and productivity in the long run, but they have negligible impact on the cyclical components within a quarter.[4]

Equations (4) and (11) describe the behavior of hours per employee. The hours per employee series is stationary across subsamples, which can be easily confirmed by standard stationarity or unit root tests. In section 1.5, I allow for a stochastic trend in the hours per employee series (which also accounts for a possible drift in the mean), but the results are virtually identical to the results obtained assuming that hours per employee are stationary.

The shocks $[\eta\ \nu\ \epsilon_y\ \epsilon_s\ \epsilon_e\ \epsilon_y]$ are the structural shocks that drive the model, and the impact coefficients λ_{ij} describe the within-period response to those shocks. The autoregressive coefficients are directly related to the persistence of the transitory movements in the series. The relative importance of permanent movements for the dynamics of employment is captured by the ratio of the volatilities of the permanent shocks, σ_η and σ_v, relative to the variances of the cyclical shocks, and by the percentage of the movements in employment that are explained by each type of shock.

In order to identify the impact coefficients, I assume that firms take sales as exogenous

[4]The non-zero correlation is meant to capture two facts. First, permanent shocks that increase productivity mean that less employees are needed to produce a unit of output. Second, productivity shocks that lead to higher income may lead to preference shifts or to simply enjoying more leisure time, if income is sufficiently high. Based on these interpretations, one would expect the correlation between the two permanent shocks to be negative. The model does not restrict the correlation to be negative a priori; rather, it is estimated from the data.

within a period, and the only shocks that affect transitory sales are permanent output shocks and transitory sales shocks. This identification scheme is standard in the literature on inventory management, and it also closely mirrors the usual VECM timing assumption that production is set based on expected sales. The transitory component of output is affected by the permanent shocks to output, the transitory sales shock, and an independent structural ϵ_y shock that can be interpreted as an inventory mistake shock that does not depend on demand. The transitory component of employment is affected by permanent output shocks and by transitory shocks to sales, inventories, and by idiosyncratic employment shocks. Within a period, hours per employee are affected by permanent shocks to output, by transitory sales, inventory, employment shocks, and by their own transitory shocks. The inclusion of inventory shocks is motivated by Galeotti et al. (2005), who use a theoretical model of employment and hours that integrates inventory and labor decisions both at the intensive and at the extensive margin for the pre-jobless recoveries period. They show that ignoring inventory decisions can lead to distorted inference about the cost of adjusting employment.

The impact coefficients λ_{ij} capture the response of the observed variables to the uncorrelated structural shocks. These coefficients play two important roles in this framework. They provide a way to orthogonalize the shocks that does not depend on restricting the responses of the cyclical components beyond the exogeneity restrictions implied by equations 8 through 11. The orthogonalization scheme used here nests the orthogonalization scheme used in the VAR models that study the response of hours to technology shocks (see, for example, Gali, 1999, or Rios-Rull et al., 2011), but it allows for the possibility that shocks other than productivity can have non-zero effects in the long run. By using impact coefficients I allow the permanent and the transitory movements to be correlated, and allow for slow cyclical adjustment to permanent movements, while keeping the structural shocks uncorrelated. Second, they are directly related to the sensitivity of the transitory components to structural shocks. In particular, the impact coefficients on employment and hours per employee can be interpreted as adjustment coefficients that measure the sensitivity of employment and hours to permanent and transitory shocks.

The polynomials $\Phi(L)$ capture the autoregressive dynamics of the transitory components. The autoregressive polynomials $\Phi_i(L) = 1 - \phi_i L - .. - \phi_{i,p_i} L^{p_i}$ are assumed to have roots strictly outside the unit circle for $i = y, s, emp,h$. As discussed in Morley et al. (2003), an unobserved components model with correlated shocks is identified given sufficiently rich dynamics. To ensure identification in this model p_i has to be greater than 1 for each $i = 1, 2, 3, 4$. Following most of the UC literature, I set $p_i = 2$ for $i = 1, 2, 3, 4$.The structural shocks are assumed to be Gaussian. The model given by equations (1)-(11) is a restricted four-variate unobserved components similar to the the unrestricted models used by Sinclair (2009) and Basistha (2009).[5]

It is well-known that there was a general decline in volatility in real macroeconomic variables around 1984, and this decline in volatility has been well documented in the literature (see, for example Kim and Nelson (1999), or McConnell and Perez-Quiros (2000)), and there is a general consensus that the break in variance occurred between 1983 and 1985. There is also a general consensus in the jobless recovery literature that the change in the behavior of employment also occurred around the mid 1980s. Since the timing of the start of the Great Moderation and the timing of the start of the jobless recoveries period is not the primary focus of this chapter, I treat the break as exogenous. Engemann and Owyang (2010) estimate a univariate smooth transition model for employment, both aggregate and sectoral, where they allow the break date in the speed of adjustment of employment to be endogenous. Their estimated break dates also coincide with the start of the Great Moderation, implying that treating the break as exogenous and occurring around 1984 is a plausible assumption. I estimate the baseline model separately for the pre-Great Moderation period (1948q1 1983q4) and for the post-Great Moderation Period (1984Q1-2012q1), allowing the autoregressive coefficients, the volatilities of the structural shocks, and the impact coefficients to change across subsamples. By allowing both the impact coefficients and the variances to change, I allow the correlation between the shocks to the cyclical components and the trend components to differ across subsamples.

[5]In particular, it is equivalent to a model where the correlation between employment trend and the cyclical components of all variables can be expressed as functions of the correlations between the output trend and the cyclical components and the correlations between the cyclical components.

In this setup, changes in employment can occur through any of the three channels implied by the theoretical models. To distinguish between the competing theories, I focus on three sets of key questions:

1. How much of the relative fluctuations in employment can be explained by permanent shocks, and how much can be explained by transitory shocks? Do cyclical movements explain movements during recessions and recoveries after 1984? If there jobless recoveries are caused by permanent shocks, then there should be significant evidence of increasing importance of permanent movements relative to cyclical movements, both across the business cycle, and in particular during recessions and recoveries.

2. Are there significant changes in the persistence of transitory movements in sales and output? Is there significant evidence of overhang at the late stages of an expansion? If there is evidence that employment is significantly above its long-run trend level at the end of the expansions, this would provide support in favor of the organizational restructuring theory. If employment is above its long run trend level at the end of the recessions, and there is evidence that the persistence of output and sales declined, this would provide evidence in favor of Bachmann's (2011) model.

3. Can the different responses of employment to output shocks be explained by the change in the sensitivity of employment and hours to overall economic activity? If employment is more sensitive to sales shocks, and initial adjustments occur on the intensive margin, this would provide evidence in favor of a switch towards just-in-time production.

My approach is somewhat similar in spirit to Gomme's (2005) probit approach, looking at the probability of finding a job and the probability of separation; as well as, Fabergman's (2008) SVAR approach using data on output, job creation, and job destruction, and defining a reallocative and aggregate productivity shock. In both papers, they find that the Great Moderation coincided with a change in the dynamics of the labor market. In contrast with Gomme's and Fabergman's approach, who each examine only if there is evidence of a change in dynamics in the labor market at the start of the Great Moderation, I formally disentangle the channels implied by the theoretical models, and analyze how much changes in each of the channels contributed to the changes in employment.

1.5. Empirical Results

1.5.1. Inference and Estimates

The data series used are quarterly U.S. real GDP and real final sales from BEA, total non-farm payroll employment (converted to quarterly frequency using arithmetic averages), and Francis and Ramey's updated measure of aggregate hours divided by the number of employees. The first subsample covers the period 1948q1-1983q4, and the second subsample covers the period 1984q1-2012q1. All series are converted to hundred times the natural logarithm of the raw data series.

The linear UC model that restricts $\mu_{it} = 0$ has 28 parameters for each subsample, and it is identical to a reduced-form vector error correction model that has both a VAR and a vector MA component. Given the fact that the model is parameter-rich, estimation is conducted using Bayesian methods. In particular, I use a multi-block random walk chain with a Student-t proposal version of the Metropolis-Hastings (MH) algorithm. There are two main advantages of using Bayesian estimation over maximum likelihood estimation in this framework. First, the Bayesian approach allows me to directly capture the uncertainty about the parameters when performing counterfactual analysis. Second, previous research that uses UC models to decompose trend and cycle movements in macroeconomic variables frequently finds that the estimated parameters for many US macroeconomic series are close to the boundary of the parameter space, which can lead to the well-known pile-up problem. For example, Sinclair (2009), uses a bivariate model for unemployment and output and finds correlations that are very close to -1, and Basistha (2009) uses a four-variate model for productivity, inflation, output per capita, and hours per capita, and finds that some of the trend variances in the labor market variables are close to 0. Using Bayesian methods allows me to find an interior mode of the posterior,even with non-informative priors, thus circumventing the pile-up problem. The parameters were estimated by casting the model into state-space form, and updating each parameter at each MH draw. The state-space representation for the UC model and a detailed description of the sampler and the priors are provided in the Appendix.

If the changes in employment behavior occurred due to increasing importance of trend movements, then the ratio of the trend volatility to volatility of the cycle in employment should increase, as should the implied correlation between the total employment trend and cycle. If there is significant evidence in favor of restructuring because of long expansions,there should be a significant employment overhang at the end of the expansions. On the other hand, if jobless recoveries can be explained by a switch to just-in-time employment practices or due to similar structural changes in the labor market, one would expect both the ratio $\lambda_{es}/\lambda_{ey}$ and the impact coefficients on hours per capita to increase significantly. Looking at the point estimates for the structural parameters allows me to directly compare if the changes in the dynamics of employment were caused by increasing importance of trend movements, or if there was a significant change in the persistence of the cyclical movements.

Table 2 presents the median estimates for the persistence of the cycles, measured by the sum of the autoregressive coefficients, and the estimates for the key volatility and impact coefficients.The autoregressive coefficients ϕ_{i1}, ϕ_{i2} (and thus the persistence of the cyclical components) are very similar across subsamples for all series. There is no evidence of a significant change in persistence of the output and sales cycle. The changes in the behavior of the responses of employment and hours to structural shocks come from the changes in the impact coefficients, rather than from changes in persistence.

The estimated volatilities of both the permanent shocks to output, η, and the cyclical components of output and sales, σ_s and σ_y, are lower for the second subsample, which is not surprising, because the break is chosen to coincide with the start of the Great Moderation. A particularly important result in this framework is not the decline in volatility, but the decline in the ratio of the volatility of the output relative to the volatility of sales. Before the start of the Great Moderation, the median estimate for the ratio of the volatility of the cyclical structural shocks of output to the volatility of the cyclical structural shocks to sales was 1.45. After 1984, the ratio drops to 0.98. Together with the change in the impact coefficients on sales, the change in this ratio explains a lot of the observed behavior in employment after 1984, as shown in the next two subsections. The impact coefficients for the permanent output shock do not change drastically after the break. It is, however, important to note that the impact coefficients for the permanent shock η on employment and hours are negative. The median estimated impact coefficient $\lambda_{e\eta}$ is -0.14 for the first subsample and -0.22 for the second subsample, and both of those parameters are significant. The median estimated impact coefficient $\lambda_{h\eta}$ is -0.18 for the first subsample, and -0.25 for the second subsample, and again the credibility intervals for both estimates do not cover zero, even when using conservative 95% credibility intervals. Contrary to predictions from RBC models, permanent shocks to productivity have a transitory negative impact on hours and employment, implying that hours and employment do not adjust to steady-state immediately following a permanent shock. Even though I use different measure of hours and productivity, the implied correlations between the shocks η and u_h are also close to the parameter estimates obtained by Basistha (2009), who looks at the correlations between productivity, hours per capita, and inflation. The negative impact coefficients are at odds with RBC predictions, but they are consistent with Gali's (1999) New-Keynesian DSGE model where productivity shocks have temporary negative effect on hours. The impact coefficients $\lambda_{e\eta}$ and $\lambda_{h\eta}$ and the correlation between the productivity shocks and the demographic shocks do not change significantly across subsamples, implying that there is no evidence that permanent technology shocks have become more important in explaining movements in employment after 1984.[6]

The largest change occurred in the impact coefficients λ_{ey}, λ_{es}, λ_{hy}, λ_{hs}, which capture

[6]The approach used here assumes that the coefficients are stable within a subsample, but it is interesting to note that the implied median estimates of the correlation between sales growth and inventory growth $\Delta i_t = y_t - s_t$ are broadly consistent with the median estimates from the TVP model used by Benati and Lubik (2012).

the sensitivity of the employment and hours cycles to transitory inventory/ non sales shocks, ϵ_y, and sales shocks, ϵ_s. The median estimate for the impact coefficient of sales on employment, λ_{es} , was 0.35 before 1984, and 0.87 after 1984. Before the Great Moderation, a positive sales shock equal to 1 increases log employment by 0.3%, and after 1984, a positive cyclical sales shocks increases employment by almost 0.9%. If one only looks at the impact coefficients, these results imply that within a quarter, the effect of a positive structural inventory shock equal to 1 on employment has gone down by approximately 50% (from 0.11 down to 0.06), and the effects of a positive structural shock to sales on employment has gone up by a factor of almost 3. The change in the impact coefficients of sales and non-sales output shocks on hours per employee is just as striking. Before the Great Moderation, hours per employee responded positively to inventory shocks within a quarter, but the response was only about 60% of the response of hours to inventory shocks after 1984 (0.2 vs 0.33). In addition to the large changes in the impact coefficients of sales and other output shocks on employment, there is a large change in the impact coefficient of employment of hours. The impact coefficient changes from 0.68 to -0.13, and the change is statistically significant. This switch from positive to negative impact coefficient is also consistent with a switch towards just-in-time employment. After 1984, a negative shock to employment is associated with an increase in hours per employee, whereas employment and hours per employee used to be positively correlated before 1984.

The changes in the relative ratios of the variances and the impact coefficients are consistent with the switch to just-in-time employment hypothesis that postulates that if firms are using more flexible labor inputs, the responses of hours and aggregate hours would increase much more after the break, and that the response of employment to sales should increase after the break. In order to further disentangle how much of the movement in employment and hours can be attributed to the permanent shocks, and how the change in the impact coefficients affected the path of the employment, I perform counterfactual simulations that allow me to address those questions directly.

1.5.2. Cyclical Shocks and Employment Overhang

Both the sectoral shift hypothesis and the job polarization hypothesis imply that at the aggregate level, permanent movements became relatively more important than cyclical movements, both on average, and during recessions and recoveries. In the unobserved components framework, this means that the variance of the overall employment trend increased relative to the variance of the overall employment cycle, that the correlation between the trend and the cycle component increased in absolute value, or both. However, the empirical results do not show a significant decline in the relative volatility. While there is an overall decline in volatility, there is no evidence that the overall employment trend became more important relative to the employment cycle. The variance of the overall trend component in employment decreased from 0.62 to 0.35, and the total variance of the cyclical component decreased from 0.94 to 0.51, leaving the ratio virtually unchanged (0.66 vs 0.68). Furthermore, the overall trend and the cycle are significantly negatively correlated in both subsamples. The correlation is -0.88 with standard deviation of 0.10 for the first subsample, and -0.91 with standard deviation of 0.7 for the second subsample. The large negative correlations indicate slow adjustment of employment to permanent shocks, but there is no evidence that the correlation, and thus the speed of adjustment has changed.

In order to examine if cyclical movements play a role during recessions and recoveries after 1984, I set up a counterfactual experiment that is similar in spirit to the historical decompositions by Gali and Rabanal (2005), Santaeulalia-Llopis (2012), and Gali, Smets, and Wouters (2012). The goal of this experiment is to examine how much of the movement in employment over time can be attributed to each structural shock. The set up is straightforward: given the observed shocks, I simulate the path of employment for the case when only permanent shocks affect the system, and for the case when only one of the structural cyclical shocks affects the system. To account for parameter uncertainty, at each iteration of the MH sampler after the burn-in I feed the observed shocks for the given parameter draw through the system using the first two quarters of 1984 as the initial values. Figure 2 plots the posterior mode for the simulated series.

Since the shocks η and ν are correlated, they are fed through the system jointly. This decomposition of the path of employment is different from the standard trend-cycle decomposition in UC models because the movements due to the permanent shocks account for movements in the stochastic trend and for movements in the cyclical component that are caused by permanent shocks. The left panel of Figure 2 plots the observed path of employment (solid blue), the median path of employment caused by permanent shocks and adjustments to permanent shocks (dashed red line), and the median path for movements due to purely cyclical shocks (patterned green, scale on the right axis). The right panel plots the decomposition of the cyclical component: solid black line is fluctuations due to sales shocks, and the dashed green line is fluctuations due to other output shocks.

As illustrated in Figure 2, there is no evidence of employment overhang at the end of the expansions; employment is very close to its trend level, with the exception of recessions and recoveries themselves. The slow adjustment of cyclical movements to permanent shocks is evidenced by the fact that employment is briefly above its trend level during the initial period after a negative shock hits the economy. There is no cyclical overhang at the end of the expansions, nor at the end of the recessions. Furthermore, movements that are caused by cyclical fluctuations are quite volatile, in particular during recessions and recoveries. The movements in employment caused by purely transitory shocks match up the NBER recessions and the period following the recession quite well. The recessions and the jobless recovery periods cannot be solely attributed to permanent shocks or adjustments to those shocks only.

The right panel of Figure 2 decomposes the part of employment driven by purely transitory shocks by type of shock: movements caused by sales shocks and movement caused by other output shocks (inventories). The overall median contribution of sales shocks to the variance caused by cyclical movements in employment increased from 24.2% before 1984 to 34.8% after 1984, while the contribution of other output shocks fell from 53% to 35.4%. This is particularly evident during the recession and recovery phase, as sales shocks drive a lot of the cyclical movements in employment, which is consistent with the just-in-time theory.

The decomposition of the hours per employee series is also consistent with this assumption. Figure 3 shows the decomposed path of hours per employee: the patterned blue line is the observed path, the red dashed line is fluctuations due to permanent shocks, the solid black line is fluctuations due to sales shocks, and the patterned green line is fluctuations due to other output shocks. A lot of the fluctuations do come from responses to the permanent shocks, but the cyclical shocks play a large role during recessions and the periods immediately following the recessions.

As illustrated in Figures 2 and 3, a large portion of the volatility in employment and hours per employee cannot be explained solely by permanent shocks and adjustments to those permanent shocks, as suggested by the sectoral shift hypothesis. Permanent shocks and adjustments to permanent shocks drive most of the dynamics during normal periods, but their relative importance has not increased after 1984, and cyclical shocks still play a significant role in explaining the dynamics of hours and employment over the business cycle, especially during recessions and the periods immediately following recessions. Furthermore, there is no evidence that employment is significantly above trend at the end of the expansions. These decompositions are consistent with the just-in-time employment scenario.

1.5.3. The Role of the Impact Coefficients and Estimated Job Losses

When disentangling the role of the different channels implied by the theoretical models, one of the key issues is to what extent the change in the impact coefficients (i.e in the sensitivity of employment and hours to real macroeconomic shocks) affected the path of employment and hours. In order to shed light on this question and to isolate the effects of each impact coefficient, I perform counterfactual analysis where I change only one of the impact coefficients to its pre-Great Moderation value, and leave all other parameters at their post-1984 values. The approach used here is similar in spirit to the approach used by Kim, Morley and Piger (2008), who use counterfactual analysis to analyze the sources of the Great Moderation, and to the approach used by Morley and Singh (2012), who use counterfactual analysis to study the role of inventory mistakes for the reduction in the volatility of output relative to the volatility of sales. The counterfactual analysis is performed as follows:

- the pre-1984 results were obtained first
- for a given parameter draw from the MH sampler for the second subsample, the impact coefficient of interest was substituted with the mode for the pre-1984 value
- the observed residuals for the draw (obtained using the MH parameter draw) were orthogonalized and fed through the system using the impact coefficient from the first subsample

Figures 4 and 5 in the Appendix plot the observed series for employment and the simulated series, and the observed series for hours and the simulated series. The observed path of the series is plotted in blue. The green "cloud" is the 90% CI for the counterfactual path when the sales impact coefficients are set to their pre-GM values, and the red "cloud" is the 90% CI for the counterfactual path when the non-sales output impact coefficients are set to their pre-GM values. These experiments illustrate the importance of the change in the impact coefficients, and they also illustrate that the slow growth in employment cannot be solely attributed to particularly adverse output shocks, as suggest by Gali, Smets, and Wouters (2012). If the impact coefficients on employment had not changed, employment would have exhibited patterns similar to those observed prior to 1984- a rapid decline at the start of the recession caused by a negative shock, and a rapid bounce back. However, because the impact coefficients did change, the observed series exhibits much smaller responses to non-sales shocks.[7] As mentioned above, this is particularly evident when large shocks hit the economy, which is exactly the case of recessions and the positive shocks that get the economy out of a recession. Hours per employee are much

[7]Even though the AR coefficients did not change significantly, and the cyclical component still plays a very important role, if one were to look only at the growth rates of employment, the result presented here would lead to a "loss of cyclicality" in the employment growth series, which is in line with the findings of Engemann and Owyang.

more procyclical after the break. When holding the impact parameters at their pre-break levels, the hours per employee cycle is much less volatile, and it takes longer for hours to return to their pre-recession levels. Again, these results are in line with the theory that firms have switched to utilizing more flexible labor inputs.

Of course, a key question for policy analysis is how many jobs were lost due to this shift towards adjusting on the intensive margin. The cumulative job losses due to the change in the impact coefficients (obtained by integrating the difference between the counterfactual path and the observed path for the three years following the trough) are quite large and significant. Three years after the 1990-1991 Recession, between 160,000-190,000 jobs were lost due to this switch. The estimate for the 2001 recession is between 0.9 and 1.31 million. The model implies that the Great Recession was caused by a mix of a a large adverse permanent shock and and negative transitory shocks. The estimates for the job losses caused by the change in the responses to cyclical shocks are slightly smaller than for the 2001 recession, but still large and significant: between 0.4 and 0.7 million. The estimates for the first two recoveries are comparable to the estimates obtained from the univariate model used by Engemann and Owyang (2010).

The counterfactual analysis illustrates that cyclical fluctuations still play an important role in explaining the movements in employment and hours per employee. Sales shocks play a much more important role in explaining employment fluctuations than other output shocks. Combined with the decline of the volatility of output relative to the volatility of sales, the results support a switch to just-in-time utilization of labor resources.

1.5.4. Robustness

Sectoral Results

The results presented in the previous section show that cyclical fluctuations still play an important role and they support the switch to just-in-time production at the aggregate level. However, in order to fully distinguish between the competing theories, one would also need to look at disaggregate industry-level data. The sectoral shift hypothesis states that cyclical fluctuations play a declining role in explaining movements in employment, and that at the sectoral level, almost all of the movements in employment for the post-break period can be explained by permanent shocks. On the other hand, if jobless recoveries can be explained by organizational restructuring, the changes should be more drastic in industries that have been expanding at fast rates before a recessionary negative shock hits the economy.

In order to test the just-in-time hypothesis at the disaggregate level, I look at the sectors that are emphasized by Groshen and Potter (2003) when explaining the movement of jobs across industries: manufacturing (durables and nondurables), services, and finance, real estate, and insurance. If the sectoral shift hypothesis holds, permanent shocks and adjustments to permanent shocks should explain most of the movements in employment in all of the sectors, and this change in the importance of permanent movements should be more pronounced when using disaggregated data. If the structural reorganization holds, then there should be significant overhang at the end of the expansion phase in the finance sector. On the other hand, if changes in the dynamics of employment occurred due to an overall switch to just-in-time labor management practices, the results should be robust to disaggregation, and the change in the relative size of the impact coefficients should be evident across industries. In order to distinguish between the competing theories, I estimate an unobserved components model for each sector for the period before 1984 and after 1984. The nominal data series for output and sales are obtained from BEA-NIPA and converted to real terms using the industry price deflators, also from BEA. The employment and hours series were obtained from BLS[8]. The sample size covered the period 1964q1-2011q2, due to the hours series being available only after 1964. Again, the model was estimated using hundred times the log of all the series.

[8]Using Francis and Ramey's series for disaggregated employment by industry leads to similar results. Since the sectoral data used here is aggregated, most firms that were classified as belonging to these industries under SIC did not change their classification under NAICS. To ensure that the results are robust, I also estimated the results by rescaling the nominal data to match at the point of conversion, and separately before and after the conversion. This did not affect the results significantly.

Manufacturing Sector

The manufacturing sector is particularly interesting because it has been used to show evidence both in favor of the sectoral shift hypothesis (Groshen and Potter, 2003), and in favor of changing labor demand and moving from adjusting labor on extensive margins to adjusting labor on extensive margins (see, for example, Ramey and Vine, 2006, and Hetrick, 2000), which is consistent with switching to just-in-time production. Furthermore, Engemann and Owyang find that most of the change in cyclicality in employment is due to large changes in the manufacturing sector.

The parameter estimates for the sectoral results are given in Tables 4 and 5 in the Appendix. The results for the manufacturing sector are similar to those obtained aggregate data, and four key results stand out. First, most of the dynamics in both subsamples is driven by permanent shocks and adjustments to permanent shocks, but the relative importance of the permanent shocks has not increased over time. Second, there is a reduction in volatility in all of the cyclical components, and a reduction in the volatility of output relative to the volatility of sales. Third, the increase in the impact coefficients of sales relative to the impact coefficients on inventories is similar, but larger than the increases for aggregate data. Fourth, the estimated impact coefficient of employment on hours per employee is positive in the first subsample, but negative in the second subsample. There is no significant reduction or increase in the persistence of the cycles. All of these results support the just-in-time hypothesis. This is consistent with the assumption that due to using more flexible labor inputs firms can adjust hours much more easily than they can adjust payroll employment, they are able to set output closer to sales, and any cyclical shocks affect hours much more than they affect employment. The results are very similar for when the series is further disaggregated into manufacturing of durables and nondurables.

Rapidly Growing Sectors with Long Expansions

Compared to other industries, employment in the services sector, in particular in the financial services sector, exhibited very rapid growth after 1984. The average growth rate for the services sector was 0.756% per quarter, and 0.423% per quarter for the financial sector, with minimal downward swings during the NBER expansion phase. If the organizational restructuring theory holds, then recoveries in this sector should be particularly lagged and drawn out. However, Figure 7 in the appendix illustrates that the dynamics of employment during the past three recoveries was comparable to the dynamics of employment in the manufacturing sector, and to the behavior of aggregate employment. The estimates from the formal econometric model, shown in Tables 6 and 7, confirm this.[9] The impact coefficients of non-sales output shocks on hours per employee increase by more than 150% in the second subsample, the impact coefficient of other output shocks on employment decreases by 70%, and the impact coefficient of employment on hours become negative in the second subsample. The increase in the impact coefficients of other output shocks on hours per employee, coupled with the change in the sign of the impact coefficient of employment on hours per employee indicates that in the second subsample, output increases on impact mostly through increasing hours per employee, and not by increasing the number of employees, which is consistent with just-in-time employment practices. It is important to note the change in the impact coefficients is not as drastic as the change in the manufacturing sector. The variance of permanent shocks increases relative to the variance of cyclical shocks, but this increase is not statistically significant. This points that movements in the services sector may include sectoral shifts or organizational restructuring, but that the cyclical component also shows strong evidence in favor of switching to just-in-time production.

Other Robustness Issues

The Effects of the Great Recession

In the past two years, there has been an ongoing debate if the Great Recession was the end of the Great Moderation. If the start of the Great Recession marked the return of a more volatile economic climate. To ensure that the results are not distorted by the

[9] The model is similar to the model given by equations (1)-(11), only adjusted to reflect the fact that in these sectors, when using data at quarterly frequency, output and sales can be treated as being roughly equal. The details of the model and the state-space representation are given in the Appendix.

presence of large outliers and change in dynamics during the past three years, I reestimate the model given by equations (1)-(11), but the second subsample is shorter and ends at 2007Q3. The last column of Table 3 gives the estimates and the standard deviations for the key volatility and impact parameters for the shorter subsample. Not surprisingly, the standard deviations of almost all the estimates are slightly larger when including the Great Recession, but none of the parameter estimates change significantly, implying that including the Great Recession data does not distort the results. This is in line with the results obtained by Stock and Watson (2012), who find that the Great Recession was driven by the same dynamics as the the previous two recessions, but the shocks were larger. The Great Recession was larger and deeper than the previous two recessions, but the recovery and the responses of employment mimic those observed during the previous two recoveries.

Alternative Measures of Hours and Stochastic Trend in Hours

The literature that studies the relationship between output shocks and hours usually focuses on aggregate hours or on hours per capita, not on hours per employee. However, there is much contention about the proper way to model hours and hours per capita, and about the presence of a stochastic or deterministic trend in the series (for a detailed literature review, see, for example, Basistha, 2009, Francis and Ramey, 2009, or Santaulalia-Llopis, 2012). Francis and Ramey's main criticism of using aggregate hours or aggregate hours per capita and not accounting for a demographic trend will tend to overestimate the effects of permanent technology shocks on transitory hours and bias them upwards. Santaeulalia-Llopis, on the other hand, finds that when he allows for preference shocks that are correlated with output shocks, hours per capita respond positively to technology shocks. The UC model used here is immune to those identification issues, because indirectly nests both Francis and Ramey's model and Santaeulalia-Llopis' model by allowing the trend in employment to have two separate components that are not perfectly correlated (output shocks and "other" shocks). The trend κ_t is allowed to mix the demographic and the preference trends, without making it necessary to identify them separately, and this trend does not enter the hours per employee series directly. Note that aggregate hours are decomposed as total employment multiplied by hours per employee, the "other" permanent shocks does not affect this series directly within a period. Furthermore, focusing on hours per employee rather than hours per capita has the advantage that the movements in hours per employee have a direct interpretation based on the theoretical models. As illustrated in the theoretical motivation in Section 1.2, if firms want to adjust total hours (which is the target variable in all standard theoretical models), they can do so by changing the number of employees, changing the hours per employee ("intensity"), or through a combination of both approaches.

Allowing for the possibility that hours per employee have a time-varying drift or low-frequency movements that might distort the results[10], the results do not change significantly. Reestimating the model by changing equation 11 to

$$\Phi_4(L)(h_t - \mu_t) = \lambda_{h\tau}\eta_t + \lambda_{hy}\epsilon_y + \lambda_{hs}\epsilon_s + \lambda_{he}\epsilon_e + \epsilon_h$$

where

$$\mu_t = \mu_{t-1} + \epsilon_{\mu t}$$

and $\epsilon_{\mu t}$ is assumed to be Gaussian with variance σ_μ and uncorrelated with all other shocks (following Basistha, 2009), leads to very small estimates for σ_μ . The median estimate is 0.051 (0.033) in the first subsample, 0.050 (0.041) in the second subsample, and all the other parameter estimates are virtually identical to those presented in Table 2. Using the BLS index of aggregate weekly hours (cutting the first subsample to 1964q1-1983q4) or Rios-Rull and Santaeulalia-Llopis' measure of aggregate hours (keeping the full sample 1948q1-2011q4) to calculate hours per employee leads to very similar qualitative results. The impact coefficients are different in magnitude (reflecting the scaling of the hours series), but the ratio of the pre-break to the post-break impact coefficients is the same as the ratios obtained using Francis and Ramey's BLS series for aggregate hours.

[10] Therefore assuming that the unit root and stationarity tests did not have correct finite sample size or power.

1.6. Conclusions

In this chapter, I have investigated the importance of a structural change in the relationship between permanent and transitory movements in real macroeconomic activity for explaining the past three jobless recoveries. A large part of the movement in employment and hours can still be explained by cyclical fluctuations, and the relative importance of transitory movements has not changed significantly after 1984. Most of the change in the dynamics of employment is due to an increase in the relative importance of sales shocks, and due to a change in the sensitivity and hours per employee to these shocks. After 1984, sales shocks play a much more important role in explaining the movements in employment.

Rather than committing to opening full-time positions, firms can accommodate temporary increases in demand by temporarily increasing hours until sales have picked up enough for the recovery to be considered robust. The results obtained using disaggregated data confirm that firms have switched towards increasing output by increasing hours rather than by increasing employment, meaning that it became easier to increase hours per current employee (or aggregate hours by utilizing flexible labor inputs) rather than by increasing employment, but this switch is not primarily driven by a change in the persistence of output or sales movements.

Further extensions of this work will explore the relationship between real economic activity, labor market variables, and the financial sector, allowing for the possibility that movements in the financial sector and financial conditions can lead to slow recoveries in the labor market, as proposed in the theoretical models used by Gu (2012) and Abo-Zaid (2012). Also, I plan to allowing for a distinct recession or bounceback phase in the cyclical components of the variable, where the switching times are estimated from the data, and not necessarily restricted to happen exactly during NBER recessions. In addition, once more data becomes available, allowing for an additional break at the time of the Great Recession will allow me to test if the behavior during the Great Recession and recovery was significantly different from the behavior during the 1990 and 2001 recession.

References

[1] Aaronson, D.E., R. Rissman and D. G. Sullivan. (2005): "Assessing the Jobless Recovery." Federal Reserve Bank of Chicago Economic Perspectives, 28.

[2] Abo-Zaid, S. (2012): "On Credit Frictions as Labor-Income Taxation." Economic Letters (forthcoming).

[3] Andolfatto, D. and G. Macdonald (2005): "Jobless Recoveries." Washington University in St. Louis Working Paper.

[4] Bachmann, R. (2011): "Understanding the Jobless Recoveries After 1991 and 2001." Working paper.

[5] Basistha, A. (2009): "Hours per Capita and Productivity: Evidence from Correlated Unobserved Components Model." Journal of Applied Econometrics, 24.

[6] Benati, L. and T. A. Lubik (2012): "Sales, Inventories, and Real Interest Rates: A Century of Stylized Facts." Federal Reserve Bank of Richmond Working Paper WP12-02.

[7] Berger, D. (2011): "Countercyclical Restructuring and Jobless Recoveries." Yale University Working Paper.

[8] Engemann, K. M. and M.T. Owyang. (2010): " Whatever Happened to the Business Cycle? A Bayesian Analysis of Jobless Recoveries." Macroeconomic Dynamics, Vol. 14.

[9] Faberman, R. J. (2008): "Job Flows, Jobless Recoveries, and the Great Moderation." Philadelphia Federal Reserve Bank Working Paper.

[10] Galeotti, M., L. J. Maccini, F. Schiantarelli. (2005): "Inventories, Employment, and Hours." Journal of Monetary Economic, Vol. 52(3), pp. 575-600.

[11] Gu, W. (2012): "Benefit Costs and Employment Dynamics in Recent U.S. Recoveries." Cornell University Working Paper.

[12] Francis, N. and V. Ramey (2009): "A Century of Work and Leisure." American Economic Journal: Macroeconomics, Vol. 1(2), pp. 189-224.

[13] Gali, J. (1999): "Technology, employment, and the business cycle: do technology shocks explain aggregate fluctuations?" American Economic Review 89.

[14] Gali J. and P. Rabanal. (2004): "Technology Shocks and Aggregate Fluctuations: How Well does the RBC model fit Postwar US data?" NBER Macroeconomics Annual 225–228.

[15] Gali, J., F. Smets, and R. Wouters. (2012): "Slow Recoveries: A Structural Interpretation." Working paper.

[16] Gomme, P. (2005): "Accounting for the Jobless Recoveries." Federal Reserve Bank of Cleveland Working Paper.

[17] Gordon, R. (1993): "The Jobless Recovery: Does it Signal a New Era of Productivity-Led Growth?" Brookings Papers on Economic Activity, Vol. 1993, No 1, pp 271-316.

[18] Gordon, R. (2012): "The Evolution of Okun's Law and of Cyclical Productivity Fluctuations in the United States and in the EU-15." Northwestern University Working Paper.

[19] Groshen, E.L. and S. Potter. (2003): "Has Structural Change Contributed to a Jobless Recovery?" Current Issues in Economics and Finance 9-1-7.

[20] Hetrick, R. L. (2000): "Analyzing the recent upward surge in overtime hours." Monthly Labor Review 123.

[21] Hodgson, A., S. L. Schreft and A. Singh. (2005): "Jobless Recoveries and the Wait-and-see Hypothesis." Federal Reserve Bank of Kansas City Economic Review 90.

[22] Jaimovich, N. and H. E. Siu. (2012): "The Trend is the Cycle: Job Polarization and Jobless Recoveries." Duke University Working Paper.

[23] Kahn J. and R. Rich (2007): "Tracking the new economy: using growth theory to detect changes in trend productivity." Journal of Monetary Economics 54: 1670–1701.

[24] Kim, C.J. and C. Nelson (1999a): "Friedman's Plucking Model of Business Fluctuations: Tests and Estimates of Permanent and Transitory Components. Journal of Money, Credit, and Banking 31(3): 317-334

[25] Kim,C. J. and C. Nelson. (1999b): "Has the U.S. Economy Become More Stable? A Bayesian Approach Based on a Markov-Switching Model of the Business Cycle." The Review Of Economics and Statistics, 81 (4), 608-616.

[26] Koenders, K. and R. Rogerson. (2005): "Organizational Dynamics and the Business Cycle: A View on Jobless Recoveries." Federal Reserve Bank of St. Louis Review 87.

[27] Lilien, D. M. (1982): "Sectoral Shifts and Cyclical Unemployment." Journal of Political Economy. 90. pp. 777-93.

[28] McConnell, M.M. and G. Perez-Quiros. (2000): "Output Fluctuations in the United States: What has Changed Since the Early 1980's?" American Economic Review 90(5), pp. 1464-1676.

[29] Morley J., C. Nelson C and E Zivot. (2003): "Why Are Beveridge–Nelson and Unobserved-Component Decompositions of GDP so Different?" Review of Economics and Statistics 85: 235–243.

[30] Morley, J. and A. Singh. (2012): "Inventory Mistakes and the Great Moderation." Working Paper, The University of Sydney Economics Working Paper Series.

[31] Owyang, M.T., J. Piger and H. J. Wall. (2008): "A State-Level Analysis of the Great Moderation." Regional Science and Urban Economics 38.

[32] Ramey, V. A. and D. J. Vine. (2006): "Declining volatility in the U.S. automobile industry." American Economic Review, 96.

[33] Rios-Rull J. V. , R. Santaeulalia-Llopis, F. Schorfheide, C. Fuentes-Albero and M. Khrysko (2012): "The Effects of Technology Shocks on Hours." Journal of Monetary Economics (forthcoming).

[34] Rios-Rull J. V. and R. Santaeulalia-Llopis. (2010). "Redistributive Shocks and Productivity Shocks."Journal of Monetary Economics, Vol. 57(8), pp. 931-948 .

[35] Santaeulalia-Llopis, R. (2012): "How Much are SVARs with Long-Run Restrictions Missing without Cyclically Moving Factor Shares?" Washington University in St. Louis Working Paper.

[36] Schreft, S. L. and A. Singh. (2003): "A Closer Look at Jobless Recoveries." Federal Reserve Bank of Kansas City Economic Review 88.

[37] Sinclair, T. (2009):The Relationships between Permanent and Transitory Movements in U.S. Output and the Unemployment Rate." Journal of Money, Credit, and Banking, Vol. 41, No 2-3.

[38] Sinclair, T. (2010): "Asymmetry in the Business Cycle: Friedman's Plucking Model with Correlated Innovations." Studies in Nonlinear Dynamics & Econometrics Vol. 14: Iss.1, Article 3.

[39] Stock, J. and M. Watson (2012): "Disentangling the Channels of the 2007-2009 Recession." Working Paper.

2. State-Dependent Effects of Fiscal Policy

2.1. Introduction

The Great Recession and the subsequent American Recovery and Reinvestment Act fiscal stimulus package reignited debate, academic and otherwise, about the stabilization role of discretionary fiscal policy. More broadly, these developments have raised questions about the relevance of aggregate demand and government spending as possible engines of economic activity. As a way to answer these questions, many recent academic studies have sought to determine whether government spending has significant effects on aggregate output and components of output, such as consumption and investment.

This debate is of central importance not only for economic policy, but also for the insights it provides into the underlying structure of modern developed economies. Theoretical models in which resources are fully employed predict that the direct effect of a positive shock to government spending, given preferences and technology, should completely crowd out private economic activity. The direct government spending multiplier arising from such models is zero, at least as a first approximation.[11] In contrast, traditional Keynesian models predict that the economy will not always fully employ available resources, possibly for extended periods of time, because of insufficient demand. If output is below its potential level due to insufficient aggregate demand, an increase in government spending can directly motivate the employment of idle resources and raise output.

Much of the recent empirical research on fiscal policy considers the effects of spending shocks on different components of output. On the one hand, a baseline neoclassical model predicts crowding out of both consumption and investment, and therefore implies negative responses of these variable to a positive shock to government spending. On the other hand, if government spending raises resource use through traditional Keynesian channels, consumption and investment should respond positively to spending shocks.

[11] These models can generate indirect allocational effects of government spending on output, but of ambiguous sign. For example, the higher interest rate or negative wealth effect (see, for example, Parker, 2011) induced by a rise in government spending could encourage higher labor supply that raises output, but higher interest rates also reduce capital accumulation that lowers output.

Indeed, these spillovers create the possibility that the government spending multiplier exceeds one because higher government spending induces increases in other components of demand.

Most of the existing empirical research on fiscal policy employs linear time series models in which the size of the response of output or other variables to government spending is independent of the state of the economy (important recent exceptions include Mittnik and Semmler, 2012, and Auerbach and Gorodnichenko, 2012a; see the following subsection). The results from such models are useful, especially if the maintained null hypothesis is the neoclassical baseline of zero effects on output and crowding out of consumption and investment. But the Keynesian alternative suggests an important state dependence (and therefore a nonlinearity) in the effect of any demand shock on output, including a government spending shock. Higher demand cannot raise output indefinitely. Eventually, resource constraints bind: even a Keynesian economy behaves like a neoclassical system if demand is sufficiently high. Threshold models provide a natural econometric framework for exploring this basic state dependence. If government spending shocks affect output through Keynesian demand channels, such effects will be be larger when the economy has significant resource slack than when it is operating at or near full capacity. The purpose of this chapter, then, is to test this simple, but fundamentally important hypothesis.

To investigate the possibility of state-dependent effects of fiscal policy, I estimate a nonlinear structural vector autoregressive model that allows parameters to switch according to whether a threshold variable crosses an estimated threshold. As candidate threshold variables, I consider several alternative measures of economic slack, as well as the debt-to-GDP ratio and a measure of the real interest rate. Various statistical and economic criteria identify capacity utilization (adjusted for a structural break) as the best threshold variable, but the main findings are robust to the other measures of slack.

The empirical results provide strong evidence in favor of state-dependent nonlinearity; specifically, government spending shocks have larger effects on output when they occur with relatively low resource utilization than when they appear at times of high resource use. Furthermore, threshold estimates for capacity utilization place half or more of its historical observations in the low-utilization regime. This evidence implies, therefore, that the "normal" state of the U.S. economy is one in which positive demand shocks have large positive and persistent effects on output and its components. I also employ simulation-based impulse-response functions to isolate the different effects of fiscal policy under particular economic conditions. I find that the responses of output and output components depend crucially on the state of the economy when a policy shock occurs.

The rest of the chapter is organized as follows. Section 2.2 reviews previous research that has estimated the aggregate effects of fiscal policy in a time-series context. Section 2.3 introduces the baseline empirical model and the estimation method. Section 2.4 presents the empirical results and extends the baseline model to models that include consumption, investment, and other variables of interest. Section 2.5 concludes.

2.2. Related Literature

The empirical literature that explores the effects of government spending on macroeconomic variables, both old and new, is divided in its findings. Most studies fall in one of four main strands: models based on traditional Keynesian theory, structural vector autoregressive (SVAR) models, dynamic stochastic general equilibrium (DSGE) models, and models based on the narrative approach first introduced by Ramey and Shapiro (1998).

Traditional Keynesian models usually relate an outcome variable such as aggregate output to different components of spending or taxes, typically with a reduced-form, linear specification. The interest rate is usually held fixed over the whole forecasting horizon, and the multipliers obtained from those models are often very large (always greater than 1, sometimes as big as 4). In particular, the American Recovery and Reinvestment Act (ARRA) fiscal stimulus package was designed based on a study of this kind by Romer and Bernstein (2009) that estimated a short-run multiplier for output of approximately 1.6.

Studies based on SVAR models in which government spending is assumed to be predetermined typically find that output, consumption, and real wages increase after a positive government spending shock. Blanchard and Perotti (2002) and Perotti (2008) find that the response of output and consumption to government spending is positive and persistent, although, perhaps surprisingly, they find a negative response of investment. This discrepancy between the positive response of consumption (implied by Keynesian models), and the negative response of investment (implied by neoclassical models) is commonly referred to as the "investment puzzle" in the fiscal policy literature. The magnitude of the estimated effects depends on the identification of the model. Blanchard and Perotti (2002) and Perotti (2008) use institutional information to identify the shocks, and they get government spending multipliers for output that are about 1.3. Mountford and Uhlig (2009) use an alternative approach based on sign restrictions, and they get a smaller government spending multiplier for output of 0.5 and a multiplier for consumption that is very close to zero. These SVAR studies are sometimes criticized because they do not allow for state-dependent responses (see Parker, 2011, in particular), an issue addressed here.

Most DSGE studies are based on a New Keynesian model with Calvo pricing frictions and make a variety of assumptions about whether interest rates adjust to an increase in government spending. Pappa (2009) uses an DSGE model with Calvo pricing and a fixed interest rate. She finds that output, real wages, and consumption rise, while investment falls, but the magnitude of the responses is very sensitive to the parameterization of the model. Some DSGE models allow for different responses to government spending shocks depending on the interest rate. For example, Cogan et al. (2010) consider the importance of interest rate responses to fiscal policy by holding the interest rate pegged at zero for four quarters, but afterwards allowing the interest rate to revert to the "natural level" determined by neoclassical first-order conditions. The estimated government spending multiplier for output from this model is only 0.4, and their simulations yield multipliers that are above one only when the interest rate is fixed exogenously.[12]

Van Brusselen (2009), Leigh et al. (2010), and Ramey (2011a) provide extensive surveys

[12] Although Barro and Redlick (2011) use a simple linear regression model with annual data instead of a DSGE model, they find that the government spending multiplier for output is just 0.3 on impact, similar to the DSGE results.

of the empirical literature on fiscal multipliers. Consistent with the summary above, they show that the estimated government spending multipliers are highly sensitive to the model and parameters. In particular, Van Brusselen (2009) compares a wide variety of empirical models and points out that, even within the same class of models (i.e., DSGE models with Calvo pricing), the government spending multiplier for output varies between -3.7 and 3.7, depending on how the increase in spending is financed, how long the interest rate is pegged, and whether the economy is closed or open. Ramey (2011a) also points out the sensitivity of estimates, but concludes that the multiplier for the U.S. economy likely lies within a range between 0.8 and 1.5. Leeper et al. (2012) examine the variability of multipliers for output and consumption in DSGE models in great detail, and they conclude that the estimates are very sensitive to the specification of the model. In particular, when the proportion of rule-of-thumb consumers is large, as in Galí et al. (2007), the multipliers are large and positive, but they are smaller when the proportion of rule-of-thumb consumers is closer to zero.

Another strand of the literature uses a narrative approach to identify exogenous government spending shocks. Ramey and Shapiro (1998), Eichenbaum et al. (1999), and Ramey (2011b) find that the response of output is small and short-lived, but they use military spending rather than total government spending. Ramey (2011b) argues that the SVAR-based multipliers are large only because they fail to capture the importance of timing of government shocks and because the combined narrative data Granger-causes the government spending shocks. However, Perotti (2007) shows that lagged government spending, tax, and GDP shocks also predict the Ramey-Shapiro narrative dates.

In this chapter, I adopt a nonlinear SVAR-based approach that allows the impact of shocks to depend on the level of resource utilization at the time of the shock. I do not impose any assumptions about the response of interest rates or about the degree of price or wage stickiness, and check the robustness of the results to including military spending separately in the model. This approach nests the possibility of empirical results consistent with neoclassical models that predict small or even negative responses of output and other macro variables to positive government spending shocks. But to the extent that the data identify Keynesian effects, the model allows these effects to vary with the state of the economy that prevails at the time of the shock; that is, I test whether positive responses are stronger when there is substantial slack in the economy. The size and difference in the multipliers across regimes is of fundamental interest, and I also consider what the data tell us about the amount of time the economy spends in the different regimes.

This nonlinear approach is strongly suggested by Parker (2011) and it is similar in motivation to recent studies by Auerbach and Gorodnichenko (2012a) and Mittnik and Semmler (2012), amongst others.[13] Auerbach and Gorodnichenko (2012a) estimate a smooth transition threshold SVAR model for government spending, taxes, and output, in which they impose the restrictions that government spending has different effects during recessions and expansions, and they calibrate the smoothness parameter based on U.S. data so that the economy spends about 20% of the time in recessions. They estimate that the effects of government spending are large and positive (1.7 over 20 quarters) when the economy is in a recession and smaller (very close to one) when the economy is not in a recession.[14] They control for the state of the business cycle by using a moving average of output growth as the threshold variable, and they impose that the threshold around which the behavior changes is equal to the mean of output growth.[15] Mittnik and Semmler (2012) estimate a bivariate threshold model for output and employment where the threshold output is lagged output growth and threshold is predetermined and equal to the mean of output growth. In their model, the responses of employment to output

[13]Recent working papers taking a nonlinear approach include Afonso, Baxa, Slavík (2011), Baum and Koester (2011), Candelon and Lieb (2011), Shoag (2013), and Owyang, Ramey, and Zubairy (2013). Also, Kirchner, Cimadomo, and Hauptmeier (2010) take a related time-varying parameter structural VAR approach.

[14]Bachmann and Sims (2012) estimate a very similar nonlinear SVAR model to Auerbach and Gorodnichenko (2012a) and find the same result that government spending shocks have larger effects during recessions than during expansions. Their additional insight is that these larger effects during recessions appear to operate largely through consumer confidence. In particular, if the response of consumer confidence to government spending shocks is shut down in the calculation of impulse response functions, the effects are much smaller and similar to the estimated effects in expansions (with or without the consumer confidence channel).

[15]In a followup to their original study, Auerbach and Gorodnichenko (2012b) find that their results for the U.S. data are largely robust across a large number of OECD countries given the same restrictions to identify recessions, but considering a panel structure and direct multi-period single-equation projections to calculate impulse response functions. Their consideration of a panel structure and single-equation projections rather than an SVAR model is motivated in part by a lower frequency of available data for many countries, making statistical identification of a nonlinear SVAR challenging.

shocks are much larger in the low regime than in the high regime.

The approach used here differs from these two other nonlinear studies of fiscal policy and aggregate demand in a number of important ways. First, I estimate the threshold that determines state-dependent effects from the data, whereas Auerbach and Gorodnichenko (2012a) and Mittnik and Semmler (2012) both assume their thresholds *a priori.* Because the threshold is estimated, the model allows the data to sort observations into possibly different multiplier regimes. Based on the estimated thresholds, I find evidence that the U.S. economy spends the majority of its time in the low-utilization/ high-multiplier state, a possibility not allowed for by the other two studies. Second, I consider a threshold SVAR model with a discrete change in regime instead of the smooth transition specification considered by Auerbach and Gorodnichenko (2012a). Although the smooth transition specification is potentially more general, estimating the smoothness parameter for such a model can be challenging, as evidenced by the fact that Auerbach and Gorodnichenko (2012a) fix this parameter (as well as the threshold) in their estimation. The difficulty is that the likelihood function for a smooth transition model is flat when the true smoothness parameter is large in the sense of implying a relatively discrete threshold, making maximum likelihood estimation and even Bayesian estimation unreliable. I circumvent this econometric problem by considering a discrete threshold only, which still allows us to focus on the primary question of whether there are state-dependent effects of fiscal policy. Third, I investigate the role of capacity constraints in generating potential state-dependent effects by considering various measures of economic slack as threshold variables, rather than just the growth rate of output, which was the focus of the other two studies. I also include each measure of slack as an endogenous variable in the SVAR model to allow the possibility that the variable is important for understanding the effects of fiscal policy, but does not necessarily relate to a state-dependent effect. Fourth, unlike the other two studies, I formally compare the nonlinear models to their linear counterparts using Bayesian marginal likelihood analysis.[16]

[16]One exception is Candelon and Lieb (2011), who consider a classical hypothesis test for nonlinearity in their threshold vector error correction model with fiscal variables and find supportive evidence for nonlinearity.

2.3. Empirical Methods

2.3.1. Model

The basic vector autoregression (VAR) model is linear, and cannot capture nonlinear dynamics such as regime switching and asymmetric responses to shocks. For the analysis used here, I consider a nonlinear version of a VAR model that extends the threshold autoregressive model of Tong (1978, 1983) to a multivariate setting. Threshold models work by splitting a time series process endogenously into different regimes. Within each regime the process is described by a linear model. In particular, I specify a threshold version of a reduced-form VAR model as follows:

$$Y_t = \Phi_0^1 + \Phi_1^1(L)Y_{t-1} + (\Phi_0^2 + \Phi_1^2(L)Y_{t-1})I[q_{t-d} > \gamma] + \varepsilon_t \tag{12}$$

where Y_t is a vector containing the first difference of the logarithm of real government spending, the first difference of the logarithm of net taxes, the first difference of the logarithm of real GDP, and an mean-adjusted measure of capacity utilization, as discussed in more detail below. This is the baseline version of the model. However, I also consider alternative versions of the model that incorporate the private-sector components (i.e., consumption, investment, exports, and imports) and other outcome variables of interest, again discussed in more detail below.

The lag polynomial matrices $\Phi_1^1(L)$ and $\Phi_1^2(L)$ capture the dynamics of the system. The disturbances ε_t are assumed to be independent and Gaussian with mean zero. Rather than assuming that the disturbances are strictly i.i.d., the covariance matrix of ε_t is set to be equal to Ω until 1984Q1 and equal to $\lambda\Omega$ afterwards to capture the Great Moderation. Because the focus of this paper is not on determining the exact break date in volatility and because there is near consensus in the literature about the general timing of the volatility break (see, for example Kim and Nelson, 1999, or McConnell and Perez-Quiros, 2000),the break date is set exogenously. By using a scale factor λ and a constant variance-covariance matrix Ω, I am also assuming that the correlations between the endogenous variables do not change over time or over regimes. This is important to ensure that any apparent state-dependent effects of fiscal policy are not merely due to different identification of structural shocks at different points in time, but rather reflect a change in the propagation of shocks. The threshold variable q_{t-d} determines the prevailing regime; γ is the threshold parameter at which the regime switch occurs. The indicator function $I[\cdot]$ equals 1 when the q_{t-d} exceeds the threshold γ and 0 otherwise. The integer d is the delay lag for the threshold switch; that is, if the threshold variable q_{t-d} crosses γ at time $t-d$, the dynamics actually change at time t. For the threshold variable, I consider capacity utilization, other measures of economic slack, and a selection of other macroeconomic variables, as discussed in more detail below.

2.3.2. Data

In addition to capacity utilization, I also consider the output gap estimated by the CBO, the unemployment rate, output growth, and employment growth to measure economic slack. The traditional Keynesian theory summarized above implies that the threshold variable should measure the level of economic activity and intensity of resource use. For this purpose, the output gap, the level of capacity utilization or the unemployment rate would seem to be good choices. However, I also consider first differences of these variables and output and employment growth to check the robustness of the results and to explore whether threshold effects might relate to growth (as in Auerbach and Gorodnichenko, 2012a, and Mittnik and Semmler, 2012) rather than to levels.

Government spending and net taxes are defined as in Blanchard and Perotti (2002). The full sample period is 1967Q1-2011Q1. All output components are measured in real terms and are seasonally adjusted by the source. The series for output, its components, including government spending, and tax revenues were obtained from NIPA-BEA, and the capacity utilization series was obtained from the Federal Reserve Statistical Releases website. I also consider data for U.S. federal government debt, the Federal Funds Rate, inflation based on the CPI (seasonally adjusted), and non-farm payroll employment, which were all obtained from the Federal Reserve Bank of St. Louis FRED website. The monthly series for capacity utilization, the unemployment rate, the Federal Funds Rate, CPI, and employment are all converted to a quarterly frequency by using simple arithmetic means.

I use growth rates rather than log-levels in the VAR because the logarithms of real GDP and output components appear to have stochastic trends according to standard unit root and stationarity tests, but there is no support for common trends amongst the variables in any version of the VAR model under consideration. Specifically, Johansen cointegration tests suggest the absence of cointegrating relationship between government spending and taxes, between government spending, taxes, and output, as well as between government spending, taxes, and output components (consumption, investment, exports, and imports). However, the results are roughly similar results when imposing cointegration between spending and taxes.

2.3.3. Specification Issues

The lag length for the VAR model is chosen based on the baseline linear VAR model estimated using maximum likelihood. For this model, both AIC and SIC select four lags, which is also the number of lags used by Blanchard and Perotti (2002), Ramey (2011b), and most other studies that use the linear SVAR approach. Unlike Mittnik and Semmler (2012), who allow the number of lags to differ across regimes, the number of lags is assumed to be the same in each regime, and we consider a model with two regimes.

To solve for the SVAR given the reduced-form VAR parameters, I impose short-run zero restrictions with government spending ordered first and taxes ordered second in all models; i.e., government spending is assumed to respond to economic conditions only with a lag, but economic conditions are allowed to respond immediately to government spending. Implicitly, the approach to solving the SVAR used here assumes that the impact matrix identifying structural shocks remains the same across regimes and throughout the entire sample period, with only the size of structural shocks allowed to undergo a structural break in 1984. This approach avoids any ambiguity about whether the dynamic effects of government spending shocks appear state-dependent because of a change in their identification rather than their propagation.

Economic theory implies several possible choices for the threshold variable. As discussed above, traditional Keynesian theory suggest that the dynamics may depend on the state of the economy, while DSGE models imply that the effects of government spending depend on the interest rate. A recent literature also suggests that the dynamics may depend on debt (see, for example, Reinhart and Rogoff, 2009, and Eggertsson and Krugman, 2012, for two very different views on the impact of debt on the efficacy of fiscal policy). Because I do not want to fix the threshold variable *a priori*, I consider a large set of possible threshold variables and select the preferred threshold variable using Bayesian model comparison. The threshold variables considered are

1. lagged output: output growth, long differences in the natural log of output, moving averages of differences in the natural log of output
2. lagged CBO output gap
3. lagged capacity utilization: level, level adjusted for long-run change in mean, first differences, and first differences of the mean-adjusted series
4. lagged unemployment rate: level, differences, mean-adjusted level, differences in the mean-adjusted series
5. lagged debt-to-GDP ratio: total Federal debt and total Federal debt held by the public, both as a percent of GDP
6. lagged real interest rate: level and change in the ex ante real interest rate based on the Federal Funds Rate and CPI inflation under the assumption of static expectations

Both capacity utilization and the unemployment rate appear to have changes in their long-run mean levels, which would make those series unsuitable for use in a stationary VAR model. Standard tests for a structural break at an unknown break date reject the null of no break in mean for both capacity utilization and the unemployment rate. Meanwhile, there is some debate about whether the unemployment rate has a unit root or whether there were just exogenous structural breaks in its mean (see, for example, Papell et al., 2000). For both series, therefore, I consider the level, first differences, the mean-adjusted levels, and the differences of the mean-adjusted levels as possible threshold variables. Table 10 summarizes the results of the test for structural breaks in mean for capacity utilization and the unemployment rate. A structural break test for capacity utilization identifies a highly significant break (F statistic of 41.7) in the level of capacity utilization in 1974Q1, which coincides with the well-known productivity slowdown. The structural break tests also identify three breaks in mean for the unemployment rate. Notably, the mean-adjusted capacity utilization series is strongly correlated with other commonly-used measures of economic activity, as shown in Figure 9, so it appears to be a highly representative measure of economic slack.[17]

2.3.4. Estimation and Inference

Because the threshold VAR model is highly parameterized, I make inferences about the threshold, the coefficients, the threshold variable, and the delay parameter using Bayesian methods; in particular, I use a multi-block Metropolis-Hastings (MH) algorithm described in detail in the Appendix to sample from marginal posterior distributions for parameters and calculate marginal likelihoods for models. The advantages of using a Bayesian approach in this setting are twofold. First, it allows us to capture the uncertainty about the parameter values when constructing the impulse response functions. Second, despite the presence of nuisance parameters in the nonlinear models, comparing the linear to the nonlinear model and examining the presence of nonlinear effects is straightforward in the Bayesian framework.

To provide an accurate approximation of the target posterior distribution of the pa-

[17] Morley and Piger (2012) also find that capacity utilization is closely related to their asymmetric measure of the business cycle based on a model-averaged forecast-based trend/cycle decomposition given a wide range of linear and nonlinear time series models of quarterly U.S. real GDP.

rameters, I follow the standard approach in the applied literature and use a tailored multivariate Student$-t$ distribution as the proposal distribution. The prior for Φ is a normal distribution, truncated to ensure stationarity. Ω is inverse-Wishart, λ is Gamma, and γ is uniform over $[q_l, q_h]$ where q_l and q_h are the highest and the lowest observed values of the the threshold variable.[18]The priors are diffuse in the sense that the likelihood strongly dominates the priors for all model parameters. The full technical details of the posterior sampler and the priors are relegated to the Appendix.

A crucial empirical question is whether the effects of government spending really do differ across regimes defined by economic slack. In a frequentist setting, to test for the presence of nonlinear effects, one would consider the null hypothesis $H_0 : \Phi_0^2 = \Phi_1^2 = 0$ that the coefficients are equal against the alternative that at least one of the elements of the matrices Φ_0^2 , Φ_1^2 is not zero. This testing problem is tainted by the fact that the threshold γ is not identified under the null. If the errors are i.i.d., a test with near-optimal power against alternatives distant from the null hypothesis is the $\sup LR$ test, but the asymptotic distribution of the test statistic is nonstandard and has to be approximated using Hansen's (1996, 1997) bootstrap procedure. Because the model is very parameter-rich, bootstrapping the asymptotic distribution is computationally prohibitive. Also, it should be noted that the 1984Q1 structural break in the variance-covariance matrix of the disturbances makes it unclear how well Hansen's procedure would perform in this setting. The Bayesian approach circumvents such problems by providing a direct method for comparing models based on the posterior odds ratios.[19] In particular, I estimate the threshold VAR model using the MH algorithm and then compare its marginal likelihood to that for a restricted linear version of the VAR model for which $\Phi_0^2 = \Phi_1^2 = 0$.

Marginal likelihoods are calculated using Chib and Jeliazkov's (2001) algorithm and I compare models based on Bayes factors, which are the ratio of marginal likelihoods and

[18]Using a truncated univariate Student$-t$ prior for γ with mean equal to the maximum likelihood estimate and 5 degrees of freedom (relatively flat over the observed values) leads to very similar posterior estimates.

[19]It should be noted, however, that a bootstrap version of the $\sup LR$ test for a simpler version of the model with only government spending, net taxes, and real GDP as endogenous variables and still using capacity utilization as the threshold variable is significant at the 5% level (under the assumption that the structural break does not distort the test).

are equal to posterior odds ratios under even prior odds (i.e., equal prior probabilities on all models under consideration).

To estimate the effects of shocks to government spending, I calculate impulse-response functions for output and and other outcome variables in response to a shock to government spending. I do this for two reasons. First, the impulse responses give us the magnitude of the response of output and its components to government spending, so they can be used to define the multiplier. Second, when it comes to designing policies, the response of output is much more important than the coefficient estimates. Because the impulse responses are nonlinear functions of the coefficients, a small asymmetry in the coefficients might correspond to a large asymmetry in the impulse responses or vice versa. Rejecting nonlinearity implies that the impulse responses are necessarily different across regimes. However, because the responses are complicated highly nonlinear functions of the coefficients, the degree of this asymmetry can only be evaluated by looking at the IRFs, rather than solely by looking at the coefficients.

For the nonlinear model, I construct two sets of impulse responses. In the first case, the economy is assumed to remain in a given state forever. Because the model is linear within a state, the impulse response functions can be obtained by using the estimated VAR coefficients for the given regime. In the second case, the state of the economy is allowed to evolve because the threshold variable itself responds to government spending shocks. When the system is allowed to evolve and switch between regimes, the impulse response function depends on the initial state and possibly on the size and the sign of the shock. Following Koop, Pesaran and Potter (1996), I consider generalized impulse response functions (GIRFs) in order to obtain the responses when the threshold variable is allowed to respond endogenously. The impulse responses are defined as the change in conditional expectation of Y_{t+k} as a result of a shock at time t:

$$E[Y_{t+k}|shock_t, \Psi_{t-1}] - E[Y_{t+k}|\Psi_{t-1}] \tag{13}$$

where Ψ_{t-1} is the information set at time t-1. Calculating the GIRFs requires specifying the nature of the shock and the initial conditions Ψ_{t-1}, and then the conditional expectations $E[Y_{t+k}|shock_t, \Psi_{t-1}]$ and $E[Y_{t+k}|\Psi_{t-1}]$ are computed by simulating the model. Similar to Kilian and Vigfusson (2011), I consider an orthogonal exogenous shock identified from the SVAR model rather than a forecast error from the reduced-form VAR, as was considered in Koop, Pesaran and Potter (1996).

In practice, the simulation-based structural IRFs are computed as follows (a detailed version of the algorithm is presented in the Appendix): First, shocks for periods 0 to 20 are simulated using the estimated variance-covariance matrix for the threshold SVAR model and, for given initial values of the variables, fed through the estimated model to produce a simulated data series. The result is a forecast of the variables conditional on initial values and a particular sequence of shocks. Next, the same procedure is repeated with the same initial values and shocks, except that the shock to government spending in period 0 is fixed at 1 percent of GDP (for that particular starting value of GDP). The shocks are fed though the model and a forecast is produced just as above. The difference between this forecast and the baseline forecast is the impulse response function for a particular sequence of shocks and initial values. This computation is repeated for five hundred draws of the shocks and averaged to produce impulse response functions conditional only on a particular history. These IRFs are then averaged over a particular subset of initial values.

Because threshold models imply that the predicted responses from the model to a shock depend on a particular history, I first simulate the responses for the evolving model, averaging over all histories when the threshold variable is above the estimated threshold and averaging over states when it is below. Then I compare those results to those obtained when the GIRFs are simulated for the recent histories between 1984-2011 when the threshold is above the threshold and when it is below, including the "New Economy" rapid expansion in the late 1990s and the "Great Recession" in the late 2000s. To capture the uncertainty about the parameter values, the credibility intervals for the impulse response functions are obtained by simulating the IRFs for all iterations of the MH algorithm.

2.4. Empirical Results

2.4.1. Model Comparisons

As discussed in Section 2.3, the formal model comparisons are based on marginal likelihoods and the implied Bayes factors. Table 11 reports marginal likelihood values for the baseline model with different threshold variables, including the restricted case of no threshold effect. Despite relatively diffuse priors on parameters, which should penalize more highly-parameterized models, the implied Bayes factors strongly favor nonlinearity when considering threshold variables related to economic slack. There is also support for nonlinearity when the threshold variable is lagged output growth, as considered in Auerbach and Gorodnichenko (2012a) and Mittnik and Semmler (2012), although neither of those studies formally tested for nonlinearity. By contrast, there is no support for nonlinearity when considering the debt-to-GDP ratio or the real interest rate. For the debt-to-GDP ratio, the estimated threshold is near the boundary of the parameter space considered, so the lack of support for nonlinearity in this case might reflect the relatively low levels of the debt-to-GDP ratio in U.S. economy since 1967, at least compared to the levels observed in other countries that have suffered debt crises. For the real interest rate, the estimated threshold is about 2 percent, which is often thought to be close to the long-run "neutral" rate. However, this estimate is quite imprecise, consistent with the lack of support for a threshold effect relating to the interest rate.

Based on the marginal likelihood results in Table 11, the preferred threshold variable for the baseline model is the first lag of capacity utilization (adjusted for a one-time structural break in the mean, as discussed in Section 2.3). The estimated threshold for the baseline model is slightly below the mean of the adjusted capacity utilization series.[20] The mean-adjusted capacity utilization series and its estimated threshold are plotted in Figure 10. The high-utilization regime dates estimated with capacity utilization as the threshold variable largely coincide with the high-utilization regime dates found when

[20] The sample mean for unadjusted capacity utilization is 85.2 for 1967Q1-1974Q1 and 80.0 for 1974Q2-2011Q1.

using the CBO's output gap estimate as a measure of slack and with the regime dates obtained when using output or unemployment as measures of slack.

Notably, more than 60% of the historical observations for mean-adjusted capacity utilization fall below the mode of posterior distribution for the threshold parameter, while close to 50% fall below the posterior mean. This result is important because it implies that, for a majority of the time since the middle 1960s, the U.S. economy has operated in a regime in which government spending shocks have relatively large effects on output. Since 2000, almost all observations have been in the low regime. A possible interpretation of these results is that demand, not supply, has been the proximate constraint on aggregate output for much of the sample period. This result also distinguishes the approach used here from Auerbach and Gorodnichenko (2012a), as their approach imposes that only 20% of the observations fall in a recessionary regime.

Although the marginal likelihood results in Table 11 strongly favor nonlinearity, it is important to address Sims' (2001) concern that evidence for time-varying parameters in VAR models may be the spurious result of failing to fully account for heteroskedasticity. Therefore, I consider some diagnostic tests for the preferred baseline model with mean-adjusted capacity utilization. The model allows for some heteroskedasticity given that it incorporates a one-time structural break in the scale of the variance-covariance matrix for the VAR residuals corresponding to the Great Moderation in 1984Q1. For this model, the standardized residuals based on the parameter values at the posterior mean pass the Jarque-Bera test for Normality of the individual residual series and the Doornik-Hansen test statistic for multivariate Normality is 10.54 (p-value 0.23). Also, there is no evidence of serial correlation in the standardized residuals based on Ljung-Box Q-tests and the ARCH-LM test does not reject the null of a constant variances for the individual residual series. Thus, the evidence for nonlinearity does not appear to be an artifact of failing to fully account for heteroskedasticity. Instead, it appears that any heteroskedasticity is successfully captured by allowing for a structural break in the scale of the variance-covariance matrix for the VAR residuals.

When I estimate the effects of government spending on output components and other variables, I substitute the outcome variable of interest (i.e., consumption, investment, exports, imports, the unemployment rate, employment, and inflation) for output in the baseline VAR model, using the first lag of mean-adjusted capacity utilization as the threshold variable.[21] As with the baseline model, I find strong evidence of nonlinearity for these models. Table 3 reports the marginal likelihood values (and the modes of the likelihood) for linear and nonlinear specifications of these alternative VAR models. In every case, the nonlinear specification is preferred. In particular, the implied Bayes factors always prefer the nonlinear specification, with posterior odds only favoring linear specifications under extremely high prior odds of more than 10 to 1 on the linear specifications. Again, the strong evidence for nonlinearity is particularly notable given relatively diffuse priors on parameters. Meanwhile, Table 13 also reports the estimated thresholds in mean-adjusted capacity utilization for these alternative VAR models and shows that they are quite consistent across the different models.

2.4.2. Responses of Output to a Government Spending Shock

As discussed above, the government spending shocks in the SVAR model is identified by assuming output and its private-sector components can respond to government spending within a quarter, but government spending does not respond to output within the same quarter. The results are similar when considering alternative identification schemes: specifically, I obtain almost identical results when I reorder taxes and government spending so that spending can respond to tax shocks, when I use Blanchard and Perotti's (2002) identification scheme that imposes short-run tax elasticities, and when adding Ramey's narrative spending variable and order it first so that the rest of government spending can respond to military spending within a quarter.[22]

[21] In preliminary analysis, I also considered these effects by adding each series as a fifth variable to the baseline model. The point estimates for the threshold and the median impulse responses were very similar for both specifications, but the 95% (and even the 75%) credibility intervals were very wide in the specification with five variables because there are too few observations per regime to precisely estimate a threshold VAR with so many variables without imposing very tight priors. Thus, the results presented in the rest of this chapter are based on four-variable versions of the VAR model.

[22] Owyang and Zubairy (2013) also find impulse response functions for SVAR models are broadly robust when considering different identification schemes, including sign restrictions. They consider a linear VAR model that includes U.S. state-level data and separates out military spending, as in Ramey (2011b).

When constructing the impulse response functions to government spending, the initial shock to government spending is set to be equal to 1 percent of GDP. The shock to government spending initiates a dynamic path of adjustment for both government spending and other variables of interest. To make the interpretation of the results more straightforward and to facilitate comparison with the linear results obtained by Blanchard and Perotti (2002), the dynamic multipliers are calculated as ratios of the cumulative dollar-for-dollar change in the variable of interest to the cumulative dollar-for-dollar change in government spending. Meanwhile, when I examine the behavior of employment, unemployment, and inflation, the responses are presented as level responses to a shock to government spending equal to 1 percent of GDP.

The primary results appear in Figure 11. The top row of the figure shows the impulse responses of output to a government spending shock for the two regimes, in both cases assuming that the economy remains in the same state forever. The response of output to spending shocks depends strongly on the regime. The shock to government spending pushes output up immediately in both the high and the low utilization regimes. However, in the low regime, output rises almost monotonically to a cumulative change in output equal to 1.6 times the cumulative change in government spending. Most of the effect takes place in the first three years. In the high-utilization regime, the pattern is substantially different. After the initial positive response, the cumulative change in output falls back towards zero. The long-term response is positive, but the multiplier is less than half of that when output is in the low-utilization regime.

When the economy is allowed to evolve from one state to another, the magnitude of the multiplier varies depending both on the state of the economy at the time of the government spending shock and on the actual history of other shocks. As shown in the bottom two rows of Figure 11, the output response for all low states peaks at 1.6 after two years and then the effects of the spending shock die out. The lower bound of the credibility interval for the low-regime impulse response is strongly positive, despite the fact that I use a fairly conservative 90% credibility interval. In comparison, the average response for all high states peaks at 0.8 after two years, and then it remains stable, but the credibility interval always covers zero.

Thus, the estimates clearly suggest that the effects of government spending on output are larger and more persistent when capacity utilization is low. In the following subsections, I examine the source of this result about the asymmetric response of output to fiscal policy shocks in more detail. In particular, I look at the responses of output components in order to determine whether the asymmetry comes from an asymmetry in the response of fiscal variables to the government spending shock or if it is due to an asymmetric response in the components of private spending.

2.4.3. Responses of Fiscal Policy to a Government Spending Shock

From Figure 12, it is clear that the response of government spending to its own shock does not depend very strongly on the prevailing regime. In this case, the IRFs are shown as cumulative dollar-for-dollar changes in government spending relative to the size of the initial shock, because the ratio of government spending to itself is necessarily equal to one. For both regimes, the peak cumulative dollar-for-dollar change is roughly 1.3, which is consistent with the results obtained in the linear case by Blanchard and Perotti (2002) and similar to the results obtained by Auerbach and Gorodnichenko (2012a). Both the credibility intervals for the regimes overlap and the actual estimated responses are similar across regimes. The similar responses across regimes clearly indicate then that the asymmetric response of output is not due to higher or more persistent government spending in the low-utilization regime.

Figure 13 shows that the peak response of tax revenues to a government spending shock is roughly 0.8 in evolving regimes, with little effect of the initial state of the economy. In the fixed low-utilization regime, tax revenues appear to increase persistently after a government spending shock, while the response of tax revenues is smaller and dies off quickly when the economy starts and remains in the high regime. But, given the wide credibility intervals for the responses at long horizons, there is no obvious evidence of an asymmetry in the response of tax revenues that could explain the asymmetry in the response of output to a government spending shock.[23] Thus, the asymmetry in the response of output to a fiscal policy shock appears be due to an asymmetry in the response of private spending, not the government sector of the economy.

2.4.4. Responses of Consumption and Investment to a Government Spending Shock

Figure 14 displays the responses of consumption to a government spending shock. The main result is that consumption increases in both regimes, but the magnitude of the response is much larger when the economy is in the low-utilization regime.[24] When starting from a low-utilization state, but allowing the state to evolve, the long-run response levels off after three years at close to 0.8, averaging over all histories (the effect is even larger when averaged over recent histories). Consumption is much less responsive when the economy starts in a high-utilization state. The peak response in this case is only around 0.4, and becomes insignificant after a year. Thus, it appears that the asymmetric response of output to government spending is at least partly due to an asymmetry in the magnitude of the response of consumption. Meanwhile, the findings of a positive response of consumption in both regimes is consistent with the linear results obtained by Blanchard and Perotti (2002), Pappa (2009), and Woodford (2011). Also, accounting for anticipated government spending by including Ramey's military spending variable and ordering it first in the linear or nonlinear versions of the SVAR does not change the response or the significance of the response in either case.

Figure 15 displays the responses of investment, which also appear to be asymmetric depending on the state of the economy. In the fixed low regime, investment increases in response to government spending, with a peak response of 0.4, although the credibility

[23] It is important to note that these results are for the responses of tax revenues, not tax rates. Tax revenues are correlated with income, so part of the increase in revenues comes from increases in income due to the positive government spending shock, indicating that spending could be partially self-financing (although further analysis would be necessary to examine this possibility given the wide credibility intervals). Another part of the increase in revenues could come from the endogenous response of tax rates to a government spending shock. The use of tax revenues also makes it difficult to interpret the responses of output and its components to changes in taxes because individuals and firms respond to marginal tax rates. Unfortunately, though, reliable data for marginal tax rates are only available at an annual frequency.

[24] This result is robust to considering consumption of nondurables and services only.

interval includes zero. When the economy is assumed to remain in the high-utilization state forever, investment drops significantly in response to a spending shock, with a cumulative decline equal to 0.9 after five years. Allowing the economy to evolve from one regime to another, the responses of investment are weakly positive when the economy starts from a low-utilization state and not different from from zero when the economy starts from a high-utilization state. These results indicate the relevance of crowding out in the high-utilization state, but provide no support for crowding out in the low-utilization state. Furthermore, these results may help explain the "investment puzzle" in linear studies such as Blanchard and Perotti (2002) because the negative response in the linear VAR is roughly a weighted average of the responses in the nonlinear model. Specifically, the apparent neoclassical behavior of investment found in these studies appears to reflect crowding out when capacity utilization is high only.

Overall, the strong state-dependence in the responses of consumption and investment suggests that a lot of the asymmetry in the response of output is due to different responses of these key components to government spending depending on the degree of resource utilization.

2.4.5. Responses of Other Macroeconomic Variables to a Government Spending Shock

Figure 16 displays the fixed-regime responses of exports and imports, which are very similar across regimes. Both exports and imports weakly increase, and the increase in exports roughly cancels out the increase in imports. Furthermore, the credibility intervals for the state-dependent responses overlap, suggesting little support for asymmetric responses of imports and exports to government spending. For brevity, I do not report responses when allowing the state of the economy to evolve, although not surprisingly given these results for the fixed regimes, those responses display very little asymmetry.

Figure 17 shows that the unemployment rate decreases in response to a spending shock in both states. In the low-utilization regime, the unemployment rate decreases monotonically, falling by a total of 2.5 percentage points after five years. The effect of a spending shock on the unemployment rate is weaker and less persistent when the economy is in the high-utilization regime. The impact response is essentially zero, and the maximum response (in magnitude) is a 1.3 percentage point decline. When analyzing the magnitude of the responses, it is important to keep in mind that the impulse responses were constructed using a relatively large spending shock (1 percent of the GDP), which might explain the large responses of the unemployment rate.

The responses of employment also exhibits state-dependence that is consistent with the responses of the unemployment rate. In Figure 18, when the economy is in a low-utilization state, employment increases by 1 percent after two years, and the long run response is equal to 0.8 percent. When the economy starts from a high-utilization state, the effect of a government spending shock on employment is only slightly positive and transitory. The credibility intervals for employment, however, are quite wide, and zero effects are not outside the 90% interval for either regime. This result is due to the fact that I use a conservative 90% interval and the fact that employment only builds up slowly after the shock. In particular, the estimated employment effects are only large at longer horizons, while credibility intervals are always wide at longer horizons for SVAR models.

Figure 19 displays the responses of the real interest rate and inflation. In the fixed low regime, there is little response of either variable. Thus, monetary policy appears to accommodate fiscal policy when capacity utilization is low, with little implication for inflation (perhaps due to a convex Phillips curve). Notably, this accomodation of fiscal policy does not just occur in a zero-lower-bound environment (see Christiano et al. 2011, and Woodford, 2011), but is the apparent response of monetary policy whenever the economy is in the low-utilization state.[25]In the fixed high regime, an increase in government spending has a more persistent effect on inflation and triggers a delayed, but large response of the interest rate. The estimated responses are consistent with the idea

[25]It is also notable that the real interest does not appear to be an important variable in linear SVAR models of fiscal policy (e.g., it is absent from Blanchard and Perotti's, 2002, model) or as a threshold variable in a nonlinear model (see our results in Table 2). The implication is that the different responses of monetary policy are primarily determined by the state of the economy as captured by capacity utilization regimes, not by other factors such as a binding zero nominal lower bound (which, of course, only occurs near the end of the sample period) that might influence the behavior of real interest rates.

that government spending can crowd out resource use, thus increasing marginal costs when the economy is close to capacity, but monetary policy responds to keep inflation under control. The response of the interest rate in the high-utilization state is large both economically and statistically, but the credibility intervals for the responses of inflation in both regimes are quite wide (due to the VAR polynomial having a unit root relatively close to one). For brevity, the responses in the evolving regimes are not displayed, as they are very similar.

2.4.6. Counterfactual Analysis

One of the main criticisms of the ARRA fiscal stimulus is that output growth remained anemic two years after it was first implemented and that the unemployment rate remained persistently high. However, in order to fully evaluate whether the stimulus had any effect on the economy, it is important to compare what output and the unemployment rate would have been if there had been no stimulus in the first place. Taking into account that in the months before the stimulus package was implemented, interest rates were already approaching the zero bound and that employment was dropping precipitously, the absence of fiscal stimulus could have resulted in even lower output growth and the unemployment rate rising even higher than its 10.1 percent peak.

In order to evaluate the effects of the stimulus, I make use of the preferred model to perform counterfactual analysis for the period 2009-2010 in which I compare the implied path of output and employment without increases in government spending to the actual path of output and spending. This is done this by orthogonalizing the shocks for each period and setting the orthogonalized government spending shocks equal to zero for the period 2009-2010.[26] Figure 20 displays the results of the counterfactual experiments. If I set the spending shocks between 2009 and 2010 equal to zero, the economy would have needed one more quarter to recover (i.e., the simulated output growth would not have

[26] Taylor (2011) notes that the overall impact of the ARRA stimulus on government spending was not as large as advertised because state and local government saved a large portion of their stimulus grants. Thus, the implied effects on output and employment reported below should not be thought of in terms of a one-time exogenous $700 billion dollar increase in federal government spending, but rather in terms of the much smaller exogenous changes in total government spending over the period under consideration.

become positive until 2009Q3) and the recovery would have been even more sluggish, with the maximum simulated growth rate reaching 0.4 percent at the end of 2009 and then dropping down to zero by the end of 2010.

The results for unemployment are similar. Without fiscal stimulus, the simulated estimate for the unemployment level is 0.7 percentage points higher in 2010Q2 than what the actual unemployment rate turned out to be. Furthermore, without the spending shocks the unemployment rate would have stayed higher (around 10.3 percent), without dropping below double digits at any point. Interestingly, the results based on the counterfactual analysis are similar to the CBO estimates for the effects of the ARRA stimulus and support the idea that, even if employment and output growth did not reach the high rates that are typical for a recovery, the fiscal stimulus was helpful in the sense that it prevented the recession from becoming deeper and longer.

2.5. Conclusions

In this chapter, I have presented strong empirical evidence in favor of state-dependent effects of fiscal policy. In particular, the estimates from a threshold structural vector autoregressive model clearly identify different responses of the economy to government spending shocks depending on whether the economy has high or low utilization of economic resources. I find that a rise in demand from the government sector causes large and persistently positive effects on output when the economy is operating with relatively low capacity utilization. This effect is much smaller and less persistent when capacity utilization is above an estimated threshold for the model.

It is particularly interesting to note that the estimated threshold for capacity utilization is such that a majority of observations for the U.S. economy over the past 40 years appear to have been in the low-utilization regime in which demand shocks have larger and more persistent effects, with any constraints from the supply side binding less tightly. This result implies that the normal state for the U.S. economy is one of significant excess capacity. Therefore, the proximate effect of a demand shock is more likely than not to be positive and persistent.

There is no empirical evidence that higher government spending crowds out consumption. Indeed, consumption rises after positive government spending shocks in both the high- and low-utilization regimes, but the increase is almost twice as large during "normal" times (i.e., low utilization) than during "booms" (i.e., high utilization). Most of the increase in the private components of output comes from the increase in consumption. These results for consumption are consistent with the linear results obtained by Blanchard and Perotti (2002), Perotti (2008) and Pappa (2009), but are at odds with the simulation results obtained using most calibrated dynamic stochastic general equilibrium (DSGE) models. Only when allowing for a high proportion of rule-of-thumb consumers do Galí et al (2007) find such large responses of consumption in an estimated (not calibrated) DSGE model. Meanwhile, the state-dependent responses of consumption are potentially related to the results obtained by Kaplan and Violante (2012), who develop a life-cycle model that endogenizes the proportion of rule-of-thumb consumers in order to examine the effect of taxes on consumption when a large proportion of the consumers' wealth is tied up in illiquid assets such as real estate. Historically, the number of credit-constrained consumers rises in recessions, and the Great Recession started with the crash of the housing market, which likely implied a large increase in the proportion of credit-constrained consumers in its aftermath. Even more directly along these lines, Canzoneri et al. (2012) calibrate a New Keynesian DSGE model with costly financial intermediation and show that countercyclical shocks to the spread between rates paid by borrowers and received by depositors implies countercyclical fiscal multipliers, although this is a fairly mechanical result given the assumptions of countercyclical spread shocks and the ability of government spending shocks to disproportionately lower borrowing costs when the level of output is lower.

In terms of investment, we also find little evidence of crowding out in the low regime. However, investment does fall in the high-utilization regime in response to higher government spending. This finding is consistent with our results for real interest rates. Specifically, interest rates do not appear to respond to government spending in the low regime, but they rise strongly with government spending when resource constraints are tight. A possible explanation for these results is that monetary policy is much more accommodative of fiscal stimulus when there is a large degree of economic slack and little pressure on inflation, with the degree of monetary accommodation determining the output multiplier, as suggested by Woodford (2011).

Regardless of the exact mechanism behind the state-dependent effects of fiscal policy, the implications for policy are straightforward and significant. Higher government spending raises output, but this effect is both larger and more persistent when capacity utilization is low. At these times, including during recessions, higher government spending reduces economic slack and increases output, consumption, and investment. Although stimulus policy to reduce slack may increase government debt, the effect is smaller than a simple calculation would suggest because higher government spending raises output, income, and therefore tax revenue, and the effect of spending stimulus on public debt is less than dollar for dollar.

Further extensions of this work will explore policy implications more deeply. In particular, because the "low-utilization" regime prevails in at least half of the sample period, it would be interesting to consider whether allowing a third regime would identify recession effects when stimulus policy might be even more effective. Also, beyond the state-dependent nonlinearities found here, there may be additional asymmetries in the response of output to the size and sign of changes in fiscal policy. Finally, I have made preliminary analysis of tax shocks and found some comparable results to those for government spending shocks. But identifying tax shocks is challenging due to a lack of availability of quarterly data on tax rates instead of tax revenues, for which movements are largely endogenous (see, for example, the May 2012 issue of the *American Economic Journal: Economic Policy* for a number of studies illustrating the challenges in identifying the effects of tax shocks, even within a linear framework). Thus,a more complete analysis of possible state-dependent effects of tax shocks is left for future research.

References

[1] Alfonso, A., J. Baxa, and M. Slavik (2011): Fiscal Developments and Financial Stress: A Threshold VAR Analysis." ECB Working Paper No. 1319.

[2] Andrews, D. and W. Ploberger (1994): "Optimal Tests when a Nuisance Parameter Is Present Only under the Alternative." *Econometrica* 62(6), 1383-1414.

[3] Auerbach, A. J. (2009): "Implementing the New Fiscal Policy Activism." *American Economic Review: Papers and Proceedings* 99-2, 543-549.

[4] Auerbach, A. J. and Y. Gorodnichenko (2012a): "Measuring the Output Responses to Fiscal Policy." *American Economic Journal: Economic Policy* 4(2), 1-27.

[5] Auerbach, A. J. and Y. Gorodnichenko (2012b): "Fiscal Multipliers in Recession and Expansion."A. Alesina and F. Giavazzi, eds. *Fiscal Policy after the Financial Crisis*, Chicago: University of Chicago Press, forthcoming.

[6] Bachmann, R. and E. R. Sims (2012): "Confidence and the Transmission of Government Spending Shocks." *Journal of Monetary Economics* 59, 235-249.

[7] Barro, R. J. and C. J. Redlick (2011): "Macroeconomic Effects from Government Purchases." *Quarterly Journal of Economics* 126(1), 51-102.

[8] Baum, A. and G. B. Koester. (2011): "The Impact of Fiscal Policy on Economic Activity over the Business Cycle: Evidence from a Threshold VAR." Deutsche Bundesbank Discussion Paper 03/2011.

[9] Blanchard, O., and R. Perotti (2002): "An Empirical Characterization of the Dynamic Effects of Changes in Government Spending and Taxes on Output." *The Quarterly Journal of Economics* 117(4), 329-1368.

[10] Calvo, G. A. (1983): "Staggered Contracts in a Utility Maximizing Framework." *Journal of Monetary Economics* 12, 383-398.

[11] Candelon, B. and L. Lieb. (2011): "Fiscal Policy in Good and Bad Times.: Masstricht University Working Paper RM/11/01.

[12] Canzoneri, M., F. Collard, H. Dellas, and B. Diba (2012): "Fiscal Multipliers in Recessions." Working Paper.

[13] Chib, S. and I. Jeliazkov (2001): "Marginal Likelihood from the Metropolis-Hastings Output." *Journal of the American Statistical Association* 96, 270-281.

[14] Christiano, L. M. , M. Eichenbaum, and S. Rebelo (2011): "When Is the Government Spending Multiplier Large?" *Journal of Political Economy* 119(1), 78-121.

[15] Cogan, J. F. , T. Cwik, J.B. Taylor, and V. Wieland (2010): "New Keynesian Versus Old Keynesian Government Spending Multipliers." *Journal of Economic Dynamics and Control* 34, 281-295.

[16] Eggertsson, G.B. and P. Krugman (2012): "Debt, Deleveraging, and the Liquidity Trap: A Fisher-Minsky-Koo Approach." *Quarterly Journal of Economics*, forthcoming.

[17] Galí, J., D. López-Salido, and J. Vallés (2007): "Understanding the Effects of Government Spending on Consumption." *Journal of the European Economic Association* 5, 227-270.

[18] Hansen, B. E. (1996): "Inference when a Nuisance Parameter is not Identified under the Null Hypothesis." *Econometrica* 64, 413-430.

[19] Hansen, B. E. (1997): "Inference in TAR Models." *Studies in Nonlinear Dymanics and Econometrics* 2(1), 1-14.

[20] Inoue, A. and L. Kilian (2012): "Inference on Impulse Response Functions in Structural VAR Models." *Journal of Econometrics*, forthcoming.

[21] Kaplan G. and G.L. Violante (2012) "A Model of the Consumption Response to Fiscal Stimulus Payments." NBER Working Paper.

[22] Kilian, L. and R.J. Vigfusson (2011): "Are the Responses of the U.S. Economy Asymmetric in Energy Price Increases and Decreases?" *Quantitative Economics* 2, 419-453.

[23] Kim, C-J. and C. R. Nelson (1999): "Has the US Economy Become More Stable? A Bayesian Approach Based on a Markov-Switching Model of the Business Cycle." *The Review of Economics and Statistics* 81(4), 608-616.

[24] Kirchner, M., J. Cimadomo, and S. Hauptmeier (2010): "Transmission of Government Spending Shocks in the Euro Area: Time Variation and Driving Forces." ECB Working Paper No. 1219. and W. Ploberger (1994): "Optimal Tests when a Nuisance Parameter Is Present Only under the Alternative." *Econometrica* 62(6), 1383-1414.

[25] Koop, G. M.H. Pesaran, and S.M. Potter (1996): "Impulse Response Analysis in Nonlinear Multivariate Models." *Journal of Econometrics* 74, 119-147.

[26] Leigh, D. , P. Devries, C. Freedman, J. Guajarado, D. Laxton, A. Pescatori (2010): "Will it Hurt? Macroeconomic Effects of Fiscal Consolidation." IMF Working Paper, October 2010.

[27] Leeper, E.M., N. Traum, and T.B. Walker (2011): "Clearing up the Fiscal Multiplier Morass." Working Paper.

[28] Lo, C. M. and J. Morley (2013):"Bayesian Analysis of Nonlinear Exchange Rate Dynamics and the Purchasing Power Parity Persistence Puzzle. " Working Paper, University of New South Wales.

[29] McConnell, M. M. and G. Perez-Quiros (2000): "Output Fluctuations in the United States: What Has Changed Since the Early 1980's? " *American Economic Review* 90(5), 1464-1676.

[30] Mittnik, S. and W. Semmler (2012): "Regime Dependence of the Fiscal Multiplier." *Journal of Economic Behavior & Organization* 83 (3), 502–522.

[31] Morley, J. and J. Piger (2012): "The Asymmetric Business Cycle." *The Review of Economics and Statistics* 91(1), 208-221.

[32] Mountford, A. and H. Uhlig (2009): "What Are the Effects of Fiscal Policy Shocks?" *Journal of Applied Econometrics* 26, 960-992.

[33] Owyang, M.T., V. A. Ramey and S. Zubairy (2013): "Are Government Spending Multipliers Greater During Periods of Slack? Evidence from 20th Century Historical Data." Working Paper.Owyang, M.T., V. A. Ramey and S. Zubairy (2013): "Are Government Spending Multipliers Greater During Periods of Slack? Evidence from 20th Century Historical Data." Working Paper.

[34] Papell, D.H., C.J. Murray, and H. Ghiblawi (2000): "The Structure of Unemployment." *The Review of Economics and Statistics* 82(2), 309-315.

[35] Pappa, E. (2009): "The Effects of Fiscal Shocks on Employment and Real Wages." *International Economic Review* 50(1), 217-244.

[36] Parker, J. A. (2011): "On Measuring the Effects of Fiscal Policy in Recessions." *Journal of Economic Literature* 49, 703-718.

[37] Perotti, R. (2008): "In Search of the Transmission Mechanism of Fiscal Policy." *NBER Macroeconomic Annual* 2007, ed. by D. Agemogly, K. Rogoff, and M. Woodford, pp. 169-226. University of Chicago Press, Chicago, Illinois.

[38] Perotti, R. (2008): "In Search of the Transmission Mechanism of Fiscal Policy." *NBER Macroeconomic Annual* 2007, ed. by D. Agemogly, K. Rogoff, and M. Woodford, pp. 169-226. University of Chicago Press, Chicago, Illinois.

[39] Ramey, V. A. (2011a): "Can Government Purchases Stimulate the Economy?" *Journal of Economic Literature* 126(1).

[40] Ramey, V. A. (2011b): "Identifying Government Spending Shocks: It's All in the Timing." *The Quarterly Journal of Economics.*

[41] Ramey, V. A. (2012) "Government Spending and Private Activity." A. Alesina and F. Giavazzi, eds. *Fiscal Policy after the Financial Crisis*, Chicago: University of Chicago Press, forthcoming.

[42] Ramey, V. A. and M. Shapiro (1998): "Costly Capital Reallocation and the Effects of Government Spending." *Carnegie Rochester Conference on Public Policy* 48, 145-194.

[43] Reinhart, C.M. and K.S. Rogoff (2009): "This Time Is Different: Eight Centuries of Financial Folly." Princeton University Press.

[44] Shoag, D. (2013): "The Impact of Government Spending Shocks: Evidence of the Multiplier from State Pension Plan Resturns." Harvard University Working Paper.

[45] Ramey, V. A. (2011a): "Can Government Purchases Stimulate the Economy?" *Journal of Economic Literature* 126(1).

[46] Ramey, V. A. (2011b): "Identifying Government Spending Shocks: It's All in the Timing." *The Quarterly Journal of Economics.*

[47] Ramey, V. A. and M. Shapiro (1998): "Costly Capital Reallocation and the Effects of Government Spending." *Carnegie Rochester Conference on Public Policy* 48, 145-194.

[48] Reinhart, C.M. and K.S. Rogoff (2009): "This Time Is Different: Eight Centuries of Financial Folly." Princeton University Press.

[49] Romer, C. and J. Bernstein (2009): "The Job Impact of the American Recovery and Reinvestment Plan." *Executive Office of The President of the United States, Council of Economic Advisers Report*, January 8, 2009.

[50] Romer, C. and D. Romer (2009): "The Macroeconomic Effects of Tax Changes: Estimates Based on a New Measure of Fiscal Shocks." *American Economic Review* 100, 763-801.

[51] Sims, C.A. (2001): "Comment on Sargent and Cogley's 'Evolving US Postwar Inflation Dynamics'." *NBER Macroeconomics Annual* 16, 373-379.

[52] Taylor, J.B. (2011): "An Empirical Analysis of the Revival of Fiscal Activism in the 2000s" *Journal of Economic Literature* 49(3), 686-702.

[53] Tong, H. (1978): "On a Threshold Model in Pattern Recognition and Signal Processing." C. H. Chen, Amsterdam: Sighoff and Nordhoff.

[54] Tong, H. (1983): "Threshold Models in Non-linear Time Series Analysis." New York: Springer-Verlag.

[55] Van Brusselen, P. (2009): "Fiscal Stabilization Plans and the Outlook for the World Economy." *NIME Policy Brief* 01-2009.

[56] Woodford, M. (2011): "Simple Analytics of the Government Expenditure Multiplier." *American Economic Journal: Macroeconomics* 3(1), 1-35.

3. The Asymmetric Effects of Fiscal Stimulus and Austerity Across the Business Cycle

3.1. Introduction

The world-wide Great Recession, the implementation of fiscal stimulus packages intended to ameliorate the recession, and the subsequent austerity measures led to renewed interest, both academic and political, about the stabilization role of discretionary fiscal policy. Some of the key questions that have arisen during the past few years are whether government spending has significant effects on aggregate output and components of output at all, and whether there is evidence of sign and size asymmetry in the responses of output and output components over the business cycle. These two questions are closely related, but distinct, especially when it comes to designing fiscal policy.

The first question is whether there is any evidence in favor of implementing a stimulus or cuts at all. Many studies find large fiscal multipliers that support positive effects of fiscal expansions and negative effects of austerity on output (see Ramey, 2011, for a recent survey). A more nuanced case for the appropriate timing of fiscal expansions comes from recent evidence that the effects of fiscal policy depend on the state of the economy. Auerbach and Gorodnichenko (2012a,b), Mittnik and Semmler (2012), and Fazzari, Morley, and Panovska (2013), find that government spending multipliers, on average, are larger during periods of macroeconomic slack. These results support the recommendation for using fiscal stimulus to increase output in times of recession or, in general, in times when the economy is below full capacity. Conditional on the existence of state dependence, the second question helps us determine the optimal size and timing for a stimulus and the optimal time and size for an austerity measure. For example, if there is evidence that a larger stimulus is more effective, dollar-for-dollar, than a small stimulus in times of recession, then it would be optimal to implement a single large stimulus package rather sequence of smaller measures. Similarly, evidence that austerity measures have a much smaller negative impact on output during robust recoveries than during weak recoveries suggest that "fiscal consolidation" has lower economic cost if implemented in good times. In this chapter, I address the second question and examine whether there is evidence that large increases in discretionary government spending affect output and output components differently than small changes, and whether cuts in government spending have effects that are different in magnitude than an increase of the same size.

In this chapter, I extend the literature on fiscal spending in two ways. First, I introduce an impulse-response based approach that allows me to formally compare the responses of output to a positive spending shock across states of the economy, and to cuts and large shocks within a state. This method can be used to directly evaluate whether there are state-dependent effects. Second, I examine whether there is evidence of sign and size asymmetry within a given state. To investigate the possibility of sign and size asymmetries within a given state, I build on the work of Fazzari, Morley, and Panovska (hereafter, FMP1), and develop an impulse response cased comparison. The framework used in FMP1 naturally lends itself to analyzing the impact of fiscal spending, because it allows for state-dependent asymmetries, and for sign and size asymmetries within a given state. For example, it allows for the possibility that the effects of a positive spending shock equal to 1% of GDP to be different in normal times and in booms, and also for the possibility that a 1% cut during a recession or a boom may not have effects that are the same in magnitude as the effects of a 1% increase.

The impulse-response model comparison presented here corroborates the results obtained using general model-based marginal likelihood comparison in FMP1. There is strong evidence in favor of state-dependent nonlinearity; specifically, government spending shocks have larger effects on output when they occur with relatively low resource utilization than when they appear at times of high resource use. We find that the responses of output and output components depend crucially on the state of the economy when a policy shock occurs. Furthermore, when the economy is not resource constrained, a large fiscal stimulus is more effective at increasing employment and output than a small stimulus. In contrast when the economy is close to capacity, a large stimulus has Less than proportional effects on output than a smaller stimulus. Consumption exhibits the same asymmetric responses as output. A government spending cut crowds in investment more than an expansion crowds out investment in the constrained state. Higher government spending crowds in investment in the unconstrained state, but the response is not significant, and there is no significant evidence of sign or size asymmetry. When the economy is very close to capacity, fiscal cuts have larger effects on payroll employment than fiscal expansions. The effects on private employment are smaller in magnitude, but qualitatively similar to the effects of cuts and expansions on overall employment.

The rest of this chapter is organized as follows. Section 3.2 reviews, very briefly the empirical literature that estimates fiscal multipliers. Section 3.3 introduces the baseline empirical model and the estimation method. Section 3.4 presents the empirical results for output and extends the baseline model to models that include consumption, investment, and other variables of interest. Section 3.5 concludes.

3.2. A Brief Review of Fiscal Multiplier Estimates

Before the Great Recession, most of the empirical research on fiscal policy employed linear time series models in which the size of the response of output or other variables to government spending is independent of the state of the economy, independent of the size of the exogenous spending shock, and cuts and increases in spending have effects that are exactly equal in magnitude, but the implication that the outcome of either stimulus or austerity measures does not depend on the size or the timing of implementation is quite restrictive. Furthermore, the size of the fiscal multipliers obtained from linear models tends to depend crucially on the model and the assumption used.

Traditional Keynesian models keep the interest rate fixed over the whole forecasting horizon, and the multipliers obtained from those models are often very large (always greater than 1, sometimes as big as 4). Studies based on SVAR models in which government spending is assumed to be predetermined typically find that output, consumption, and real wages increase after a positive government spending shock. Blanchard and Perotti (2002) and Perotti (2008) find that the response of output and consumption to government spending is positive, large, and persistent, although, perhaps surprisingly, they find a negative response of investment. This discrepancy between the positive response of consumption (implied by Keynesian models), and the negative response of investment (implied by neoclassical models) is commonly referred to as the "investment puzzle" in the fiscal policy literature. The magnitude of the estimated effects depends on the identification of the model. Blanchard and Perotti (2002) and Perotti (2008) use institutional information to identify the shocks, and they get government spending multipliers for output that are about 1.3. Mountford and Uhlig (2009) use an alternative approach based on sign restrictions, and they get a smaller government spending multiplier for output of 0.5 and a multiplier for consumption that is positive, but very close to zero. Studies that augment the linear VAR model with a measure of expected government spending or with a narrative measure of military spending (EVARs) typically find much smaller and less persistent output multipliers. For example, Ramey (2011b, 2012) finds that the peak response of output to the narrative military spending variable is 1.2, and the response of consumption is negative. Leeper et al. (2012) examine the variability of multipliers for output and consumption in DSGE models in great detail, and they conclude that the estimates are very sensitive to the specification of the model. The estimates for output multipliers obtained vary very widely, from -4 to 4, based on the assumptions imposed on the model parameters. Parker (2012) and Ramey (2011a) provide an extensive survey of the results obtained from linear studies.

One of the reasons for the wide range of empirical estimates for the spending multiplier is the possible existence of state-dependent effects. If the size of the multiplier depends on the state of the economy, then a linear model would average over both high and low responses, or possibly over positive and negative responses, leading both to wide confidence intervals and to results that are very strongly affected by the sample period used. Recently the focus shifted towards nonlinear models where the effects of fiscal spending are allowed to vary over the business cycle. Fazzari, Morley, and Panovska (2013), FMP1 hereafter, find that the effects of government spending depend very strongly on the state of the business cycle. In particular, we estimate and formally compare whether there are threshold effects and find that the effects of a spending shock are significantly larger during periods when the economy has underutilized resources. It is important to note that 60% of the historical observations fall below the estimated threshold. This result implies that, for a majority of the time since the middle 1960s, the U.S. economy has operated in a regime in which government spending shocks have relatively large effects on output. Since 2000, almost all observations have been in the low regime. Auerbach and Gorodnichkenko (2012a,b) and Mittnik and Semmler (2012) impose a threshold that splits the sample into NBER recessions and "normal times" and they also find evidence in favor of state-dependent effects of government spending on output, with output multiplier being close to 1.8 in times of recessions. Candelon and Lieb (2012) extend the model used by Auerbach and Gorodnichenko by imposing long run equilibrium conditions. They also find strong evidence of state-dependence, but their output multiplers are smaller than those obtained by FMP1 and Auerbach and Gorodnichenko. Baum and Koester (2011) find that there is strong evidence in favor of state-dependent effects of state-dependent effects of fiscal policy in Germany, but the multipliers are smaller than the estimated multipliers for the U.S. Bachmann and Sims (2012) estimate a similar model to Auerbach and Gorodnichenko (2012a) and find the same result that spending shocks have large effects during recessions, but in their model the larger effects operate largely through consumer confidence. Shoag (2013) obtains much higher output multipliers for state level government spending during periods of slack in the labor market than during normal periods (2.12 vs 1). Owyang, Ramey, and Zubairy (2013) combine the approach used by Auerbach and Gorodnichenko (2012b) and Ramey (2011a) by imposing a threshold and augmenting the model with a narrative measure of military spending. They find state-dependent effects for Canada, but no evidence of state-dependence for the U.S.

Very few papers explored size and sign asymmetries. Baum and Koester (2011) compare the response of output to a 2% positive shock, 5% positive shock, and 2% cut to government spending for Germany. They find that large increases have proportionally larger effects in recessions, and cuts have slightly smaller effects than increases in all periods. The literature that uses U.S. has not investigated size or sign asymmetries.

3.3. Empirical Methods

3.3.1. Model

Before introducing the impulse-response-based comparison of multipliers, I start by reviewing the model used in FMP1. This model can be used both to assess if there is any evidence of state-dependence in the average response of output over the business cycle, and whether there is evidence of sign and size asymmetry. The basic vector autoregression (VAR) model is linear, and cannot capture nonlinear dynamics such as regime switching and asymmetric responses to shocks. I consider a nonlinear version of a VAR model that extends the threshold autoregressive model of Tong (1978, 1983) to a multivariate setting. I specify a threshold version of a reduced-form VAR model as follows:

$$Y_t = \Phi_0^1 + \Phi_1^1(L)Y_{t-1} + (\Phi_0^2 + \Phi_1^2(L)Y_{t-1})I[q_{t-d} > \gamma] + \varepsilon_t \tag{14}$$

where Y_t is a vector containing the first difference of the logarithm of real government spending, the first difference of the logarithm of net taxes, the first difference of the logarithm of real GDP, and a mean-adjusted measure of capacity utilization. This is the baseline version of the model. In the following subsection I also consider alternative versions of the model that incorporate the private-sector components of output (i.e., consumption and investment) and labor market variables.

The lag polynomial matrices $\Phi_1^1(L)$ and $\Phi_1^2(L)$ capture the dynamics of the system. The disturbances ε_t are assumed to be independent and Gaussian with mean zero. Rather than assuming that the disturbances are strictly i.i.d., I set the covariance matrix of ε_t equal to Ω until 1984Q1 and equal to $\lambda\Omega$ afterwards to capture the Great Moderation. As elaborated in Chapter 2 and in FMP1, the break date is set exogenously. By using a scale factor λ and a constant variance-covariance matrix Ω, I am also assuming that the correlations between the endogenous variables do not change over time or over regimes. The threshold variable q_{t-d} determines the prevailing regime; γ is the threshold parameter at which the regime switch occurs. The indicator function $I[\cdot]$ equals 1 when the q_{t-d} exceeds the threshold γ and 0 otherwise. The integer d is the delay lag for the threshold switch; that is, if the threshold variable q_{t-d} crosses γ at time $t-d$, the dynamics actually change at time t.

3.3.2. Data

Government spending and net taxes are defined as in Blanchard and Perotti (2002). The full sample period is 1967Q1-2012Q3. All output components are measured in real terms and are seasonally adjusted by the source. The series for output, its components, including government spending, and tax revenues were obtained from NIPA-BEA, and the capacity utilization series was obtained from the Federal Reserve Statistical Releases website. I also consider non-farm payroll employment and employment in the private sector, both of which were all obtained from the Federal Reserve Bank of St. Louis FRED website. The monthly series for capacity utilization and both employment series are all converted to a quarterly frequency by using simple arithmetic means.

I use growth rates rather than log-levels in the VAR because the logarithms of real GDP and output components exhibit nonstationarity. Johansen cointegration tests suggest the absence of cointegrating relationship between government spending and taxes, between government spending, taxes, and output, as well as between government spending, taxes, and output components (consumption, investment, employment, and private employment).

3.3.3. Specification Issues and Using Capacity Utilization as a Measure of Slack

The lag length for the VAR model is chosen based on AIC (for the baseline linear VAR model, estimated using maximum likelihood as a starting point). Both AIC and SIC select four lags, which is also the number of lags used by Blanchard and Perotti (2002), Ramey (2011b), and most other studies that use the linear SVAR approach.

To solve for the SVAR given the reduced-form VAR parameters, I impose the same short-run zero restrictions used in the previous chapter: government spending is ordered first and taxes are ordered second in all models; i.e., government spending is assumed to respond to economic conditions only with a lag, but economic conditions are allowed to respond immediately to government spending.

I use capacity utilization as the threshold variable for three important reasons. First, it is a natural proxy for the level of utilization of resources. Second, FMP1 consider a large set of threshold variables, and we show that the preferred threshold variable for the baseline model is the first lag of capacity utilization, adjusted for a one-time structural break in mean (see Tables 10, 11, and 12 for details). Last, as shown in Figures 9 and 21 in the appendix and in Figure 1 in FMP1, the mean-adjusted capacity utilization series is strongly correlated with other commonly-used measures of economic activity, and it is highly correlated with the asymmetric measure of the business cycle based on a model-averaged forecast-based trend/cycle decomposition, so it appears to be a highly representative measure of economic slack.

3.3.4. Estimation and Inference

I make inferences about the threshold and the coefficients using Bayesian methods; in particular, I use a multi-block Metropolis-Hastings (MH) algorithm described in detail in FMP1 to sample from marginal posterior distributions for parameters and calculate marginal likelihoods for models. An advantage of using a Bayesian approach in this framework is that it allows us to directly capture the parameter uncertainty when comparing the impulse responses. The estimated threshold for the baseline model is -0.21, slightly below the mean of the adjusted capacity utilization series.

There are two ways to compare whether the effects of government spending differ across regimes in this framework. First, we can evaluate if the model exhibits state dependence by using model comparison. Second, we can explore state dependence, and size or sign asymmetry by directly comparing impulse responses. In FMP1 we use the model comparison approach. In particular, we estimate the threshold VAR model using the MH algorithm and then we compare its marginal likelihood to that for a restricted linear version of the VAR model, specified as follows:

$$Y_t = \Phi_0^0 + \Phi_1^0(L)Y_{t-1} + \varepsilon_t \tag{15}$$

The marginal likelihoods are calculated using Chib and Jeliazkov's (2001) algorithm and we compare models based on Bayes factors, which are the ratio of marginal likelihoods and are equal to posterior odds ratios under even prior odds (i.e., equal prior probabilities on all models under consideration). The model comparison approach indicated that there is strong evidence in favor of non-linearity.

A disadvantage of the model comparison approach is that can only discern whether the coefficients are different across regimes, but it does not disentangle the sources of the nonlinearity, and it does not provide a formal comparison of the responses to positive and negative shocks, nor of the responses to shocks of different size. Rejecting nonlinearity using model comparison implies that at least one of the impulse responses to at least one of the structural shocks is necessarily different across regimes, but the degree of this asymmetry can only be evaluated by looking at the impulse response functions. This approach is better suited for examining the question at hand for two reasons. First, the impulse responses give us the magnitude of the response of output and its components to any kind of government spending shock for any history of interest, so they can be used both to define the multiplier in the usual sense and to examine the response to cuts and to increases of different sizes. Second, when it comes to designing policies, the response of output is much more important than the coefficient estimates, and policy makers are usually more concerned with the response of output or a variable of interest conditional on current economic conditions, rather than with the average response. The IRF comparison approach allows us to construct both average responses, and precisely estimated conditional responses.

To construct the impulse responses when the threshold variable is allowed to respond endogenously, I consider generalized impulse response functions (GIRFs). A GIRF is defined as the change in the conditional expectation of Y_{t+k} as a result of a shock at time t:

$$GIRF[shock_t, \Psi_{t-1}] = E[Y_{t+k}|shock_t, \Psi_{t-1}] - E[Y_{t+k}|\Psi_{t-1}] \tag{16}$$

where Ψ_{t-1} is the information set at time t-1. Calculating the GIRFs requires specifying the nature of the shock and the initial conditions Ψ_{t-1}, and then the conditional expectations $E[Y_{t+k}|shock_t, \Psi_{t-1}]$ and $E[Y_{t+k}|\Psi_{t-1}]$ are computed by simulating the model. Similar to Kilian and Vigfusson (2011), and FMP1, I consider an orthogonal exogenous shock identified from the SVAR model rather than a forecast error from the reduced-form VAR, as considered in Koop, Pesaran and Potter (1996). A detailed version of the algorithm used to compute the impulse responses is presented in the appendix. Because threshold models imply that the predicted responses from the model to a shock depend on a particular history, I can simulate the responses for the evolving model for a particular history of interest, or averaging over all histories when the threshold variable is above or below the estimated threshold. To capture the uncertainty about the parameter values, the credibility intervals for the impulse response functions are obtained by simulating the GIRFs for all iterations of the MH algorithm. As discussed in detail in the Appendix, in addition to producing a measure of the multiplier for any kind of spending shock that directly accounts for parameter uncertainty, this approach allows me to compare the impulse responses across states or the responses to different kinds of spending shocks by looking at the posterior of the distribution of the difference

$$\begin{aligned} \Delta GIRF[shock_{1t_1}, shock_{2t_2}, \Psi_{t_1-1}, \Psi_{t_2-1}] = \\ = GIRF[shock_{t_1}, \Psi_{t_1-1}] - GIRF[shock_{t_2}, \Psi_{t_2-1}]. \end{aligned} \tag{17}$$

I can evaluate if the difference between the response is significant simply by checking if zero is within a given quantile of the posterior. This approach is similar in spirit to the approach used by Kilian and Vigfuson (2011), who test for size and sign asymmetry in a frequentist setting by looking at the distribution of the impulse responses, but slightly more general, because it allows both for state-dependence and for sign and size asymmetry within a state. It is also more general than Jorda's (2005) approach used by Auerbach and Gorodnichenko (2012b) and Owyang, Ramey, and Zubairy (2013), because it allows for state-dependency, sign and size asymmetry, and parameter uncertainty.

3.4. Empirical Results

3.4.1. State-Dependence: Are the Average Responses Different?

In order to formally compare if the response of output is different in times of slacks than in times of high utilization, I construct the average impulse response for all low states and for all high states, and look at the posterior distribution for the difference in the impulse responses.

As shown in Figure 22, when the shock to government spending is fixed to equal 1% of GDP in both states, the difference between the mode of the average response in the low state and the mode of the average response in the high state peaks at 0.8, and the 90% credibility interval does not include zero during the first two years. When using less conservative 68% credibility intervals, common in the fiscal literature, the credibility interval includes zero only after 12 quarters. The results for the average responses corroborate the results obtained using the formal posterior likelihood comparison and the informal impulse response comparison in FMP1. During periods of slack, the average response of output to a spending shock peaks at 1.6. When the economy is close to being constrained with capacity utilization above the threshold, the average response is 0.8. The new result here is that the difference in the responses is significant, implying that even when accounting for parameter uncertainty, the average response in the low utilization state is larger, both statistically and economically, than the average response in the high utilization state.

The results are similar when I consider alternative identification schemes: specifically, I obtain almost identical results when we taxes and government spending are reorderedso that spending can respond to tax shocks, when using Blanchard and Perotti's (2002) identification scheme that imposes short-run tax elasticities, when adding Ramey's narrative spending variable and order it first so that the rest of government spending can respond to military spending within a quarter, and when we using Barro and Redlick's (2011) marginal tax rate, interpolated to quarterly frequency.

3.4.2. Output: Size and Sign Asymmetry within a State

As discussed above, I identify government spending shocks in the SVAR model by assuming output and its private-sector components can respond to government spending within a quarter, but government spending does not respond to output within the same quarter.

When constructing the impulse response functions to government spending, the initial shock to government spending is set to equal 1 percent of GDP. The shock to government spending initiates a dynamic path of adjustment for both government spending and other variables of interest. To make the interpretation of the results more straightforward and to facilitate comparison with the linear results obtained by Blanchard and Perotti (2002), I calculate the dynamic multipliers as ratios of the cumulative dollar-for-dollar change in the variable of interest to the cumulative dollar-for-dollar change in government spending. To account for the fact that the ratio G_t/Y_t changes over time and may not be appropriate for evaluating the average responses, as pointed by Owyang, Ramey, and Zubairy (2013), we use G_t/Y_t at each point in time. In order to assess whether there are any size and sign asymmetries within a state, I focus on two histories of particular interest and compare the responses within a state, and across those two histories. The first history Iconsider is the last quarter of 1995, and the second history we consider is the second quarter of 2008. As shown in Figure 21, the mid-late 1990s was a period of rapid growth, when all commonly used measure of economic slack were above their historic averages. The middle of the Great Recession was a time when capacity was below its average, and there was a significant level of slack. Even when accounting for uncertainty in the estimation of the threshold, the first period is above the threshold, and the second chosen history always falls below the estimated threshold, so the responses conditional on these histories have a clear interpretation as responses in the "constrained" state and in the "unconstrained" state.

The primary results appear in Figures 23 and 24. When the economy is far below capacity, as captured by the history observed in the Great Recession, a positive shock to government spending equal to 2% has bigger dollar-for-dollar effects than a shock equal to 1% of GDP. The 2% shock increases output by $1.9, whereas the smaller shock increases output by $1.6. This difference is significant during the second year, and significant during the first 3 years if we consider 68% credibility intervals. Since the responses are scaled to account for the difference in size, this means that a single larger spending shocks would increase output more than a sequence of smaller spending shocks that add up to the value of the large shock. To the extent that the cost of the fiscal stimulus is proportional to the size of the shock (for example, because the public debt created depends on the size of the shock) this result implies that it is better to implement fiscal stimulus all at once rathen than piecemeal over time. [27]

Spending cuts have larger effects than increases when capacity is far below the threshold value. As shown in the bottom panel of Figure 23, a cut decreases output by $2, a stimulus increases output by $1.6, and this difference is significant during the first 12 quarters. The sign asymmetry indirectly confirms the state-dependence found in FMP1. Following a stimulus, output and capacity increase, eventually getting closer to the constrained state, which dampens the size of the multiplier. But a decrease in spending decreases output, pushing capacity further below the threshold, leading to an output response that is very similar in shape and size to the fixed-regime responses presented in FMP1. The significance of the size asymmetry also implies that unless a cut is deliberately offset by

[27] Of course, there could be benefits of a sequence of small shocks relative to a single large shock that are not captured in our model. For example, there may be uncertainty about how much total stimulus to inject that may be partially resolved by waiting for the economy to evolve.

an alternative measure (either fiscal or monetary), it would lead to a disproportionate decrease in output if implemented during a deep recession.

When the economy is closer to the limits imposed by resource constraints (in the late 1990s), the asymmetries look different. Beginning from a state of high capacity utilization, the effect of a large spending shock is proportionally less than the effect of a small shock (the modal peak responses are 0.9 vs 1.2) . Again, this finding confirms the interpretation of the basic state-dependence results from FMP1. A large stimulus would push the economy towards full employment faster, and the converted dollar-for-dollar response is very close to the dollar-for-dollar fixed regime response.

Furthermore, cuts have smaller effects during periods of high capacity utilization. A cut in government spending decreases output by \$0.4 in the long run, implying that when the economy is doing fairly well, a spending cut is not likely to push it down in a recession. The sign asymmetry implies that while cuts do decrease output, the timing when they are implemented matters, and it is possible to minimize the negative impact of cuts.

The history-dependent sign and size asymmetries both confirm the results from FMP1 that the effects of government spending on output are larger and more persistent when capacity utilization is low, and confirm the economic intuition behind the state-dependence. In the following subsections, I examine the source of this result about the asymmetric response of output to fiscal policy shocks in more detail. In particular, I look at the responses of consumption and investment, and at the responses of employment.

3.4.3. Consumption and Investment: Size and Sign Asymmetry within a State

Figures 25 through 28 display the response of consumption of nondurables and services and of fixed plus nonresidential investment to a government spending shock.[28] Consumption increases both in the constrained and in the unconstrained state, but the magnitude of the response is different. In a rapid growth phase like the mid-late 1990s, a \$1 increase

[28] I use consumption of nondurables and services and fixed plus nonresidential investment so that the results can be easily compared to the results obtained by Ramey (2012b). The shape of the responses of consumption and gross invesment is very similar to the results shown in Figures 25 and 26.

in government spending increases consumption by \$0.4, and a \$1 increase during a deep recession increases consumption by \$0.9. The difference in the size of the consumption multiplier is consistent with the state-dependence results.

When starting in a deep recession, a large stimulus increases consumption by \$1.1, compared to the \$0.9 increase from a smaller stimulus. Cuts in government spending decrease consumption, and the response is larger than the response to increases. This difference is significant during the second year following the stimulus. There is no evidence that expansions crowd out consumption, or that cuts crowd in consumption during deep recessions. The responses of consumption when starting in a rapid expansion are very similar, both qualitatively and quantitatively, to the responses of output. Large increases in government spending have smaller impact, dollar-for-dollar, as do cuts. A 2% increase in spending increases consumption, dollar-for-dollar by \$0.15, and a 1% increase in government spending increases consumption by \$0.4. A cut initially crowds out consumption by \$0.2, then crowds it in, then crowds it out, with a long run response of \$0.1.

Figures 27 and 28 display the responses of investment. The modal responses are asymmetric depending on the state of the economy, but this asymmetry is weak and not statistically significant. When starting in a deep recession, an increase in spending increases investment during the first year (peak response \$0.9), then investment falls slightly below zero only to rise again to a long run response after 5 years of \$0.7, but the response is not significant. A cut in spending decreases investment, but there is also no evidence that this decrease is significant, nor that there is significant sign asymmetry. Similarly, there appears to be no difference in the responses of investment to large and small spending shocks. When the economy starts above the threshold (i.e. in a rapid expansion), investment is crowded out. Furthermore, a cut initially crowds in investment, then crowds it out. Even though the difference in the responses of investment is not statistically significant, it is important to note that both the shape of the impulse response and the difference in responses both for consumption and for investment are quite similar in shape and magnitude to the fixed-regime-response from FMP1, and also very similar in magnitude and shape to responses obtained from linear DSGE models.

Overall, the strong state-dependence in the responses of consumption suggest that a lot of the asymmetry in the response of output is due to different responses of these key components to government spending depending on the degree of resource utilization.

3.4.4. Employment: Size and Sign Asymmetry within a State

Figures 29 and 30 display the impulse responses of total employment, and Figures 31 and 32 display the responses of private employment. The responses of employment also exhibit state-dependence. In Figure 29, when the economy is in a deep recession, in response to a government spending shock equal to 1% employment increases by 1 percent after two years, and the long run response is equal to 0.8 percent. In response to a spending shock equal to 2% of GDP, employment increases by 2.6% (meaning that the jobs-per-dollar increase is larger for a larger spending shock). The size asymmetry is significant during the first two years. A spending cut decreases employment by 1.6%, and the difference in the size of the responses between a cut and a stimulus is significant in the second year following the spending shock.

When the economy starts in a high-utilization state, the effect of a government spending shock on employment is only slightly positive and transitory. The credibility intervals for employment, however, are quite wide, and zero effects are not outside the 90% interval for either regime. There is no evidence of size and sign asymmetry, and there is no evidence that cuts decrease employment in the high utilization state. The state dependence and the size and sign asymmetry imply that when there are underutilized physical resources in the economy, there is also a large degree of underutilized labor resources, and spending increases increase employment. Furthermore, the fact that the response to a cut in the low state is larger than the response to an increase confirms the state-dependence results.

Ramey (2012) uses a linear model where the military spending variable is ordered first, and she argues that most of the increase in employment comes from the increase in government employment, with the response of private employment being negligible.

In order to examine if the negligible response of private employment is driven by state-dependence, I also look at the responses of private employment across states. Figures 11 and 12 display the responses of private employment.

The response of private employment to a 1% spending shock in the low state is large and positive, peaking at 1.2%, and leveling off after 2 years. The cumulative response after 5 years in the late 1990s is weakly negative, and not significantly different from zero, indicating that the responses of private employment are also state-dependent. The response to a 2% spending shock is 1%, making the jobs-per-dollar effect smaller for a larger stimulus. A 1% cut in the low state decreases private employment by 0.8. In the high state, a large spending increase decreases employment by 0.5. It is interesting to note that a cut also decreases employment by 0.2% after 5 years. The results for private employment indicate that the state-dependent response of employment is not driven by the government sector. Also, the fact that in the high state a cut decreases employment, and a large increase in spending decreases employment more than a smaller increase, indicates that there are trade-offs in the high state. The full-employment state determined by the model corresponds closely to the definition of "full-employment" from Neoclassical models. A cut decreases private employment in the unconstrained state, and a large increase increases employment less than a smaller increase, but the scaled response to a large shock is still significant at all horizons, which stands in contrast to Ramey's (2012) results. This indicates that while there are trade-offs in the unconstrained state, they are much smaller than in the constrained state.

3.4.5. Cuts during a Weak Recovery

In the previous subsections I examined the responses of output, output components and labor market variables in extreme states- a deep recession, and a rapid expansion. As discussed in the previous subsections, if the goal is to make inferences about asymmetries in the relationship between government spending and the private sector, it is natural to look at the extreme cases when we are fairly certain that there is a significant degree of underutilization, or when the capacity constraints are close to binding. Similarly, if one wanted to make inference about implementing a certain policy, a stimulus is most likely to be implemented during a deep recession, and cuts are most likely to be implemented during a rapid expansion. While the extreme cases are very important, both because of their theoretical importance and because of the importance for policy, in 2010 through early 2013 the U.S. economy has presented a case study that does not fall into either category. The recovery following the recession has been slow and drawn out, and the U.S. government cut spending during a weak recovery.

While evaluating the full spectrum of policies undertaken in the period following the Great Recession would entail a much more detailed model that takes into account the individual cuts and the individual tax increases, the model used here can help us develop an intuition about the optimal time to implement austerity measures, and it can help us evaluate what kinds of effects spending cuts are likely to have during a weak recovery. Figure 33 shows the response to an increase in government spending equal to 1% of GDP, and to a decrease in government spending equal to 1% in GDP when the economy starts out in 2012Q3 (the last observation in the sample). Positive fiscal stimulus increases output by $1.5 (dollar-for-dollar), and this increase is significant the first two years following the shock. A cut decreases output by $1.8, and the negative impact does not die out even after five years. This indicates that responses during a weak recovery are much more similar to responses during recessions than to responses during rapid expansions.

3.5. Conclusions

In this chapter, I have presented strong empirical evidence in favor of state-dependent effects of fiscal policy and in favor of size and sign asymmetries. In particular, I use an impulse response based comparison to confirm the state-dependence found by Fazzari, Morley, and Panovska (2013). I find that a rise in demand from the government sector causes large and persistently positive effects on output when the economy is operating with relatively low capacity utilization. This effect is much smaller and less persistent when capacity utilization is above an estimated threshold for the model. There is no evidence that higher government spending crowds out consumption. Indeed, consumption rises after positive government spending shocks in both the high- and low-utilization regimes, but the increase is almost twice as large during “normal” times (i.e., low utilization) than during “booms” (i.e., high utilization). Both total payroll employment and private employment also exhibit strong state-dependence.

When the economy is not resource constrained, a large fiscal stimulus is more effective at increasing employment and output. In contrast, when the economy is close to capacity, a large stimulus has smaller effects on output than a smaller stimulus, consistent with the view that our results arise from traditional demand channel, but constrained when resources are limited . Consumption exhibits the same asymmetric responses as output. A cut crowds in investment more than an expansion crowds out investment in the high-utilization state. Investment rises in the unconstrained state, but the response not significant, and there is no significant evidence of sign or size asymmetry. When the economy is very close to capacity, fiscal cuts have larger effects on payroll employment than fiscal expansions. The effects on private employment are smaller in magnitude, but qualitatively similar to the effects of cuts and expansions on overall employment. The size and sign asymmetry imply that when there are underutilized physical resources in the economy, there is also a large degree of underutilized labor resources, and spending increases increase employment.

The state-dependence of the multipliers, and the sign and size asymmetries within a state indicate that both the timing, and the size of stimulus and austerity measures matters, and should be taken into account when designing policy. Further extensions of this work will explore policy implications more deeply. In particular, I plan to explore the effects of higher government spending on the dynamics of government debt in more detail.

References

[1] Auerbach, A. J. and Y. Gorodnichenko (2012a): "Measuring the Output Responses to Fiscal Policy." *American Economic Journal: Economic Policy* 4(2), 1-27.

[2] Auerbach, A. J. and Y. Gorodnichenko (2012b): "Fiscal Multipliers in Recession and Expansion."A. Alesina and F. Giavazzi, eds. *Fiscal Policy after the Financial Crisis*, Chicago: University of Chicago Press, forthcoming.

[3] Bachmann, R. and E. R. Sims (2012): "Confidence and the Transmission of Government Spending Shocks." *Journal of Monetary Economics* 59, 235-249.

[4] Barro, R. J. and C. J. Redlick (2011): "Macroeconomic Effects from Government Purchases." *Quarterly Journal of Economics* 126(1), 51-102.

[5] Baum, A. and G. B. Koester. (2011): "The Impact of Fiscal Policy on Economic Activity over the Business Cycle- Evidence from a Threshold VAR." Deutsche Bundesbank Discussion Paper 03/2011.

[6] Blanchard, O., and R. Perotti (2002): "An Empirical Characterization of the Dynamic Effects of Changes in Government Spending and Taxes on Output." *The Quarterly Journal of Economics* 117(4), 329-1368.

[7] Candelon, B. and L. Lieb. (2012): "Fiscal Policy in Good and Bad Times." Working Paper.

[8] Chib, S. and I. Jeliazkov (2001): "Marginal Likelihood from the Metropolis-Hastings Output." *Journal of the American Statistical Association* 96, 270-281.

[9] Davig, T. and E. M. Leeper (2011): "Monetary-Fiscal Policy Interaction and Fiscal Stimulus." *European Economic Review* 55(2): 211-227.

[10] Eggertsson, G.B. and P. Krugman (2012): "Debt, Deleveraging, and the Liquidity Trap: A Fisher-Minsky-Koo Approach." *Quarterly Journal of Economics*, forthcoming.

[11] Fazzari, S. M., J. Morley and I. Panovska (2013): "State-Dependent Effects of Fiscal Policy." Working Paper.

[12] Inoue, A. and L. Kilian (2012): "Inference on Impulse Response Functions in Structural VAR Models." *Journal of Econometrics*, forthcoming.

[13] Kilian, L. and R.J. Vigfusson (2011): "Are the Responses of the U.S. Economy Asymmetric in Energy Price Increases and Decreases?" *Quantitative Economics* 2, 419-453.

[14] Kim, C-J. and C. R. Nelson (1999): "Has the US Economy Become More Stable? A Bayesian Approach Based on a Markov-Switching Model of the Business Cycle." *The Review of Economics and Statistics* 81(4), 608-616.

[15] Koop, G. M.H. Pesaran, and S.M. Potter (1996): "Impulse Response Analysis in Nonlinear Multivariate Models." *Journal of Econometrics* 74, 119-147.

[16] Jorda, O. (2005): ""Estimation and Inference of Impulse Responses by Local Projections." *American Economic Review* 95 (1), 161-82.

[17] Leeper, E.M., N. Traum, and T.B. Walker (2011): "Clearing up the Fiscal Multiplier Morass." Working Paper.

[18] Lo, C. M. and J. Morley (2013):"Bayesian Analysis of Nonlinear Exchange Rate Dynamics and the Purchasing Power Parity Persistence Puzzle. " Working Paper, University of New South Wales.

[19] McConnell, M. M. and G. Perez-Quiros (2000): "Output Fluctuations in the United States: What Has Changed Since the Early 1980's? " *American Economic Review* 90(5), 1464-1676.

[20] Mittnik, S. and W. Semmler (2012): "Regime Dependence of the Fiscal Multiplier." *Journal of Economic Behavior & Organization* 83 (3), 502–522.

[21] Morley, J. and J. Piger (2012): "The Asymmetric Business Cycle." *The Review of Economics and Statistics* 91(1), 208-221.

[22] Mountford, A. and H. Uhlig (2009): "What Are the Effects of Fiscal Policy Shocks?" *Journal of Applied Econometrics* 26, 960-992.

[23] Owyang, M.T. and S. Zubairy (2013): "Who Benefits from Increased Government Spending? A State-Level Analysis." Regional *Science and Urban Economics*, forthcoming.

[24] Owyang, M.T., V. A. Ramey and S. Zubairy (2013): "Are Government Spending Multipliers Greater During Periods of Slack? Evidence from 20th Century Historical Data." Working Paper.

[25] Parker, J. A. (2011): "On Measuring the Effects of Fiscal Policy in Recessions." *Journal of Economic Literature* 49, 703-718.

[26] Perotti, R. (2008): "In Search of the Transmission Mechanism of Fiscal Policy." *NBER Macroeconomic Annual* 2007, ed. by D. Agemogly, K. Rogoff, and M. Woodford, pp. 169-226. University of Chicago Press, Chicago, Illinois.

[27] Ramey, V. A. (2011a): "Can Government Purchases Stimulate the Economy?" *Journal of Economic Literature* 126(1).

[28] Ramey, V. A. (2011b): "Identifying Government Spending Shocks: It's All in the Timing." *The Quarterly Journal of Economics.*

[29] Ramey, V. A. (2012) "Government Spending and Private Activity." A. Alesina and F. Giavazzi, eds. *Fiscal Policy after the Financial Crisis*, Chicago: University of Chicago Press, forthcoming.

[30] Ramey, V. A. and M. Shapiro (1998): "Costly Capital Reallocation and the Effects of Government Spending." *Carnegie Rochester Conference on Public Policy* 48, 145-194.

[31] Reinhart, C.M. and K.S. Rogoff (2009): "This Time Is Different: Eight Centuries of Financial Folly." Princeton University Press.

[32] Shoag, D. (2013): "The Impact of Government Spending Shocks: Evidence of the Multiplier from State Pension Plan Resturns." Harvard University Working Paper.

[33] Tong, H. (1978): "On a Threshold Model in Pattern Recognition and Signal Processing." C. H. Chen, Amsterdam: Sighoff and Nordhoff.

[34] Tong, H. (1983): "Threshold Models in Non-linear Time Series Analysis." New York: Springer-Verlag.

A. Appendix for Chapter 1

A.1. State-Space Representation for the Unobserved Components Model

The measurement equation for the model given by equations (1)-(11) is

$$\begin{bmatrix} y_t \\ s_t \\ e_t \\ h_t \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} * \beta_t \tag{18}$$

$\beta_t = [\tau_t \ \kappa_t \ c_{yt} \ c_{yt-1} \ c_{st} \ c_{st-1} \ c_{et} \ c_{et-1} \ c_{ht} \ c_{ht-1}]$ and the transition equation is given by

$$\beta_t = \mu_t + F\beta_{t-1} + u_t \tag{19}$$

where $\mu = [\mu_1 \ \mu_2 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ \mu_3 \ 0]'$ and $u_t = [\eta_t \ v_t \ u_{yt} \ 0 \ u_{st} \ 0 \ u_{et} \ 0 \ u_{ht} \ 0]'$ where u_{xt} is the linear combination of structural shocks corresponding to the right-hand of the equation that defines behavior of the cyclical component of variable x, given by equations (8)-(11). The transition matrix F is given by

$$F = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \phi_{y1} & \phi_{y2} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \phi_{s1} & \phi_{s2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \phi_{e1} & \phi_{e2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \phi_{h1} & \phi_{h2} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}. \tag{20}$$

When hours per capita are assumed to have a time-varying mean, the measurement equation for the estimated model is given by

$$\begin{bmatrix} y_t \\ s_t \\ e_t \\ h_t \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} * \beta_t \tag{21}$$

where $\beta_t = [\tau_t \; \kappa_t \; \mu_{ht} \; c_{yt} \; c_{yt-1} \; c_{st} \; c_{st-1} \; c_{et} \; c_{et-1} \; c_{ht} \; c_{ht-1}]$ and the transition equation is given by

$$\beta_t = \mu + F\beta_{t-1} + u_t \tag{22}$$

where $\mu = [\mu_1 \; \mu_2 \; 0 \; 0 \; 0 \; 0 \; 0 \; 0 \; 0 \; 0 \; 0]'$ and $u_t = [\eta_t \; v_t \; \epsilon_{ht} \; u_{yt} \; 0 \; u_{st} \; 0 \; u_{et} \; 0 \; u_{ht} \; 0]'$ where u_{xt} is the linear combination of structural shocks corresponding to the right-hand of the equation that defines behavior of the cyclical component of variable x, given by equations (8)-(11). The transition matrix F is given by

$$F = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \phi_{y1} & \phi_{y2} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \phi_{s1} & \phi_{s2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \phi_{e1} & \phi_{e2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \phi_{h1} & \phi_{h2} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}. \tag{23}$$

The state-space representation is identical for each subsample. The variance-covariance matrix Q is equal to $E[u_t u_t^{'}]$. For the sake of brevity, the equations that describe each element of Q as a function of the variances of the structural shocks and the impact coefficients is omitted, but it is available upon request from the author.

Model and State-Space Representation for Services

Since the services sector does not typically hold physical inventories (at least at the quarterly level), the model is adjusted to reflect that output and sales are equal within a quarter. The new model is given by the following equations:

$$y_t = \tau_t + c_y \tag{24}$$

$$e_t = \zeta_t + c_e \tag{25}$$

$$h_t = \mu + c_h \tag{26}$$

where y_t is hundred times the logarithm of real GDP, e_t is hundred times the logarithm of employment, and h_t is the logarithm of aggregate hours per employee. The stochastic trends in employment and output and the permanent shocks have the same interpretation as in the baseline model. The cyclical components are assumed to be stationary, and their dynamics can be described by the following equations:

$$\Phi_1(L)(y_t - \tau_t) = \lambda_{y\tau}\eta_t + \epsilon_y. \tag{27}$$

$$\Phi_3(L)(e_t - \tau_t - \kappa_t) = \lambda_{e\tau}\eta_t + \lambda_{ey}\epsilon_y + \epsilon_e \tag{28}$$

$$\Phi_4(L)(h_t - \mu_3) = \lambda_{h\tau}\eta_t + \lambda_{hy}\epsilon_y + \lambda_{he}\epsilon_e + \epsilon_h \tag{29}$$

where the impact coefficients are also defined as in the baseline model.

The state-space representation for the trivariate model is given by:

$$\begin{bmatrix} y_t \\ e_t \\ h_t \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} * \beta_t \tag{30}$$

$\beta_t = [\tau_t \; \kappa_t \; c_{yt} \; c_{yt-1} \; c_{et} \; c_{et-1} \; c_{ht} \; c_{ht-1}]$ and the transition equation is given by

$$\beta_t = \mu + F\beta_{t-1} + u_t \tag{31}$$

where $\mu = [\mu_1 \; \mu_2 \; 0 \; 0 \; 0 \; 0 \; 0 \; 0 \; 0]'$ and $u_t = [\eta_t \; v_t \; u_{yt} \; 0 \; u_{et} \; 0 \; u_{ht} \; 0]'$ where u_{xt} is the linear combination of structural shocks corresponding to the right-hand of the equation that defines behavior of the cyclical component of variable x, given by equations (21)-(23). The transition matrix F is given by

$$F = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \phi_{y1} & \phi_{y2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \phi_{e1} & \phi_{e2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \phi_{h1} & \phi_{h2} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}. \tag{32}$$

A.2. Cointegration

Figure 8 illustrates the motivation for modeling output and sales as cointegrated series. Both output and sales are nonstationary, which can be easily confirmed by pre-testing using standard unit root or stationarity tests. Both the first differences and the series are stationary. Indeed, as shown in Table 9, I reject the null of no cointegration if I assume that the cointegrating vector is $[1 \; -1]$.

Estimating the cointegrating vector leads to estimates that are very close to $[1 \; -1]$, and again I reject the null of no cointegration, implying that output and sales share the same stochastic trend.

A.3. Bayesian Estimation

The results presented in the text are obtained using a multi-block Metropolis-Hastings algorithm with a tailored proposal distribution. In order to ensure identification, I restrict the sign of $\lambda_{y\eta}$ and $\lambda_{s\eta}$. Assuming that $\lambda_{y\eta}$ and $\lambda_{s\eta} < 0$ is fairly innocuous in this context and is in line with previous studies, as discussed below. For convenience, the prior for $\lambda_{y\eta}$ and $\lambda_{s\eta}$ were truncated normal distributions on $(-\infty, 0)$ with mean -0.5 and variance equal to 1, and the prior for λ_{ys} was a truncated normal distribution on $(0, \infty)$ with mean 1 and variance 1. These priors are based on the MLE estimates and on results of previous studies (for example, Morley et al., 2003, Sinclair 2009, Morley and Singh, 2012, or Basistha, 2009). Restricting λ_{ys} to be positive simply means that output responds positively to positive sales shocks. This restriction is based on the MLE estimates and is consistent both with theory and with estimated responses from a baseline VECM model. It is important to note that the truncated priors are merely a convenient tool to ensure slightly faster convergence- the results were robust to the choice of priors, and using more diffuse priors or different families of priors does not affect the results significantly. The only restriction that is needed to identify the coefficients is the sign restrictions.

The priors for the initial values for the stochastic trends were Gaussian distributions that were centered at the initial observations and had variance equal to 10. The MH algorithm was implemented as follows:

1. Start with arbitrary values for the parameter coefficients.

2. Conditional on the parameter coefficients, obtain the smoothed estimates for the state variables.

3. At the i^{th} iteration, conditional on the parameter vector $\theta^{(i)}$ and the state variables, draw a new value for the J-dimensional parameter block θ_b from a Student-t proposal

with mean $\theta_b^{(i)}$, scale equal to the J-by-J subblock of the inverse of the Hessian matrix evaluated at $\theta^{(i)}$ and 15 degrees of freedom. If the value is accepted, update θ^i to reflect the update when calculating the Hessian for the other blocks.

4. Repeat step 2 until all parameters are updated, update θ^{i+1}

The results presented in the text were obtained using 80,000 iterations of the MH chain, after a burn-in of 20,000 draws. To ensure convergence, the chain was started from several different values.In order to obtain the trend and cycle estimates, I use the Kalman filter and the related prediction-error decomposition.

A.4. Tables

Table 1: Mean and standard deviations of growth rates by subsample

	Average		Recoveries	
	1948Q1- 1983Q4	1984q1-2012Q1	1948Q1- 1983Q4	1984q1-2012Q1
Output	0.87 (1.13)	0.67 (0.64)	1.60 (0.81)	0.65 (0.31)
Sales	0.87 (0.89)	0.67 (0.60)	1.17 (0.54)	0.37 (0.46)
Employment	0.50 (0.79)	0.33 (0.49)	0.70 (0.49)	−0.14 (0.31)
Aggregate Hours	0.40 (1.42)	0.29 (0.62)	1.01 (1.00)	−0.05 (0.47)
Hours per Employee	−0.10 (0.74)	−0.03 (0.26)	0.31 (0.61)	0.08 (0.25)

Table 2: Posterior medians for key parameters (standard deviations in parentheses)

	1948q1-1983q4	1984q1-2011q4	Change in Parameter
$\phi_{11} + \phi_{12}$	0.45(0.11)	0.62 (0.16)	0.16 (0.13)
$\phi_{21} + \phi_{22}$	0.66 (0.15)	0.59 (0.11)	-0.06 (0.09)
$\phi_{31} + \phi_{33}$	0.60 (0.12)	0.69 (0.18)	0.10 (0.25)
$\phi_{41} + \phi_{42}$	0.85 (0.17)	0.89 (0.18)	0.05 (0.20)
σ_η	1.72 (0.22)	1.29 (0.02)	-0.42 (0.21)
σ_y	0.49 (0.14)	0.28 (0.03)	-0.21 (0.09)
σ_s	0.34 (0.10)	0.29 (0.08)	-0.06 (0.11)
σ_e	0.21 (0.09)	0.12 (0.01)	-0.10 (0.05)
σ_h	0.53 (0.09)	0.24 (0.09)	-0.30 (0.18)
$\lambda_{e\eta}$	-0.14 (0.08)	-0.22 (0.01)	-0.08 (0.06)
$\lambda_{h\eta}$	-0.18 (0.04)	-0.25 (0.02)	-0.07 (0.05)
λ_{es}	0.35 (0.06)	0.87 (0.03)	0.52 (0.07)
λ_{hs}	0.27 (0.17)	0.22 (0.07)	0.06 (0.05)
λ_{ey}	0.11 (0.01)	0.06 (0.01)	-0.05 (0.02)
λ_{hy}	0.20 (0.02)	0.33 (0.02)	0.14 (0.07)
λ_{he}	0.68 (0.21)	-0.13 (0.11)	-0.81 (0.16)

Table 3: Volatility and impact coefficients (standard deviations in parentheses)

	1948q1-1983q4	1984q1-2011q4	1984q1-2007q3
σ_η	1.716 (0.221)	1.288 (0.105)	0.914 (0.101)
σ_v	0.741 (0.144)	0.739 (0.051)	0.678 (0.153)
σ_y	0.491 (0.141)	0.282 (0.029)	0.196 (0.072)
σ_s	0.338 (0.092)	0.287 (0.083)	0.250 (0.132)
σ_e	0.210 (0.082)	0.119 (0.009)	0.090 (0.020)
σ_h	0.527 (0.092)	0.239 (0.092)	0.223 (0.025)
$\rho_{\eta\nu}$	-0.752 (0.130)	-0.689 (0.125)	-0.835 (0.141)
$\lambda_{y\eta}$	-0.785 (0.120)	-0.720 (0.075)	-0.773 (0.098)
$\lambda_{s\eta}$	-0.834 (0.152)	-0.717(0.082)	-0.768 (0.112)
$\lambda_{e\eta}$	-0.136 (0.081)	-0.221 (0.014)	-0.218 (0.099)
$\lambda_{h\eta}$	-0.184 (0.036)	-0.254 (0.015)	-0.235 (0.052)
λ_{ys}	0.325 (0.027)	0.257 (0.095)	0.645 (0.152)
λ_{es}	0.349 (0.061)	0.872 (0.028)	0.783 (0.114)
λ_{hs}	0.269 (0.167)	0.218 (0.072)	0.242 (0.098)
λ_{ey}	0.113 (0.007)	0.059 (0.006)	0.211 (0.054)
λ_{hy}	0.198 (0.017)	0.326 (0.022)	0.321 (0.093)
λ_{he}	0.683 (0.214)	-0.126 (0.112)	-0.150 (0.055)

Table 4: Sums of the estimated AR coefficients: Manufacturing (standard deviations in parentheses)

	Total Manufacturing		Durables		Nondurables	
	1964q1-1983q4	1984q1-2011q2	1964q1-1983q4	1984q1-2011q2	1964q1-1983q4	1984q1-2011q2
$\phi_{y1}+\phi_{y2}$	0.788 (0.120)	0.777 (0.045)	0.851 (0.137)	0.647 (0.081)	0.746 (0.217)	0.628 (0.087)
$\phi_{s1}+\phi_{s2}$	0.811 (0.024)	0.774 (0.132)	0.739 (0.057)	0.660 (0.078)	0.814 (0.132)	0.724 (0.092)
$\phi_{e1}+\phi_{e2}$	0.896 (0.012)	0.895 (0.120)	0.912 (0.019)	0.748 (0.170)	0.901 (0.093)	0.897 (0.074)
$\phi_{h1}+\phi_{h2}$	0.781 (0.114)	0.899 (0.120)	0.804 (0.117)	0.899 (0.182)	0.853 (0.098)	0.827 (0.105)

Table 5: Estimated variances and impact coefficients: Manufacturing (standard deviations in parentheses)

	Manufacturing- aggregate		Durables		Nondurables	
	1964q1-1983q4	1984q1-2011q2	1964q1-1983q4	1984q1-2011q2	1964q1-1983q4	1984q1-2011q2
σ_η	6.369 (0.323)	1.344 (0.182)	3.595 (0.324)	1.432 (0.235)	2.302 (0.253)	1.439 (0.128)
σ_v	6.263 (2.257)	2.512 (0.619)	3.411 (0.521)	3.042 (0.133)	2.218 (0.416)	1.868 (0.425)
σ_y	1.821 (0.421)	0.834 (0.594)	2.669 (0.262)	1.103 (0.235)	1.480 (0.209)	1.181 (0.056)
σ_s	1.505 (0.122)	0.796 (0.082)	2.339 (0.116)	1.182 (0.210)	1.101 (0.210)	0.984 (0.112)
σ_e	0.801 (0.126)	1.055 (0.221)	0.953 (0.313)	0.454 (0.128)	0.512 (0.043)	0.376 (0.42)
σ_h	0.617 (0.111)	0.414 (0.062)	0.706 (0.073)	0.502 (0.027)	0.841 (0.117)	0.382 (0.057)
$\rho_{\eta\nu}$	-0.798 (0.211)	-0.687 (0.233)	-0.693 (0.078)	-0.621 (0.221)	-0.855 (0.025)	-0.532 (0.115)
$\lambda_{y\eta}$	-0.731 (0.124)	-0.827 (0.065)	-0.998 (0.026)	-0.557 (0.080)	-0.636 (0.149)	-0.481 (0.117)
$\lambda_{s\eta}$	-0.678 (0.012)	-0.657 (0.032)	-0.651 (0.183)	-0.759 (0.012)	-0.786 (0.062)	-0.930 (0.042)
$\lambda_{e\eta}$	-0.217 (0.042)	-0.238 (0.051)	-0.262 (0.082)	-0.198 (0.023)	-0.166 (0.024)	-0.188 (0.050)
$\lambda_{h\eta}$	-0.265 (0.010)	-0.117 (0.013)	-0.279 (0.110)	-0.084 (0.052)	-0.127 (0.120)	-0.214 (0.162)
λ_{ys}	0.899 (0.014)	0.771 (0.079)	0.668 (0.160)	0.763 (0.133)	0.396 (0.025)	0.768 (0.104)
λ_{es}	0.159 (0.011)	0.653 (0.068)	0.04 (0.267)	0.232 (0.056)	0.450 (0.063)	1.811 (0.222)
λ_{hs}	0.210 (0.320)	0.228 (0.117)	0.123 (0.073)	0.218 (0.155)	0.261 (0.052)	0.221 (0.083)
λ_{ey}	0.167 (0.025)	0.101 (0.131)	0.358 (0.282)	0.087 (0.102)	0.728 (0.054)	0.219 (0.052)
λ_{hy}	0.221 (0.104)	0.259 (0.014)	0.394 (0.034)	0.418 (0.145)	0.450 (0.051)	0.590 (0.021)
λ_{he}	0.158 (0.041)	-0.302 (0.174)	-0.104 (0.234)	-0.257 (0.066)	0.034 (0.100)	-0.326 (0.071)

Table 6: Sums of the estimated AR coefficients: Services and FIRE (standard deviations in parentheses)

	Services		FIRE	
	1964q1-1983q4	1984q1-2011q2	1964q1-1983q4	1984q1-2011q2
$\phi_{y1}+\phi_{y2}$	0.700 (0.118)	0.801 (0.199)	0.953 (0.069)	0.843 (0.164)
$\phi_{e1}+\phi_{e2}$	0.782 (0.155)	0.889 (0.073)	0.836 (0.087)	0.987 (0.018)
$\phi_{h1}+\phi_{h2}$	0.902 (0.058)	0.862 (0.099)	0.873 (0.143)	0.755 (0.156)

Table 7: Estimated variances and impact coefficients: services and FIRE (standard deviations in parentheses)

	Services		FIRE	
	1964q1-1983q4	1984q1-2011q2	1964q1-1983q4	1984q1-2011q2
σ_η	1.243 (0.156)	1.217 (0.234)	2.451 (0.021)	2.136 (0.354)
σ_v	2.073 (0.560)	3.487 (0.412)	1.893 (0.245)	1.831 (0.052)
σ_y	1.688 (0.043)	1.253 (0.229)	1.584 (0.098)	0.978 (0.137)
σ_e	0.798 (0.062)	0.906 (0.305)	0.678 (0.125)	0.304 (0.061)
σ_h	0.204 (0.052)	0.309 (0.140)	0.706 (0.073)	1.038 (0.119)
$\rho_{\eta\nu}$	-0.750 (0.300)	-0.319 (0.210)	-0.836 (0.144)	-0.362 (0.115)
$\lambda_{y\eta}$	-0.869 (0.128)	-0.958 (0.062)	-0.552 (0.166)	-0.725 (0.215)
$\lambda_{e\eta}$	-0.0473 (0.102)	-0.185 (0.158)	-0.137 (0.096)	-0.115(0.087)
$\lambda_{h\eta}$	-0.176 (0.113)	-0.241 (0.052)	-0.034 (0.007)	-0.165 (0.058)
λ_{ey}	0.379 (0.005)	0.121 (0.014)	0.363 (0.125)	0.067 (0.032)
λ_{hy}	0.286 (0.041)	0.450 (0.056)	0.343 (0.006)	0.449 (0.095)
λ_{he}	0.278 (0.102)	-0.048 (0.031)	0.321 (0.064)	-0.165 (0.005)

Table 8: Priors for the Bayesian Estimation (UC Model)

	Family	Mean	Variance
$(\phi_{i1}\ \phi_{i2})$	Truncated Normal, to ensure stationarity	(0 0)	I_2
σ_X^2	Inverse Gamma	1	5
$\lambda_{y\tau}$	Truncated Normal	-0.5	1
$\lambda_{s\tau}$	Truncated Normal	-0.5	1
λ_{ys}	Truncated Normal	1	1
$\lambda_{i,,j_{i\neq y,s}}$	Normal	0	10
$\mu_{1,2}$[29]	Normal	0	10

Table 9: Testing for Cointegration Between Output and Sales

	Normalized Cointegrating Coefficient on Sales	ADF, Residual Series
OLS	-1.0004 (0.0006)	-7.953 (<0.001)
Imposing $[1\ -1]$ as cointegrating vector	-	-8.08 (<0.001)

A.5. Figures

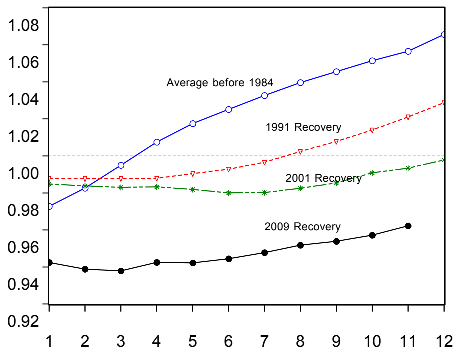

Figure 1: Behavior of payroll employment the 12 quarters following the NBER trough (1=Level at the NBER Peak)

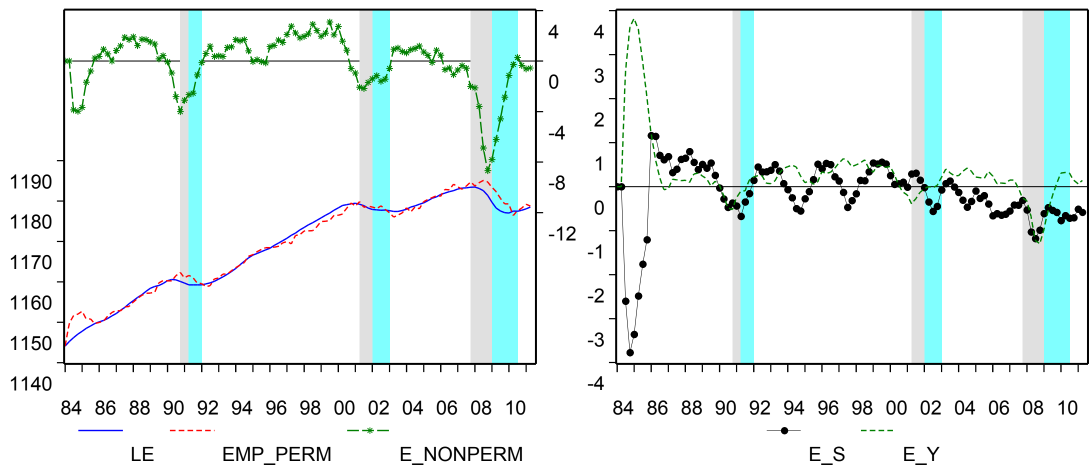

Figure 2: Decomposition of the path of employment by type of shock

Notes: Left panel: observed path of log employment (blue) versus movements causes by permanent shocks (red) and movements caused by purely transitory shocks (green). The right panel plots the decomposition of the cyclical part of employment: fluctuations due to sales shocks (dashed green line) and fluctuations due to non-sales output shocks.NBER recessions are shaded in gray, the blue shading indicates the four quarters following the NBER trough.

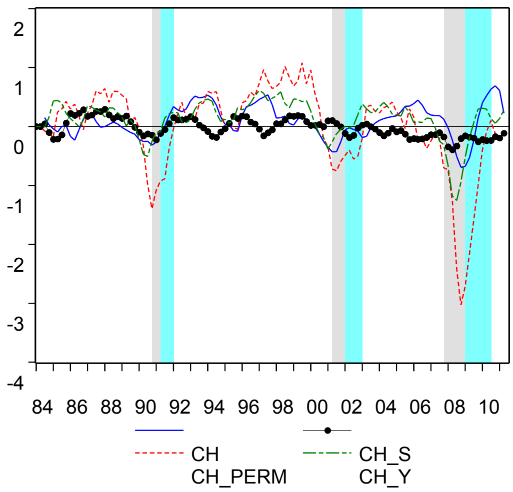

Figure 3: Decomposition of the path of mean-adjusted hours per employee by type of shock

Notes: Observed path of mean adjusted hours per employee (blue line) versus movements causes by permanent shocks (dashed red line), movements caused by purely transitory sales shocks (patterned black line), and movements caused by transitory non-sales output shocks (dashed green line).

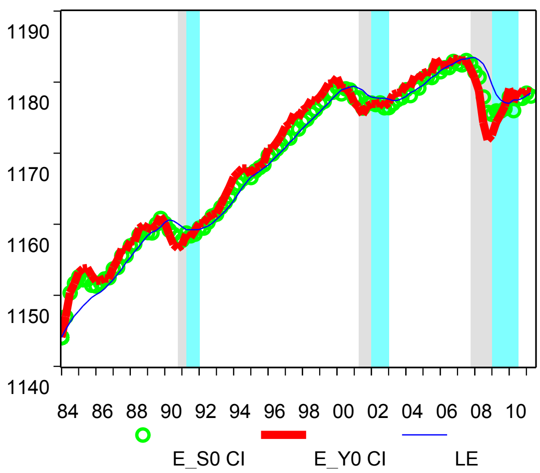

Figure 4: Counterfactual paths for employment.

Notes: Observed path of log employment (blue line), 90% credibility interval for the counterfactual path of employment when the sensitivity coefficient λ_{es} is drawn from the pre-1984 posterior distribution (green), and 90% credibility interval for the counterfactual path of employment when the sensitivity coefficient λ_{ey} is drawn from the pre-1984 posterior distribution (red).

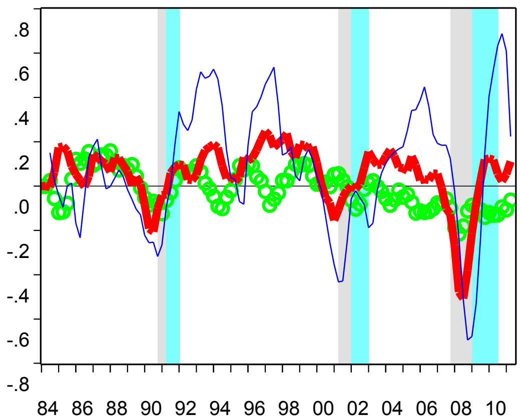

Figure 5: Counterfactual paths for hours per employee (mean-adjusted)

Notes: Observed path of mean-adjusted hours per employee (blue line), 90% credibility interval for the counterfactual path of mean-adjusted hours per employee when the sensitivity coefficient λ_{hs} is drawn from the pre-1984 posterior distribution (green), and 90% credibility interval for the counterfactual path of employment when the sensitivity coefficient λ_{hy} is drawn from the pre-1984 posterior distribution (red).

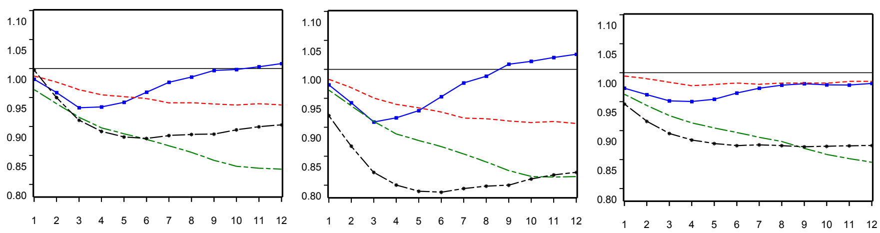

Figure 6: Employment in the manufacturing sector after the NBER trough(1=Level at the NBER Peak)

Notes: Left: total manufacturing, middle: durables, right: nondurables. Blue: average before 1984, red:recovery from 1990 recession, green: recovery from the 2001 recession, black: recovery from the Great Recession. Observed path of log employment (blue line), 90% credibility interval for the counterfactual path of employment when the sensitivity coefficient λ_{es} is drawn from the pre-1984 posterior distribution (green), and 90% credibility interval for the counterfactual path of employment when the sensitivity coefficient λ_{ey} is drawn from the pre-1984 posterior distribution (red).

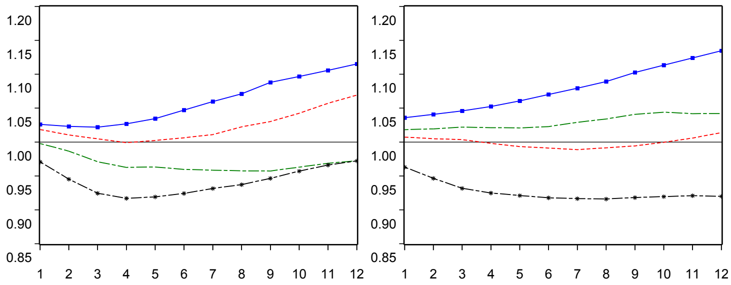

Figure 7: Employment in the services and FIRE sector after the NBER trough(1=Level at the NBER Peak).

Notes: Left: service, right: FIRE. Blue: average before 1984, red:recovery from 1990 recession, green: recovery from the 2001 recession, black: recovery from the Great Recession.

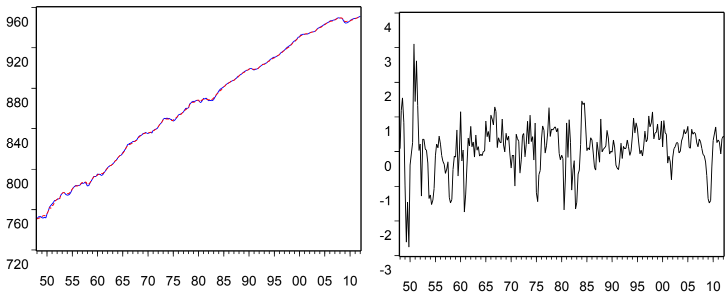

Figure 8: Cointegration: real GDP and real sales. Left: Real GDP (blue), real final sales (red). Right: $d = y - s$. The sample size is 1948q1:2012Q1.

B. Appendix for Chapter 2

B.1. Bayesian Estimation

For the baseline linear model, I assume that the prior for the VAR parameters is multivariate normal, the prior for the variance matrix is an inverse Wishart distribution, and the prior for the scale parameter λ is a Gamma distribution. Under these assumptions, I can sample directly using a Gibbs step. Specifically, recall that the linear model is given by (1):

$$Y_t = \Phi_0 + \Phi(L)Y_{t-1} + \lambda_t \varepsilon_t$$

where $\Phi(L)$ is an autoregressive matrix polynomial with roots strictly outside the unit circle, t = 1 for t = 1967q1,...,1983q4 and equal to λ for t = 1984q1,... 2010q4, and ε_t is iid Gaussian random variable with mean 0 and variance- covariance matrix Ω that does not change over time. Then, letting $\Phi = vec(\Phi_0)|vec(\Phi_1)...|vec(\Phi_j)$, I assume that the prior for Φ is a normal distribution, truncated to the stationarity region, with mean equal to 0, and variance-covariance matrix equal to V_n. The scaling parameter λ is assumed to have a gamma prior with parameters α and β, and I impose an inverted Wishart prior with $v_0 = 25$ degrees of freedom and a scale matrix R_0. For brevity, let $rhs_t = [1\, y_{t-1,1} \ldots y_{t-1,k} \ldots y_{t-p,k}]$ and

$$x_t = \begin{bmatrix} rhs_t & 0 & 0 & 0 \\ 0 & rhs_t & 0 & 0 \\ 0 & 0 & rhs_t & 0 \\ 0 & 0 & 0 & rhs_t \end{bmatrix}.$$

It is straightforward to see that $\Phi|\Omega, \lambda, y$ is Gaussian with variance

$$V = (V_n^{-1} + \sum_{t=p+1}^{T} x_t'(\lambda_t\Omega)^{-1}x_t)^{-1}$$

and mean

$$\mu = V^{-1}(\sum_{t=p+1}^{T} x_t'(\lambda_t \Omega)^{-1} y_t).$$

Similarly, $\Omega|\Phi, \lambda, y \sim IW(\nu_1, R_1)$, where $\nu_1 = \nu_0 + T - 4$ and $R_1 = [R_0^{-1} + \sum_{t=5}^{T}(y_t - x_t\Phi)'\lambda_t^{-1}(y_t - x_t\Phi)]^{-1}$. The inverse Wishart distribution is a standard distribution, so I can sample Ω conditional on the other parameters directly. Conditional on the other parameters and the data, λ has a gamma distribution with parameters $\alpha_1 = \alpha + T_p$ and $\beta_1 = \beta + 0.5\sum_{t=t_1}^{T}(y_t - x_t\Phi)'\Omega^{-1}(y_t - x_t\Phi)$ where $t_1 = 1984Q1$, and $T_p = 108$ (the number of quarters from 1984Q1 to 2010Q4).

The threshold model is given by

$$Y_t = \Phi_0^1 + \Phi_1(L)Y_{t-1} + (\Phi_0^2 + \Phi_2(L)Y_{t-1})I[q_t \leq \gamma] + \lambda_t \varepsilon_t$$

where q_t is the threshold variable, γ is the threshold around which the dynamics of the model changes, and the other variables are defined similarly to the variables in the linear model. Let $rhs_{1t} = [rhs_t \ rsh_t * I[q_t \leq \gamma]$ and assume that the prior for $\Phi = vec(\Phi_0^1)|vec(\Phi_1)|vec(\Phi_2^0)|vec(\Phi_2)$ is normal with mean 0 and variance 0, truncated so that $\Phi_1(L)$ and $\tilde{\Phi}(L) = \Phi_1(L) + \Phi_2(L)$ have roots strictly outside the unit circle (i.e. so that the VAR is stationary in each regime).

Similar to the linear case, it is straightforward to show that $\Phi|\Omega, \lambda, \gamma, y$ is Gaussian with variance $V = (V_n^{-1} + \sum_{t=p+1}^{T} x_{1t}'(\lambda_t\Omega)^{-1}x_{1t})^{-1}$and mean $\mu = V^{-1}(\sum_{t=p+1}^{T} x_{1t}'(\lambda_t\Omega)^{-1}y_t)$, where x_{1t} is defined the same way as x_t,except I use the vector rhs_{1t} is used in place of rhs_t. Likewise, the conditional distribution of Ω is inverse Wishart, and the conditional posterior distribution of λ is gamma, and I can sample from these posteriors using a Gibbs step. The conditional distribution for γ is nonstandard, and it has to be sampled using an MH step.

Following the standard approach in the literature, the proposal density is Student$-t$ with 15 degrees of freedom. To obtain the mode for the proposal distribution for the first draw, I use concentrated maximum likelihood and grid search over the middle 70% of the threshold variable in order to obtain the posterior mode of the parameter γ. Because ε_t

is assumed to be Gaussian, the ML estimators can be obtained by using least squares estimation. For this maximization γ is restricted to a bounded set $\Gamma = [\underline{\gamma}, \overline{\gamma}]$ that covered the middle 70% of the threshold variable.

Conditional on γ and the threshold variable, the model is linear in Φ and Ω. Estimating the linear model by splitting the sample into two subsamples yields the conditional estimators $\widehat{\Phi}$ and $\widehat{\Omega}$. The estimated threshold value (conditional on the threshold variable and the delay lag) can be identified uniquely as

$$\widehat{\gamma} = \arg\max_{\gamma \in \Gamma_n} llik_n(\gamma | q, d) \tag{33}$$

where Γ is approximated by a grid search on $\Gamma_n = \Gamma \cap \{q_1, q_2, ..., q_n\}$. To ensure identification, the bottom and top 15% quantiles of the threshold variable are trimmed. I use the estimated value $\widehat{\gamma}$ for constructing the proposal for the first draw of the MH algorithm. Given a sufficiently large burn-in, the value of $\hat{\gamma}$ does not affect the Bayesian estimates, but it provides us with a plausible starting value for the mode and it enables us to easily compare the Bayesian mode with the maximum likelihood estimate.

Note, however, that the grid search makes it infeasible to obtain the variance of the estimate of γ. To address this issue, I use the approach proposed by Lo and Morley (2013). In particular, I obtain a measure of the curvature of the posterior with respect to γ by inverting the likelihood ratio statistics for the threshold parameters, based on the assumption that the parameter estimate is normally distributed and the LR statistics is $\chi^2(1)$. I use the 95% CI for the likelihood ratio statistics to obtain a corresponding standard error for γ. Again, I do not attempt to perform frequentist inference. This is simply a fast way of obtaining a plausible value for the curvature. It it important to note that this approach is only an approximation, and it not asymptotically accurate for obtaining standard errors of the ML estimate for γ, but it is much faster than bootstrapping and inverting the LR test in order to obtain the scale for the first draw of the MH algorithm. Because this approximation affects only the first draw, it will not affect the estimates of the parameters if the burn-in is large enough.

At the i^{th} iteration, the transition density for $\gamma^{(i+1)}$ is a Student-t distribution with mean equal to $\gamma^{(i)}$ and variance equal to $\kappa\hat{\sigma}_{\gamma}^{2}$, where $\hat{\sigma}_{\gamma}^{2}$ is obtained by inverting the LR test. The parameter κ is calibrated on the fly to ensure acceptance rate between 20 and 60%.

To ensure that the results are robust to the choice of priors, the model is estimated by using different hyperparameters for the priors, and by using different functional forms for the priors (when the priors are not conjugate to the posteriors, all parameters are drawn using a multi-block MH step). Also, to check for convergence for each combination of priors, I start the algorithm from different points, and we use a large burn-in for all runs of the MH algorithm. In particular, I use a burn-in sample of 20,000 draws and make inference based on an additional 50,000 MH iterations The results presented and discussed in the paper are based on the following priors:

	Type of Prior	Mean	Variance/ Scale
Φ	Multivariate Gaussian	0	$100 * I_k$
Ω	Inverse Wishart	$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$	$25 * \mu$
λ	Gamma	1	0.75
γ	Uniform	0.165	0.365^2

B.2. Generalized Impulse Response Function

The procedure for computing the generalized impulse response functions (GIRFs) follows Koop, Pesaran and Potter (1996). The generalized impulse response is defined as the effect of a one-time shock on the forecast of variables in the model, and the response is compared against a baseline “no shock” scenario.

$$GIRF_y(k, \varepsilon_t, \Psi_{t-1}) = E[Y_{t+k}|\varepsilon_t, \Psi_{t-1}] - E[Y_{t+k}|\Psi_{t-1}] \tag{34}$$

where k is the forecasting horizon, ε_t is the shock, and Ψ_{t-1} are the initial values of the variables in the model. The impulse response is then computed by simulating the model. The shock to government spending is normalized to be equal to 1% of GDP (at the time the shock occurs).

The $GIRF_y$ response for a given draw $\Theta^{(i)}$ of the MH algorithm is generated using the following steps:

1. Pick a history Ψ_{t-1}. The history is the actual value of the lagged endogenous variable at a particular date.

2. Pick a sequence of 4-dimensional shocks ε_{t+k}, $k = 0, 1, ..., 20$. The shocks are simulated assuming an independent Gaussian process with mean zero and variance-covariance matrix equal to $\lambda_t^{(i)} * \Omega^{(i)}$.

3. Using Ψ_{t-1}^r and ε_{t+k}, simulate the evolution of Y_{t+k} over $l+1$ periods. Denote the resulting path $Y_{t+k}(\varepsilon_{i0}, \Psi_{t-1}, \varepsilon_{t+k})$ for $k = 0, 1, ...l$.

4. Substitute ε_{i0} for ε_{t+k}, using the Cholesky decomposition of Ω_t to othogonalize the shocks. Simulate the evolution of Y_{t+k}over $l+1$ periods. Denote the resulting path $Y_{t+k}(\Psi_{t-1}, \varepsilon_{t+k})$ for $k = 0, 1, ..., l$.

5. Repeat steps 2 to 4 B times, with $B = 500$ to obtain a consistent estimate of the impulse response function conditional on the history.

6. To obtain the average response for a subset of histories, repeat steps 1-5 for a the subset of interest (I compute it for all low states, all high states, the rapid expansion of the late 1990s and the Great recession), and report the response averaged over all histories.

Because the impulse responses are nonlinear functions of the parameters, their distribution is nonstandard and it is not symmetric around the mean. In cases like this, reporting the median value is not adequate, as the median may not be a valid measure of central tendency, and the median impulse response may not correspond to a well-defined structural model. In order to circumvent this problem, I use the approach proposed by Inoue and Kilian (2012). For a given history, I evaluate the impulse response function for each draw of the MH algorithm, drawing the entire impulse response function for periods 1 through 20. Then I average over histories, and evaluate the posterior likelihood of the impulse response for that draw of the algorithm (averaged over the histories of interest). The impulse response function with the highest average posterior likelihood is then used for inference. To construct the $(1-\alpha)*100\%$ credibility interval, I order the posterior likelihood values, and include the impulse responses whose posterior likelihood was in the upper $(1-\alpha)*100$ percentile. This method results in a "credibility cloud" with a shot gun pattern because we draw entire impulse responses rather than responses for each individual point in time. For easy interpretation, only the outer points of the cloud are reported.

B.3. Tables

Table 10: Structural breaks in capacity utilization and unemployment

	Break Date(s)	F-statistics	p-value
cap	1974Q1	41.66	<0.001
un	1974Q3, 1981Q4, 1994Q4	32.34,16.43,6.21	<0.01 ,<0.01,0.03

Notes: The break dates were obtained using a sequential Quandt-Andrews test. The estimated break dates coincided with the break dates obtained using Bai-Perron's sequential procedure under the assumption that the mean is the only parameter that has a structural break.

Table 11: Marginal likelihood values and estimated thresholds: baseline model

Type of Switching Variable	q_{t-d}	ML	Threshold
Linear Model	-	-997.18	-
Lagged output growth and moving averages of output growth	Δy_{t-2}	-720.69	1.33 (0.12)
Output Gap	gap_{t-1}	-682.52	-0.594 (0.41)
Capacity Utilization, not adjusted for breaks (level and growth)	cap_{t-1}	-800.52	81.1 (1.42)
Capacity Utilization, adjusted for breaks (level and growth)	$\hat{cap}_{t-1}$	-673.69	-0.21 (0.37)
Unemployment, not adjusted for breaks (level and first differences)	un_{t-1}	-703.25	4.83 (0.33)
Unemployment, adjusted for breaks in mean	$\hat{u}_{t-2}$	-760.63	-0.26 (0.11)
Total Federal Debt Held By the Public as % of GDP	$ratio_{t-1}$	-1100.24	34.5 (1.3)
Debt Outstanding, Federal Government Sector, as % of GDP	$ratio2_{t-2}$	-1020.23	47.2 (1.15)
Real Short-Term Interest Rate (FFR-CPI)	$R_{real,t-2}$	-1004.42	2.10 (1.35)

Table 12: Marginal likelihoods and likelihood at the mode for the linear model and the TVAR using $\hat{cap}_{t-1}$ as the switching variable

	ML_{LIN}	lik_M	ML_{TVAR}	$lik_{m,TVAR}$
M1 (Y)	-997.18	-1461.28	-673.69	-1338.53
M2 (C)	-851.95	-1270.47	-576.91	-1209.77
M3(I)	-2011.98	-2561.17	-1473.50	-2216.87
M4 (X)	-6414.28	-6962.87	-4060.67	-4812.772
M5 (M)	-7011.89	-7519.62	-4311.90	-5062.58
M6 (UR)	-813.43	-1205.82	-549.33	-1197.53
M7 (E)	-802.77	-1190.28	-544.12	-1199.82
M8 (π)	-1563.58	-2064.27	-1156.22	-1995.35

Table 13: Estimated thresholds when mean-adjusted capacity utilization is the switching variable

	Posterior Mode for γ (Standard Deviation in Parentheses)
M1 (Y)	-0.21 (0.37)
M2 (C)	-0.54 (0.37)
M3(I)	-1.39 (1.32)
M4 (X)	-1.001 (0.12)
M5 (M)	-0.177 (0.37)
M6 (UR)	-0.461 (0.39)
M7 (E)	-0.5133 (0.45)
M8 (π)	-0.521 (0.35)

B.4. Figures

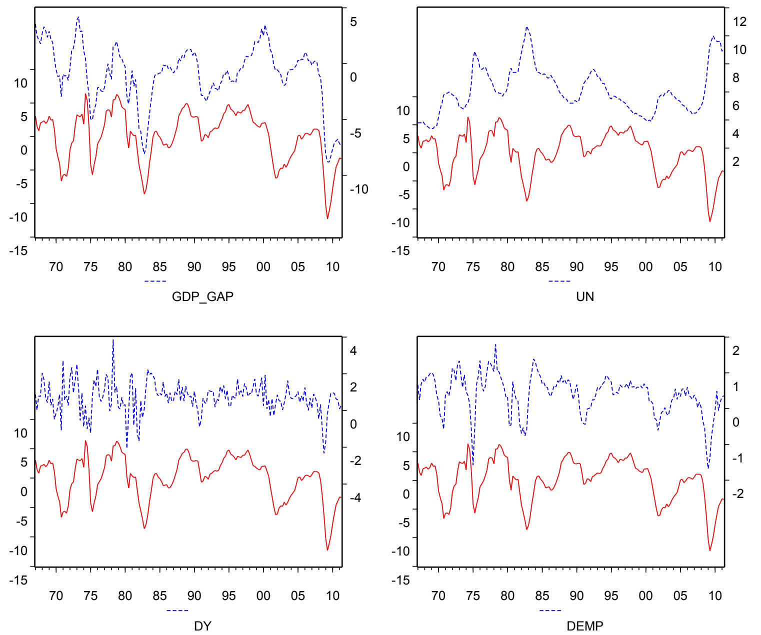

Figure 9: Mean-adjusted capacity utilization versus other measures of slack

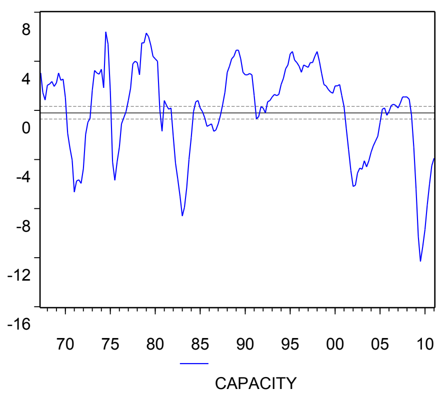

Figure 10: Mean-Adjusted Capacity Utilization and Estimated Threshold (with 90% credibility interval)

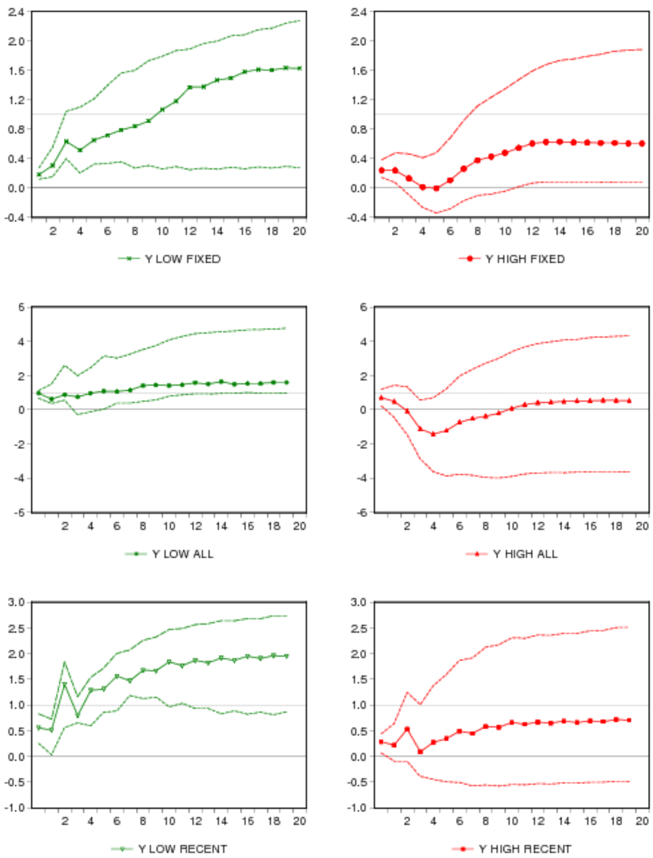

Figure 11: Responses of output to a government spending shock

Left: low initial state, right: high initial state. Top: fixed states, middle: evolving states, averages over all histories, bottom: evolving states, averages over recent histories (post 1984)

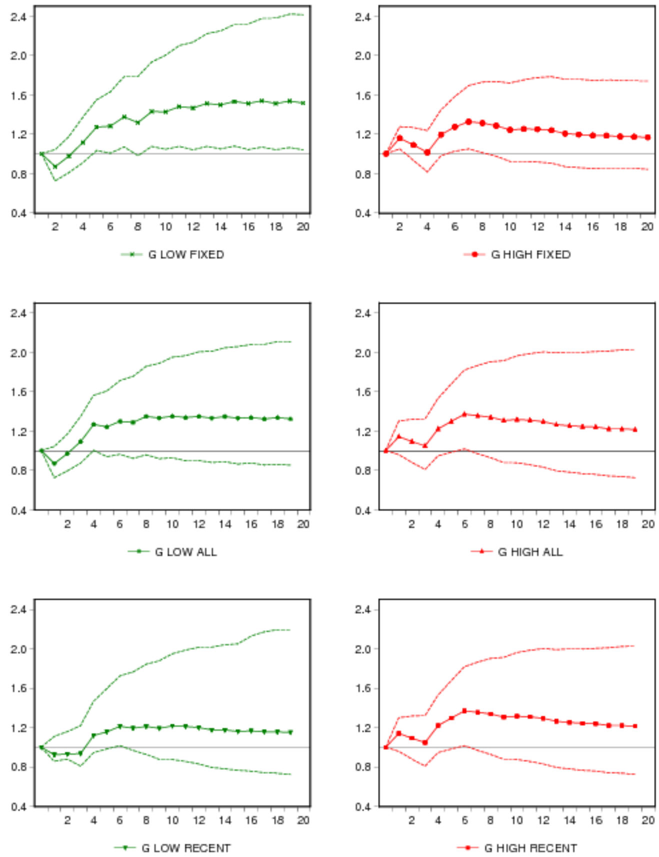

Figure 12: Responses of government spending to a government spending shock
Left: low initial state, right: high initial state. Top: fixed states, middle: evolving states, averages over all histories, bottom: evolving states, averages over recent histories (post 1984)

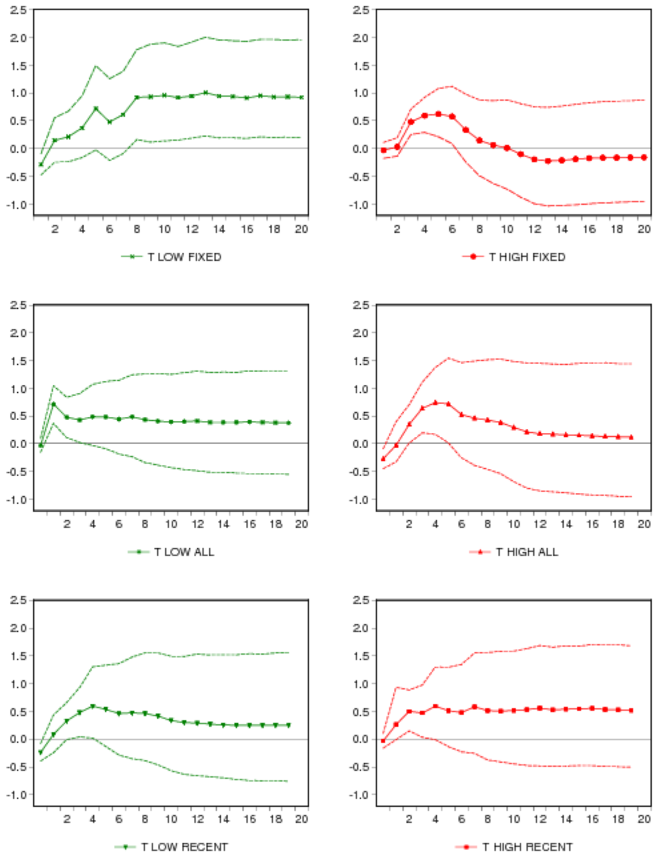

Figure 13: Responses of tax revenues to a government spending shock
Left: low initial state, right: high initial state. Top: fixed states, middle: evolving states, averages over all histories, bottom: evolving states, averages over recent histories (post 1984)

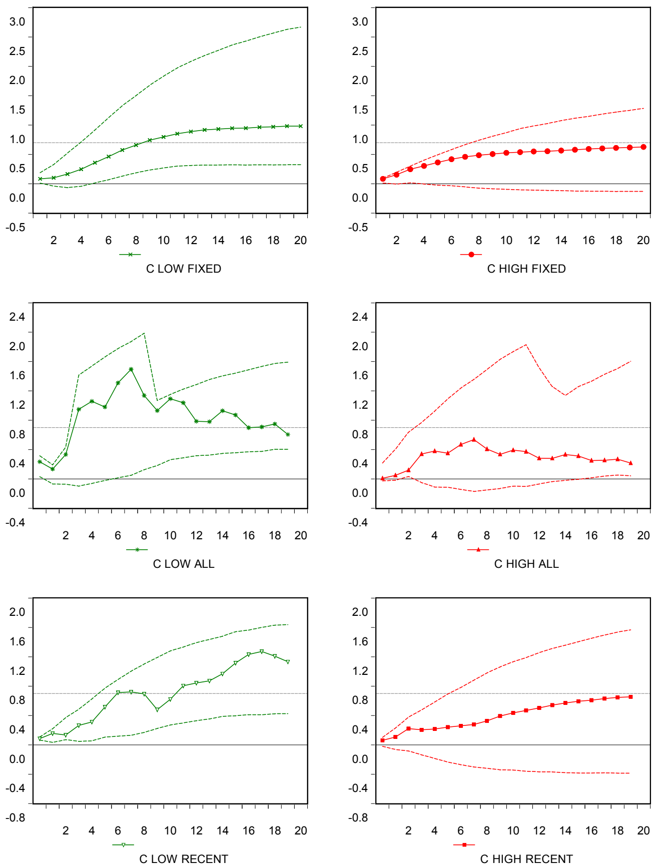

Figure 14: Responses of sonsumption to a government spending shock

Left: low initial state, right: high initial state. Top: fixed states, middle: evolving states, averages over all histories, bottom: evolving states, averages over recent histories (post 1984)

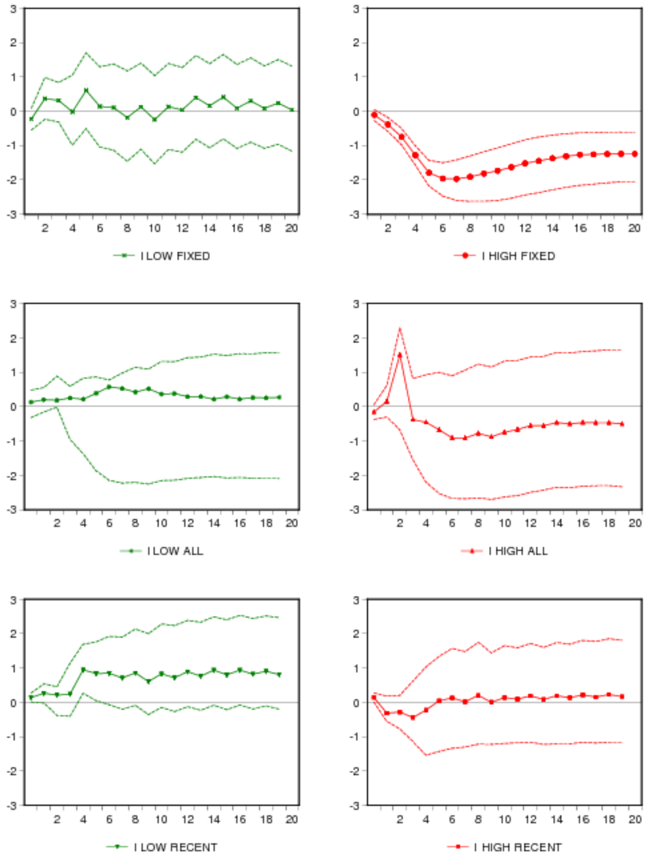

Figure 15: Responses of investment to a government spending shock
Left: low initial state, right: high initial state. Top: fixed states, middle: evolving states, averages over all histories, bottom: evolving states, averages over recent histories (post 1984)

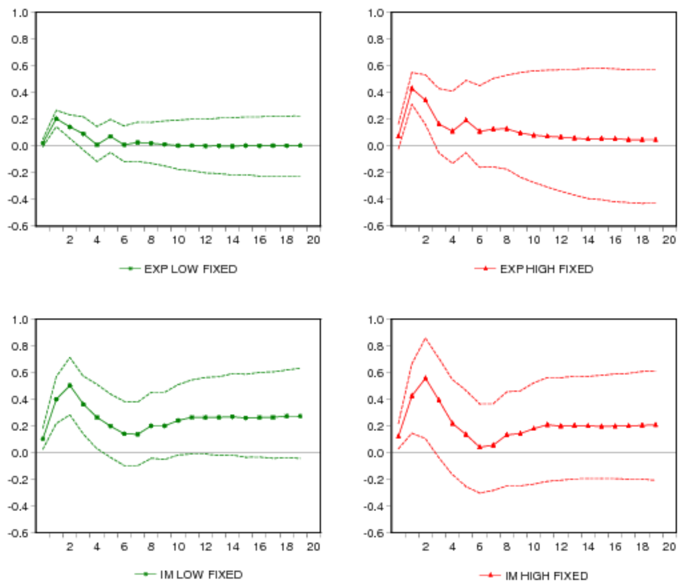

Figure 16: Fixed-regime responses of exports and imports to a government spending shock
Left: low regime, right: high regime. Top: exports, bottom: imports

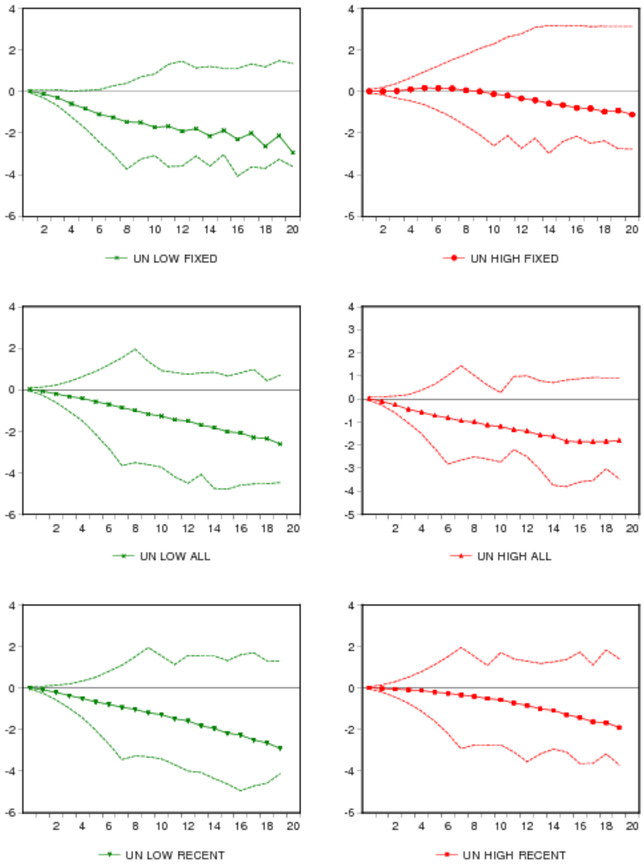

Figure 17: Responses of the unemployment rate to a government spending shock
Left: low initial state, right: high initial state. Top: fixed states, middle: evolving states, averages over all histories, bottom: evolving states, averages over recent histories (post 1984)

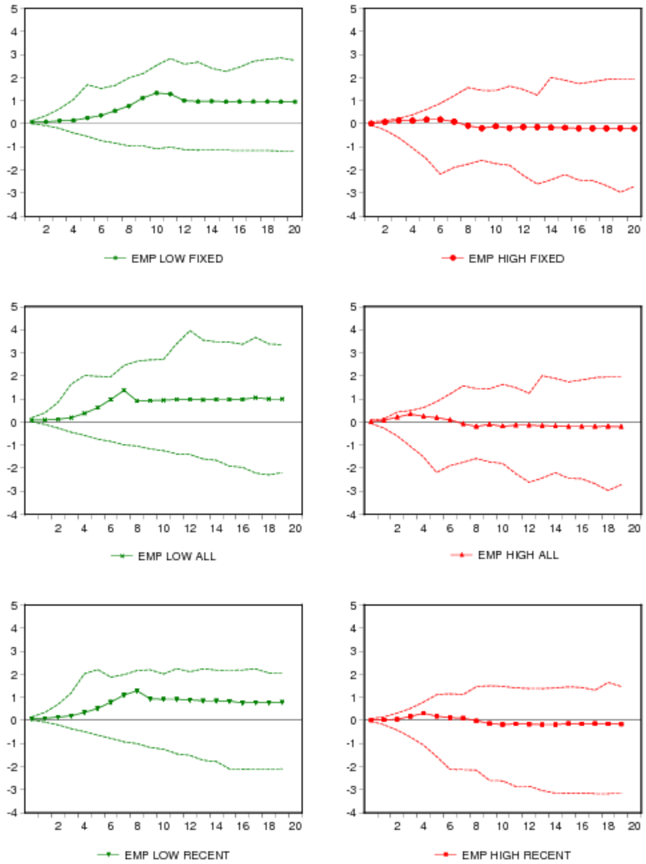

Figure 18: Responses of employment to a government spending shock

Left: low initial state, right: high initial state. Top: fixed states, middle: evolving states, averages over all histories, bottom: evolving states, averages over recent histories (post 1984)

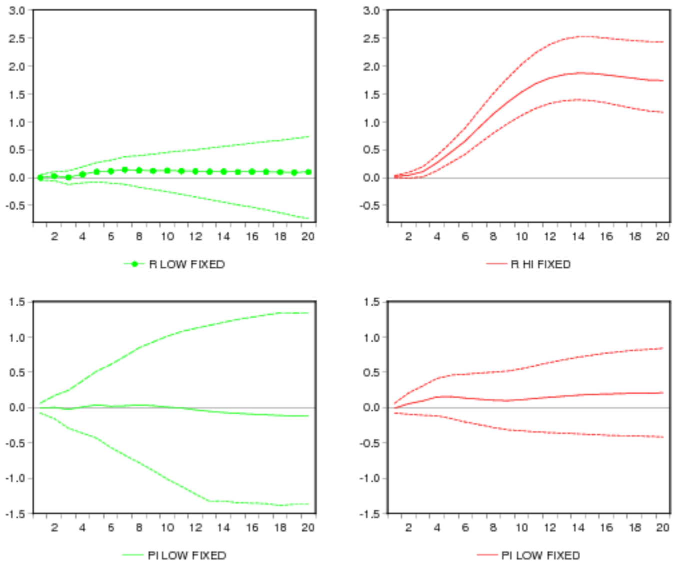

Figure 19: Fixed-regime responses of real interest rates and inflation to a government spending shock

Left: low initial state, right: high initial state. Top: real interest rate, bottom: inflation.

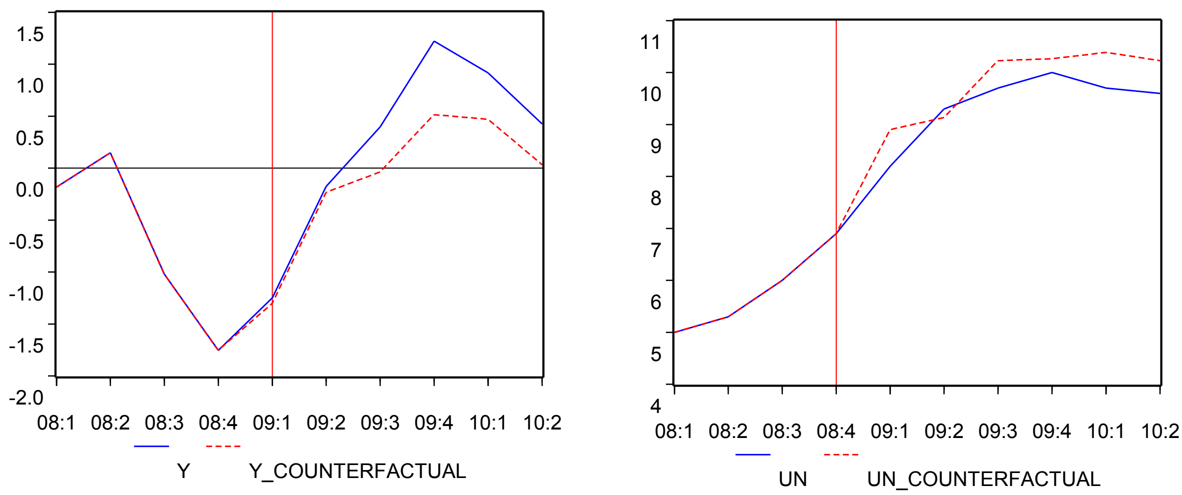

Figure 20: Counterfactual output and unemployment rate following the ARRA Stimulus

Notes: Actual Path (solid blue), Median Simulated Counterfactual Path (dashed red).

C. Appendix for Chapter 3

C.1. Generalized Impulse Response Function Comparison

The procedure for computing the generalized impulse response functions (GIRFs) follows Koop, Pesaran and Potter (1996), with the modification of considering an orthogonal structural shock, as in Kilian and Vigfusson (2011). The generalized impulse response is defined as the effect of a one-time shock on the forecasted level of variables in the model, and the response is compared against a baseline “no shock” scenario.

$$GIRF_y(k, shock_t, \Psi_{t-1}) = \alpha * [E[Y_{t+k}|shock_t, \Psi_{t-1}] - E[Y_{t+k}|\Psi_{t-1}]] \tag{35}$$

where k is the forecasting horizon, Ψ_{t-1} denotes the initial values of the variables in the model, and α is the scaling factor, discussed below. The impulse response is then computed by simulating the model. The positive shock to government spending is normalized to be equal to 1 percent of GDP (at the time the shock occurs), the large positive shock is normalized to be equal to 2 percent of GDP, and the negative shock is normalized to be equal to 1 percent of GDP.

The $GIRF_y$ response for a given draw $\Theta^{(i)}$ of the MH algorithm is generated using the following steps:

1. Pick a history Ψ_{t-1}. The history is the actual value of the lagged endogenous variable at a particular date.

2. Pick a sequence of 4-dimensional forecast errors ε_{t+k}, $k = 0, 1, ..., 20$. The forecast errors are simulated assuming an independent Gaussian process with mean zero and variance-covariance matrix equal to $\lambda_t^{(i)} * \Omega^{(i)}$.

3. Using Ψ_{t-1}^r and ε_{t+k}, simulate the evolution of Y_{t+k} over $l + 1$ periods. Denote the resulting path $Y_{t+k}(\varepsilon_{t+k}, \Psi_{t-1})$ for $k = 0, 1, ...l$.

4. Using the Cholesky decomposition of Ω_t to orthogonalize the shocks, solve for the

government spending shock at time t, replace it with a shock equal to 1 percent of GDP, and reconstruct the implied vector of forecast errors. Denote the implied vector of forecast errors as $\varepsilon_t^{shock_t}$, the sequence of forecast errors as $\varepsilon_{t+k}^{shock_t}$, and the resulting simulated evolution of Y_{t+k} over $l+1$ periods as $Y_{t+k}(\varepsilon_{t+k}^{shock_t}, \Psi_{t-1})$ for $k = 0, 1, ..., l$.

5. Construct a draw of a sequence of impulse responses as $Y_{t+k}(\varepsilon_{t+k}^{shock_t}, \Psi_{t-1}) - Y_{t+k}(\varepsilon_{t+k}, \Psi_{t-1})$ for $k = 0, 1, ..., l$.

6. Repeat steps 2 to 5 for B times, with $B = 500$, and average the sequences of responses to obtain a consistent estimate of the impulse response function conditional on the history and the size of the shock.

   a) For spending shocks equal to 1% of GDP, I do not rescale the response

   b) For large shocks (taken to be equal to 2% of GDP), I rescale the response using $\alpha = 0.5$

   c) For negative shocks, I rescale the response using $\alpha = -1$.

7. To obtain the average response for a subset of histories, repeat steps 1-6 for a the subset of histories of interest, and report the response averaged over all histories. In the second chapter, the average impulse responses are reported for all low states, all high states, the rapid expansion of the late 1990s and the Great recession.

8. In order to compare the responses for two types of shocks for a fixed history and, or the responses for two different histories,i construct the difference

$$\Delta IRF = GIRF_y(k, shock_t^1, \Psi_{t-1}^1) - GIRF_y(k, shock_t^2, \Psi_{t-1}^2).$$

Because the impulse responses are nonlinear functions of the parameters, their distribution of both the generalized impulse responses and the significance ΔIRF is nonstandard and it is not necessarily symmetric around the mean. In this case, reporting the median value is unlikely to be adequate, as the median may not be a valid measure of central tendency, and the median impulse response may not correspond to a well-defined structural model. In order to circumvent this problem, i adapt the approach proposed by Inoue and Kilian (2012). For a given history, i evaluate the impulse response function for each draw of the MH algorithm, drawing the entire impulse response function for periods 1 through 20. Then i average over the histories of interest, and i evaluate the posterior likelihood of the impulse response for that draw of the algorithm. The impulse response function with the highest average posterior likelihood is then used for inference. To construct the $(1-\alpha)*100\%$ credibility interval, i order the posterior likelihood values, and i include the impulse responses whose posterior likelihood was in the upper $(1-\alpha)*100$ percentile. This method results in a "credibility cloud" with a shot gun pattern because I draw entire impulse responses rather than responses for each individual point in time. For easy interpretation,I report only the outer points of the cloud. To convert the responses to dollar-for-dollar or jobs-for-dollar responses, all the impulse responses are converted to cumulative responses, and then scaled using the ratio $G_t/Variable_t$ for every t.

C.2. Graphs

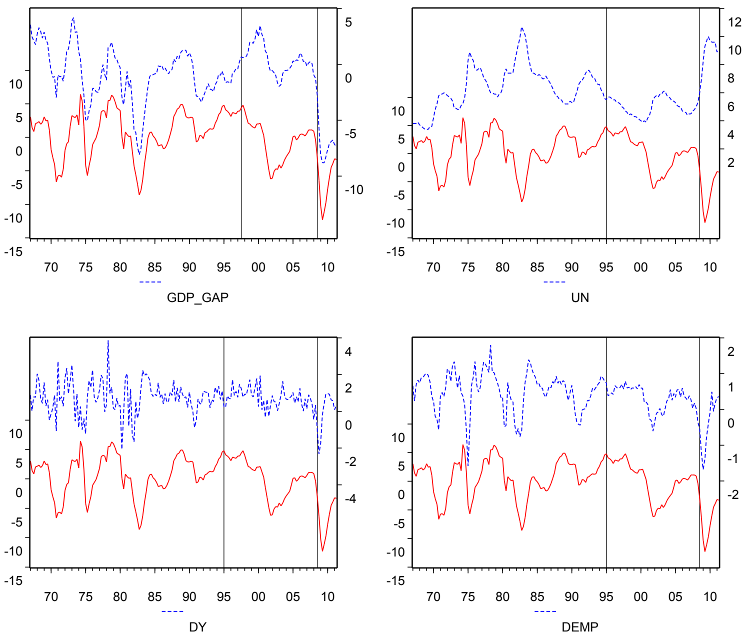

Figure 21: Capacity utilization vs other measures of slack, vertical lines at particular histories of interest

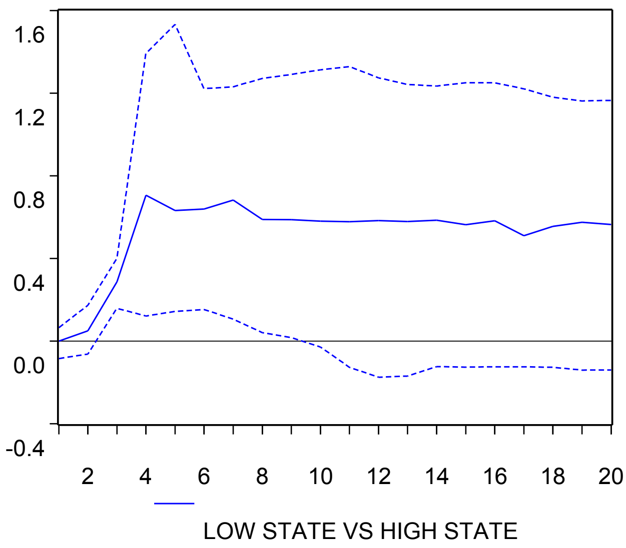

Figure 22: Significance: low state vs high state, 90% CI

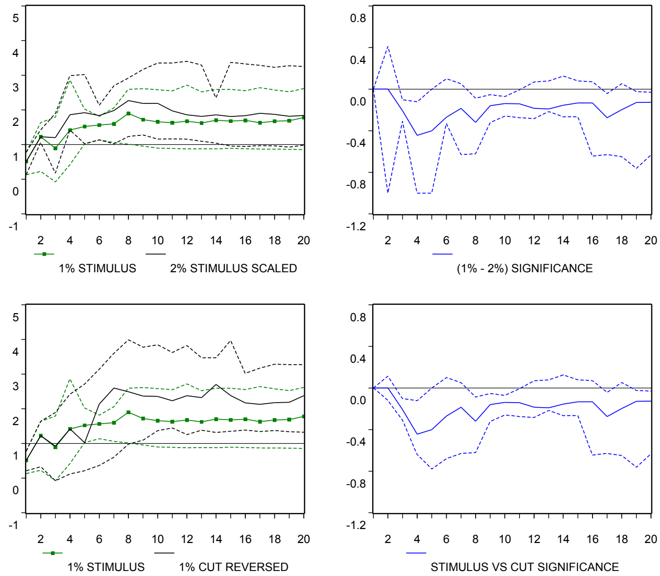

Figure 23: Responses of output to government spending: Great Recession, 90% CI

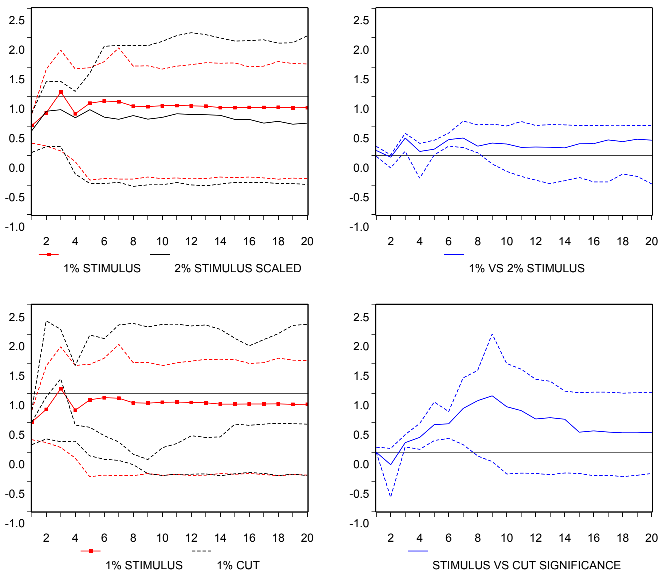

Figure 24: Responses of output to government spending: Late 1990s, 90% CI

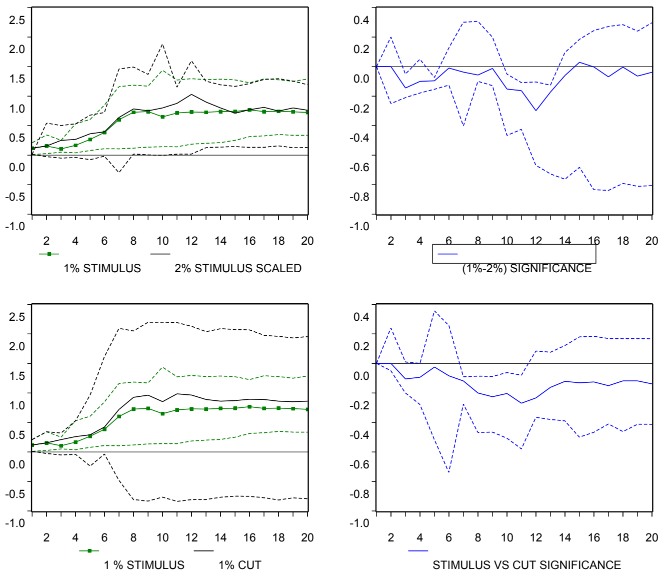

Figure 25: Responses of consumption to government spending: Great Recession, 90% CI

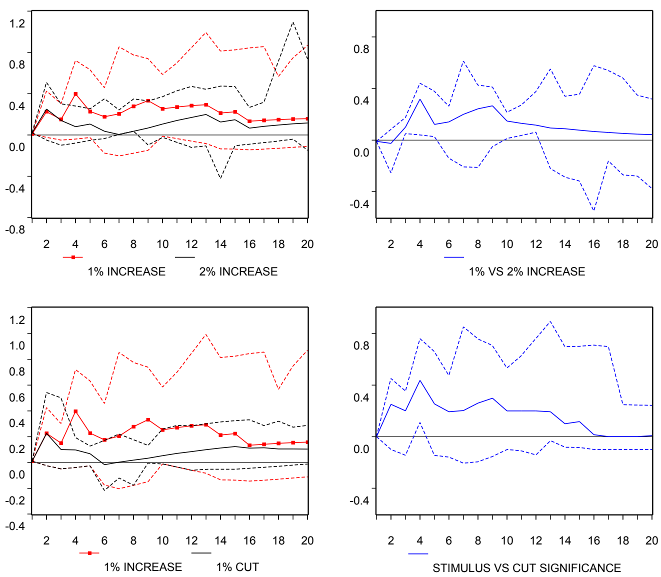

Figure 26: Responses of consumption to government spending: Late 1990s, 90% CI

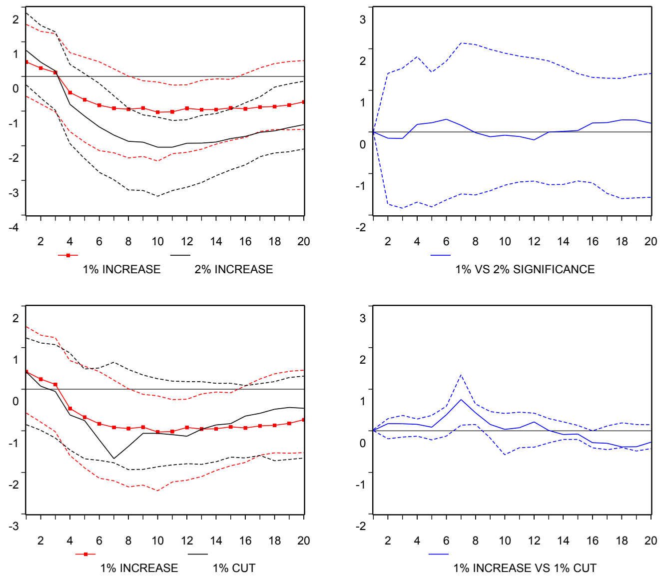

Figure 28: Responses of investment to government Spending: Late 1990s, 90% CI

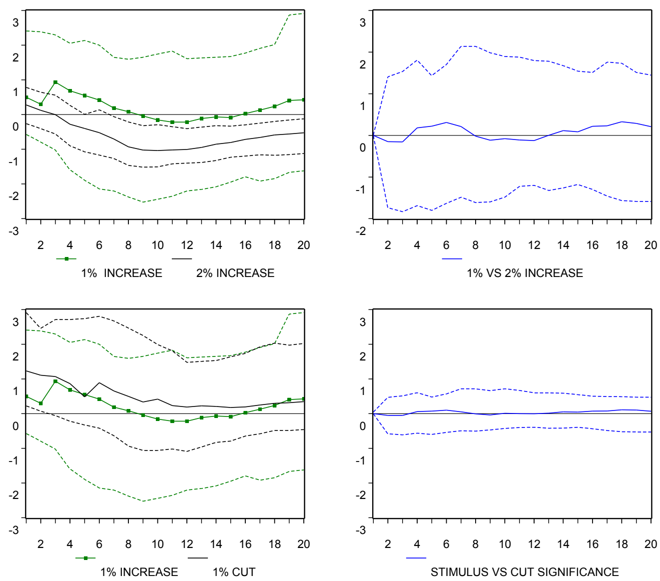

Figure 27: Responses of investment to government spending: Great Recession, 90% CI

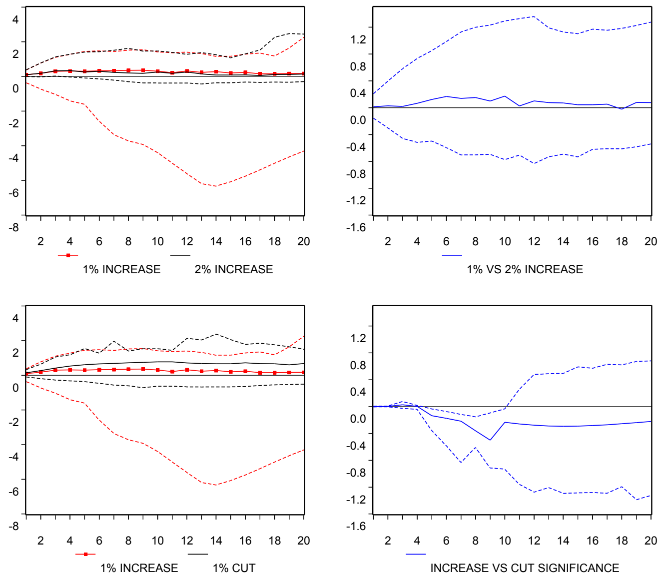

Figure 30: Responses of payroll employment to government spending: Late 1990s, 90% CI

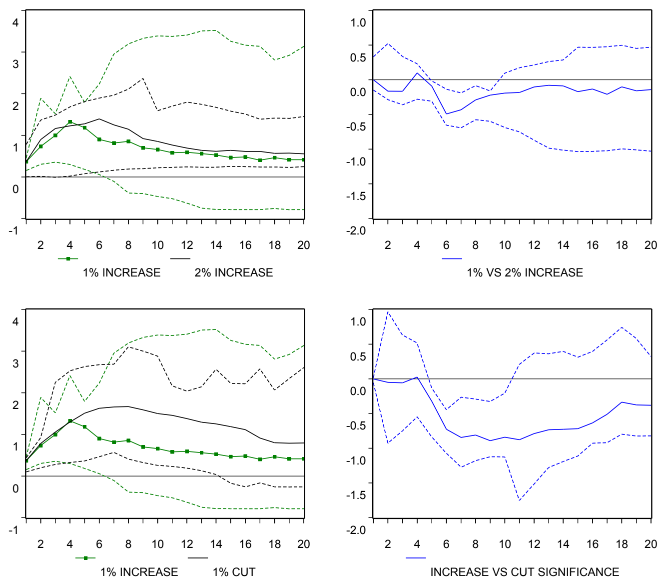

Figure 29: Responses of Payroll employment to government spending: Great Recession, 90% CI

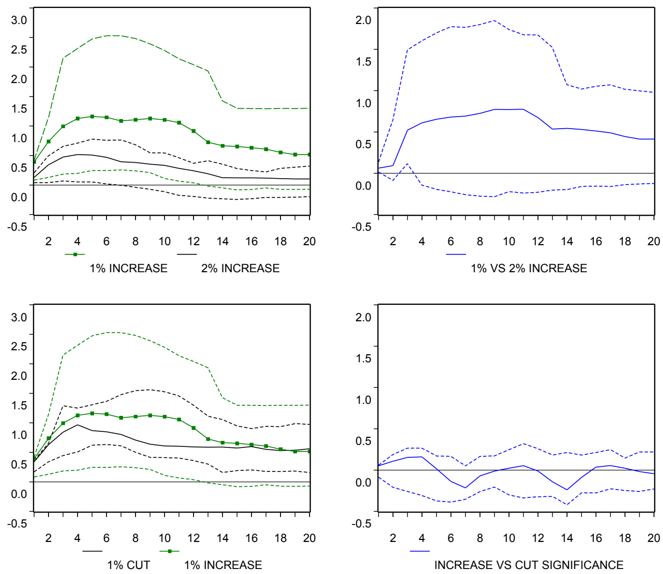

Figure 31: Responses of private employment to government spending: Great Recession, 90% CI

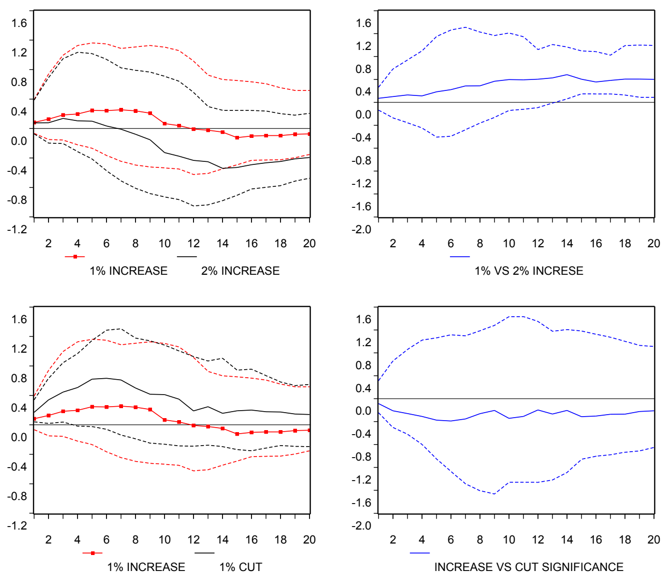

Figure 32: Responses of private employment to government spending: late 1990s, 90% CI

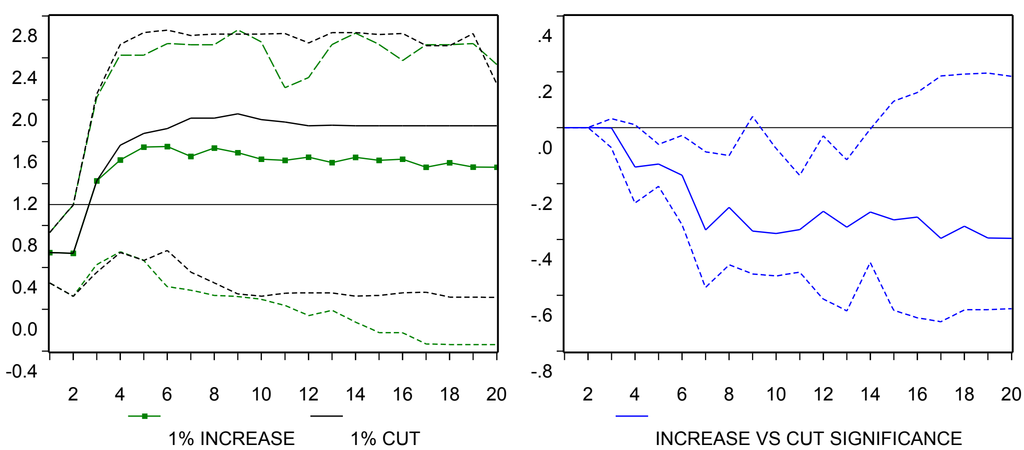

Figure 33: Responses of output to 1% increase and 1% cut: 2012Q3, 90% CI