Assessing the Effects of Fiscal Shocks*

Craig Burnside[†] Martin Eichenbaum[‡] Jonas D.M. Fisher[§]

November 1999

Abstract

This paper investigates the response of real wages and hours worked to an exogenous shock in fiscal policy. We identify this shock with the dynamic response of government purchases and tax rates to an exogenous increase in military purchases. The fiscal shocks that we isolate are characterized by highly correlated increases in government purchases, tax rates and hours worked as well as persistent declines in real wages. We assess the ability of standard Real Business Cycle models to account for these facts. They can–but only under the assumption that marginal income tax rates are constant, a standard assumption in the literature. Once we abandon this counterfactual assumption, RBC models cannot account for the facts. We argue that our empirical findings pose a challenge to a wide class of business cycle models.

*We would like to thank Lawrence J. Christiano and Lars Hansen for helpful conversations. The views expressed in this paper do not necessarily represent the views of the Federal Reserve Bank of Chicago, the Federal Reserve System or the World Bank. Martin Eichenbaum gratefully acknowledges the financial support of a grant from the National Science Foundation to the National Bureau of Economic Research.

[†]The World Bank

[‡] Northwestern University, Federal Reserve Bank of Chicago, NBER

[§]Federal Reserve Bank of Chicago

1. Introduction

This paper investigates the response of real wages and hours worked to an exogenous shock in fiscal policy. We identify this shock as the dynamic response of government purchases and tax rates to an exogenous increase in military purchases. The fiscal shocks that we isolate are characterized by highly correlated increases in government purchases, tax rates and hours worked as well as persistent declines in real wages. We assess the ability of standard Real Business Cycle models to account for these facts. They can - but only under the assumption that marginal income tax rates are constant, a standard assumption in the literature. Once we abandon this counterfactual assumption, RBC models cannot account for the facts.

The most basic failure of these models is that they predict hours worked and government purchases are sharply *negatively* correlated after a fiscal policy shock. In reality, after such a shock, hours worked and government purchases are strongly *positively* correlated. The traditional salve of technology shock driven RBC models - a high elasticity of labor supply - does not help improve their ability to account for the affects of a fiscal shock. In fact it exacerbates their failings. We infer that changes in the models' structure whose major effect is to increase the elasticity of labor supply will not remedy their shortcomings. After reviewing other model perturbations, we argue that our empirical findings pose a challenge to a wide class of business cycle models.

Why does allowing for empirically plausible movements in tax rates have such an important effect on the performance of RBC models? In the data, a fiscal policy shock leads to highly correlated hump-shaped movements in tax rates and government purchases. A rise in government purchases raises the present value of agents' taxes, thus triggering an increase in aggregate labor supply. But agents prefer to work harder when tax rates are low and less hard when tax rates are high. So a hump shaped rise in tax rates has both intratemporal and intertemporal substitution effects on labor supply. Once these substitution effects are taken into account, RBC models counterfactually predict that, after a fiscal policy shock, government purchases are strongly negatively correlated with hours worked. The mismatch between model and data is worse the more elastic labor supply is assumed to be: the more elastic labor supply is, the more agents wish to substitute hours worked towards periods when tax rates are low and away from periods when tax rates are high. While this magnifies the volatility of hours worked, it also magnifies the predicted negative conditional correlation between hours worked and government purchases.

The idea that the consequences of a fiscal policy shock depend on how increases in government consumption are financed is certainly not new. For example, Baxter and King (1993) forcefully demonstrate this point in the context of standard RBC models. In their model, when an increase in government purchases is financed by lump sum taxes, a fiscal shock generates a rise in employment and output and a fall in real wages. But, when an increase in government purchases is financed entirely by distortionary income taxes, employment, output and after-tax real wages all fall. In a similar vein, Mulligan's (1998) argument that neoclassical models cannot account for the rise in U.S. employment during WWII rests critically on the observation that marginal income tax rates rose dramatically during the war. Finally, Ohanian's (1997) analysis of the consequences of a rise in government purchases is predicated on the fact that the U.S. financed WWII and the Korean war in fundamentally different ways.

These observations notwithstanding, many studies proceed under the assumption that increases in government purchases are entirely financed by lump sum taxes.[1] The results in Baxter and King (1993), Mulligan (1998) and Ohanian (1997) suggest that this assumption *may* give rise to misleading results. The only way to know whether it actually does is to confront models with an experiment that is commensurate with what occurred in the data. But this requires that we

[1] See for example Christiano and Eichenbaum (1992), Devereaux, Head and Lapham (1996), Edelberg, Eichenbaum and Fisher (1999), Ramey and Shapiro (1998) and Rotemberg and Woodford (1992).

know how actual average marginal tax rates historically responded to exogenous increases in government consumption. Providing this information and showing how to use it are two of the main objectives of this paper.

Given data on average marginal tax rates, the key empirical problem is identifying *exogenous* changes in fiscal policy. The literature has pursued various approaches.[2] We build on the approach used by Ramey and Shapiro (1998) who focus on changes associated with exogenous movements in defense spending. To isolate such movements, Ramey and Shapiro (1998) identify three political events, arguably unrelated to developments in the domestic U.S. economy, that led to large military buildups. We refer to these events as 'Ramey-Shapiro episodes'. The weakness of this approach is that Ramey and Shapiro only identify three episodes of exogenous shocks to fiscal policy. In our view this weakness is more than offset by the compelling nature of their assumption that the war episodes are exogenous. Certainly their assumption seems plausible relative to the assumptions required to isolate the exogenous component of statistical innovations in government purchases and tax rates.

The remainder of this paper is organized as follows. Section 2 presents our evidence on the effects of a fiscal shock. Section 3 discusses a limited information strategy for using our results to assess the empirical plausibility of competing business cycle models. Section 4 reports the results of implementing this strategy on a simple, prototypical RBC model. Section 5 discusses the effects of various perturbations to the benchmark model. Section 6 contains concluding remarks.

2. Evidence on the Effects of a Shock to Fiscal Policy

In this section we accomplish two tasks. First, we describe our strategy for estimating the effects of an exogenous shock to fiscal policy. Second, we present the results of implementing this strategy. The main contribution relative to Ramey and Shapiro's empirical analysis is that (i) we abandon their assumption that the

[2] See Blanchard and Perotti (1998) and Edelberg, Eichenbaum and Fisher (1999) for discussions of alternative approaches.

episodes which they isolate were of equal intensity and (ii) we bring tax rates into the analysis.

2.1. Identifying the Effects of a Fiscal Policy Shock

Ramey and Shapiro (1998) pursue a 'narrative approach' to isolate three arguably exogenous events that led to large military buildups and increases in total government purchases: the Korean War, the Vietnam War and the Carter-Reagan defense buildup. Based on their reading of history, they date these events at 1950:3, 1965:1 and 1980:1.

To estimate exogenous movements in government purchases, g_t, and average marginal tax rates, τ_t, and the resulting effects on other macro variables, we use the following procedure. Define the three dummy variables D_{it}, $i = 1, 2, 3$, where

$$D_{it} = \begin{cases} 1, & \text{if } t = d_i \\ 0, & \text{otherwise} \end{cases}$$

and d_i denotes the i^{th} element of

$$d = \begin{pmatrix} 1950{:}3 & 1965{:}1 & 1980{:}1 \end{pmatrix}'.$$

We include the D_{it}'s as explanatory variables in a vector autoregression (VAR). Suppose that the stochastic process Z_t has the representation:

$$Z_t = A_0 + A_1(L)Z_{t-1} + \sum_{i=1}^{3} A_2(L)\psi_i D_{it} + u_t, \tag{2.1}$$

where $Eu_t = 0$,

$$Eu_t'u_{t-s} = \begin{cases} 0, \text{ for all } s \neq 0 \\ \Sigma, \text{ for } s = 0, \end{cases}$$

Σ is a positive definite matrix of dimension equal to the number of elements in Z_t and $A_j(L)$, $j = 1, 2$ are finite ordered vector polynomials in nonnegative powers of the lag operator L. In addition, the ψ_i are scalars with ψ_1 normalized to unity. The parameters ψ_2 and ψ_3 measure the intensity of the second and third Ramey-Shapiro episodes relative to the first. This specification allows us to depart from the assumption in Ramey and Shapiro (1998) and Edelberg, Eichenbaum and Fisher (1999) that the episodes are of equal intensity, i.e. $\psi_i = 1$, $i = 1, 2, 3$.

We estimated (2.1) by maximum likelihood assuming a Gaussian likelihood function. A consistent estimate of the response of Z_{it+k}, the i^{th} element of Z at time $t + k$, to the onset of the i^{th} Ramey-Shapiro episode is given by an estimate of the coefficient on L^k in the expansion of $\psi_i \left[I - A_1(L)L\right]^{-1} A_2(L)$. So while the episodes may differ in intensity, their dynamic effects are the same, up to a scale factor, ψ_i.

2.2. Empirical Results

In this subsection we present the results of implementing the procedure discussed above. Unless otherwise noted, the vector Z_t contains the log of time t real GDP, the net three month Treasury bill rate, the log of the producer price index of crude fuel, the log of a measure of the average marginal income tax rate, the log of real government purchases, the log of hours worked, and after tax real wages.[3] The VAR has six lagged values of all variables. This lag length was chosen using the modified likelihood ratio test described in Sims (1980).[4] All estimates are based on quarterly data from 1947:1 to 1994:4. The Appendix describes the data used in our analysis.

As background to our analysis, the first row of Figure 1 displays the log of real defense expenditures and the log of real government purchases, along with vertical lines at the dates of Ramey-Shapiro episodes. The lower left hand panel of Figure 1 reports the share of government purchases in GDP. The time series on real defense expenditures is dominated by three events: the large increases in real defense expenditures associated with the Korean war, the Vietnam war and the Carter-Reagan defense buildup. The Ramey-Shapiro dates essentially mark the beginning of these episodes. In the models that we explore it is total government

[3] Our measure of government purchases includes real defense expenditures but excludes non-military investment.

[4] Sims corrects the likelihood ratio statistic to account for small sample bias.

purchases, rather than military purchases that is relevant. Figure 1 shows that the Ramey-Shapiro episodes also coincide with rises in real government purchases. It is apparent that these rises associated with the different episodes are not of equal magnitude. Our econometric procedure, which allows for different intensities of the Ramey-Shapiro episodes, is consistent with this observation.

The lower right hand panel of Figure 1 displays a measure of the average marginal tax rate, taken from Stephenson (1998), along with vertical lines at the Ramey-Shapiro dates. This tax rate measure is an updated version of the one constructed by Barro and Sahasakul (1983, 1986). Their tax rate measure is a weighted average of statutory marginal tax rates, where the weights are the shares of adjusted gross income subject to each statutory rate.[5] The measure of the average marginal tax can move for a variety of reasons. The U.S. tax system is progressive and, over much of our sample period, tax brackets were set in nominal terms. Consequently, the average marginal tax rate was affected by inflation as well as real growth. This implies that there are at least two ways to implement tax rate changes after a fiscal policy shock: (i) undertake explicit legislative action to raise rates, and (ii) abstain from undertaking actions necessary to prevent rises in marginal tax rates due to inflation or growth. From the perspective of the agents in our models, all that matters is whether tax rates change, not how they are changed.

A number of interesting facts emerge from the lower right hand panel of Figure 1. First, tax rates rise substantially after the first Ramey-Shapiro date. This is consistent with Ohanian's (1998) argument that the Korean war was financed by a contemporaneous rise in distortionary taxes. Second, tax rates rise one or two quarters after the second Ramey-Shapiro date and at the onset of the third Ramey-Shapiro date. These observations suggest the potential importance of taking into account the response of average marginal tax rates to a fiscal policy

[5] See Stephenson (1998) for refinements to the Barro-Sahasakul measure. We find that our results are qualitatively insensitive to working with Stephenson's refined tax rate measure, as well as to working with another tax rate measure suggested by Seater (1985).

shock when evaluating a structural model.

As mentioned above, our econometric procedure allows us to distinguish between the intensities of the different Ramey-Shapiro episodes. Recall that we normalize the first episode (Korea) to be of unit intensity. Our point estimates of the intensities of the second and third episodes are equal to 0.20 and 0.38, respectively. The 95% confidence intervals are 0.09 to 0.32 and 0.26 to 0.50, respectively.[6] In all cases below, we report the dynamic response function of various aggregates to an episode of unit intensity. This simply scales the size of the impulse response functions.

Column 1 of Figure 2 reports the responses of real government purchases and the average marginal income tax rate. The solid lines display point estimates of the coefficients of the dynamic response functions.[7] The dashed lines correspond to 95% confidence interval bands.[8] Note that the onset of a Ramey-Shapiro episode leads to a large, persistent, hump-shaped rise in total government purchases with a peak response of about 23% seven quarters after the shock.[9] Next consider

[6] As an additional check, we sequentially redid our analysis using only two of the three Ramey - Shapiro episodes. Our qualitative results were robust to ignoring either the second or third episode.

[7] With one exception, the impulse response functions are reported as percentage deviations from a variable's unshocked path. The exception is the impulse response function of the average marginal tax rate (Figure 2), which is reported as the deviation of the tax rate from its unshocked level, measured in percentage points.

[8] These were computed using a bootstrap Monte Carlo procedure. Specifically, we constructed 500 time series on the vector X_t as follows. Let $\{\widehat{u}_t\}_{t=1}^T$ denote the vector of residuals from the estimated VAR. We constructed 500 sets of new time series of residuals, $\{\widehat{u}_t(j)\}_{t=1}^T$, $j = 1, ..., 500$. The t^{th} element of $\{\widehat{u}_t(j)\}_{t=1}^T$ was selected by drawing randomly, with replacement, from the set of fitted residual vectors, $\{\widehat{u}_t\}_{t=1}^T$. For each $\{\widehat{u}_t(j)\}_{t=1}^T$, we constructed a synthetic time series of Z_t, denoted $\{Z_t(j)\}_{t=1}^T$, using the estimated VAR and the historical initial conditions on Z_t. We then re-estimated the VAR using $\{Z_t(j)\}_{t=1}^T$ and the historical initial conditions, and calculated the implied impulse response functions for $j = 1, ..., 500$. We then calculated the 13^{th} lowest and 487^{th} highest values of the corresponding impulse response coefficients across all 500 synthetic impulse response functions. The boundaries of the confidence intervals in the figures correspond to a graph of these coefficients.

[9] Working with an equal intensity specification ($\psi_i = 1$, $i = 1, 2, 3$) Ramey and Shapiro (1998) show that the response of real defense purchases is larger in size but similar in shape to the response of total government purchases. This remains the case if we allow the Ramey-Shapiro episodes to be of different intensities.

the response of the average marginal tax rate. A number of key results emerge here. First, the tax rate rises in a hump-shaped pattern, mirroring the dynamic response of government purchases, with the peak occurring roughly seven quarters after the onset of a Ramey-Shapiro episode. Indeed the conditional correlation between the tax rate and government purchases is equal to 0.97, with a 95% Monte Carlo confidence interval $(0.54, 0.99)$. Second, the rise in the tax rate is large, peaking at 3.3 percentage points after seven quarters. This represents a rise of roughly 19% in the tax rate relative to its value in 1949.

Viewed overall, these results indicate that, for these episodes, a fiscal policy shock is characterized by a large persistent rise in government purchases *and* a rise in average marginal tax rates. Accordingly, both features must be incorporated into the experiment that we conduct in our model economy.

Column 2 of Figure 2 reports the response of private business hours and a measure of after-tax manufacturing real wages to the onset of a Ramey-Shapiro episode.[10] A number of interesting results emerge. First, paralleling the response of total government purchases, hours worked has a delayed hump shaped response. The conditional correlation between these two variables equals 0.93 with a 95% confidence interval of $(0.48, 0.98)$. The peak response in hours worked is roughly 6.2% and occurs about 6 quarters after the onset of a Ramey-Shapiro episode.[11] Second, we see that after-tax manufacturing real wages fall after the fiscal shock, with a peak decline of 10% roughly 8 quarters after the shock. The conditional correlation between government purchases and after-tax real wages equals -0.97 with a 95% confidence interval of $(-0.86, -0.98)$. The response of before-tax manufacturing real wages to the onset of a Ramey-Shapiro episode looks similar to that of after-tax real wages, but it is smaller. The peak decline is only 5.5% and occurs roughly 8 quarters after the shock. The conditional correlation be-

[10] The latter is the broadest measure of real wages available over our sample period.

[11] See Ramey and Shapiro (1998) and Edelberg, Eichenbaum and Fisher (1999) for the responses of real GDP and various other measures of hours worked obtained under the assumption that the Ramey-Shapiro episodes are of equal intensity. The qualitative nature of these responses is unaffected by allowing for different intensities.

tween government purchases and before-tax real wages equals -0.94 with a 95% confidence interval of $(-0.75, -0.97)$.[12]

In sum, a Ramey-Shapiro episode is marked by statistically significant falls in after-tax real wage rates and rises in government purchases, average marginal tax rates, total output and employment.[13]

3. A Limited Information Diagnostic Procedure

The previous section displayed our estimates of the dynamic consequences of a fiscal policy shock to government purchases and average marginal tax rates. In addition we displayed the corresponding movements in hours worked and real wages. In this section we discuss a limited information procedure for using these results to assess the empirical plausibility of competing models.

Suppose that the fiscal authority sets the time t value of $F_t = (\ g_t \quad \tau_t\ )'$ according to the rule,

$$F_t = f(\Omega_t) + M_F(L)D_t + \varepsilon_{Ft}. \tag{3.1}$$

Here Ω_t is the information set available to the fiscal authority when it sets F_t, and ε_{Ft} is a serially uncorrelated shock that is orthogonal to the elements of Ω_t. The variable D_t is an exogenous stochastic process whose realizations are unaffected by all the other variables in our model. The particular realization of $\{D_t\}$ that we have as data is $D_t = \psi_i$ at the i^{th} Ramey and Shapiro date, $i = 1, 2, 3$, and

[12] Nominal wages are deflated using the CPI. Our results are very similar if we deflate using the GDP deflator. Ramey and Shapiro (1998) and Edelberg, Eichenbaum and Fisher (1999) show that various measures of gross real wages computed using various deflators fall after the onset of fiscal policy shock. These results are not sensitive to allowing for different intensities of the episodes.

[13] The Monte Carlo methods that we used to quantify the importance of sampling uncertainty do not convey any information about 'date' uncertainty. This is because they take as given the Ramey and Shapiro dates. One simple way to assess the importance of date uncertainty is to redo the analysis perturbing the Ramey and Shapiro dates. Edelberg, Eichenbaum and Fisher (1999) document the robustness of inferences under the assumption that the different episodes are of equal intensity.

$D_t = 0$ otherwise. Finally, $M_F(L)$ is a finite-ordered, 2×1 matrix polynomial in the lag operator.

Consider the $k \times 1$ vector Z_t which we partition as

$$Z_t = \begin{pmatrix} \bar{Z}_t \\ F_t \end{pmatrix}.$$

The class of models that we consider implies that the equilibrium law of motion for Z_t takes the form of, or can be well approximated by, the system of linear difference equations:

$$B_0 Z_t = \kappa + B(L) Z_{t-1} + M(L) D_t + \varepsilon_t. \tag{3.2}$$

Here $B(L)$ is a finite ordered polynomial in the lag operator L, $M(L) = [\ 0'_{k-2} \quad M_F(L)'\ ]'$, 0_{k-2} is a $k-2$ vector of zeroes, and the elements of $\varepsilon_t = (\ \varepsilon'_{\bar{z}t} \quad \varepsilon'_{Ft}\ )'$ are uncorrelated with each other, with D_t, and with lagged values of Z_t. The last two rows of (3.2) consist of the policy rule (3.1), assuming that $\bar{Z}_t$ contains all the elements of Ω_t and the matrix consisting of the lower right hand 2×2 sub-block of B_0 is equal to the identity matrix. With this specification the only variables that are directly affected by D_t are those in F_t. The onset of a Ramey-Shapiro episode sets off a chain of movements in F_t which leads to movements in $\bar{Z}_t$ through the mechanisms embedded in the particular model under consideration.

We can illustrate the equivalence between our theoretical model, (3.2), and a modified VAR of the form (2.1) by multiplying both sides of (3.2) by B_0^{-1}. Notice that the two representations are equivalent when

$$A_0 = B_0^{-1}\kappa, \quad A_1(L) = B_0^{-1}B(L), \quad A_2(L) = B_0^{-1}M(L), \quad u_t = B_0^{-1}\varepsilon_t. \tag{3.3}$$

To characterize impulse response functions we use the moving average representation corresponding to (3.2) given by

$$Z_t = \pi_0 + \pi(L)\varepsilon_t + H(L) D_t \tag{3.4}$$

where

$$\pi_0 = \begin{pmatrix} \pi_0^1 \\ \pi_0^2 \end{pmatrix}, \quad \pi(L) = \begin{pmatrix} \pi^1(L) \\ \pi^2(L) \end{pmatrix}, \quad H(L) = \begin{pmatrix} H^1(L) \\ H^2(L) \end{pmatrix}.$$

By assumption, $\{\pi^i\}$ and $\{H^i\}$ form square-summable sequences. Note that $H(L)$ completely characterizes the dynamic response path of the vector Z_t to the time t realization of D_t. In particular, the response of Z_{t+j} is given by the coefficient on L^j in $H(L)$.

It is useful to write the last two rows of (3.4) as

$$F_t = \pi_0^2 + \pi^2(L)\varepsilon_t + H^2(L)D_t \tag{3.5}$$

Relationship (3.5) expresses the time t values of $F_t = (\begin{array}{cc} g_t & \tau_t \end{array})'$ as a function of current and past values of all the shocks to the economy, as well as current and past values of D_t. We did not identify the elements of ε_t in our empirical analysis, since there is no need to. Under our assumptions, D_t is orthogonal to ε_t. So we can study the effects of a change in D_t abstracting from movements in ε_t, i.e. we can proceed under the assumption that $\varepsilon_t = 0$. This is equivalent to working with the representation for F_t given by

$$F_t = \pi_0^2 + H^2(L)D_t. \tag{3.6}$$

To assess the empirical plausibility of a model's implications for an exogenous shock to fiscal policy we can proceed as follows.

1. Estimate the VAR given by (2.1)

$$Z_t = A_0 + A_1(L)Z_{t-1} + \sum_{i=1}^{3} A_2(L)\psi_i D_{it} + u_t.$$

2. Use the estimates of A_0, $A_1(L)$ and $A_2(L)$ to obtain a moving average representation for Z_t that is equivalent to (3.4)

$$Z_t = \hat{\pi}_0 + \hat{\pi}(L)u_t + \hat{H}(L)D_t. \tag{3.7}$$

Notice that $\hat{H}^1(L)$ characterizes the dynamic responses of the non-fiscal variables, $\bar{Z}_t$, to the onset of an intensity-weighted Ramey-Shapiro episode.

3. Use the last two rows of (3.7) to characterize fiscal policy in the theoretical model.

4. Using this specification of policy, and calibrating the parameters of the theoretical model, calculate the dynamic response functions of the non-fiscal variables to a Ramey-Shapiro episode. Denote the polynomial in the lag operator that characterizes these responses as $\tilde{H}^1(L)$.

5. Compare the theoretical model's responses to their empirical counterparts, estimated in the second step. Abstracting from sampling uncertainty and the linearity assumptions implicit in the VAR analysis, the two sets of response functions should be the same, i.e. it should be the case that $\tilde{H}^1(L) = \hat{H}^1(L)$.

This last conclusion follows from the fact that an equilibrium for an economy with policy rule for F_t given by (3.1) is also an equilibrium for the economy with the policy rule given by (3.6).[14] Since D_t is orthogonal to ε_t and we are only interested in responses to D_t we can work with the version of (3.4) given by (3.6). Suppose that we solve a model assuming that fiscal policy is given by the estimated version of (3.6). For the models which we consider, there is a unique equilibrium. The dynamic response function of $\bar{Z}_t$ in this equilibrium is the same as the dynamic response function in the equilibrium when the policy rule is given by (3.1). So, sampling uncertainty aside, if the model has been specified correctly, the model based and estimated dynamic response functions of $\bar{Z}_t$ to the onset of an intensity weighted Ramey-Shapiro episode should coincide.

In practice there are two sources of sampling uncertainty. The first concerns the structural parameters of the model describing preferences and technology. Results in Burnside and Eichenbaum (1996) suggest that this source of uncertainty is unlikely to significantly affect inference for the models discussed below.[15] The

[14] For a more general discussion of the relationship between these two ways of representing policy in the context of monetary policy see Christiano, Eichenbaum and Evans (1998).

[15] Burnside and Eichenbaum calculate confidence intervals for the dynamic response functions

second source of sampling uncertainty pertains to the estimated response of the U.S. economy to a fiscal policy shock. Uncertainty about the response of U.S. government purchases and tax rates affects inference via its effect on the relevant experiment to be conducted in the model. Uncertainty about the actual dynamic response of employment and real wages affects how we assess the results of a given experiment in the model.

The second source of uncertainty is of particular concern to us because of the small number of Ramey-Shapiro episodes. To assess its significance we adopt the following procedure for testing hypotheses. We are interested in assessing the ability of the model to account for various conditional moments of the data, i.e. moments pertaining to the behavior of the economy conditional on a fiscal shock having occurred. One way to estimate such a moment is to use a point estimate, $\hat{\theta}_H$, of the vector of coefficients, θ_H, characterizing $H(L)$, in a way that does not involve the use of an economic model. We let $d(\hat{\theta}_H)$ denote the point estimate of a conditional moment obtained in this way. A different way to estimate the conditional moment is to use an economic model along with values for the parameters describing agents' preferences and technology and an estimate of the coefficients characterizing the exogenous variable policy rule, θ_H^2. Note that θ_H^2 is a subset of θ_H. We denote by $m(\hat{\theta}_H^2)$ the point estimate of the conditional moment in question derived from the economic model.

Let

$$s(\theta_H) = d(\theta_H) - m(\theta_H^2).$$

We are interested in testing hypothesis of the form:

$$H_0 : s(\theta_H) = 0.$$

An implication of results in Eichenbaum, Hansen and Singleton (1984) and Newey

in a standard RBC model to a shock in government purchases. They argue that the size of the confidence intervals is determined primarily by sampling uncertainty regarding the law of motion for government purchases, rather than the other parameters of the models, at least when the latter are estimated using the generalized method of moment techniques employed in Burnside and Eichenbaum (1996).

and West (1987) is that the test statistic

$$J = s(\widehat{\theta}_H)'\widehat{\text{var}}[s(\widehat{\theta}_H)]^{-1}s(\widehat{\theta}_H) \tag{3.8}$$

is asymptotically distributed as a chi-squared distribution with 1 degree of freedom, where $\widehat{\text{var}}[s(\widehat{\theta}_H)]$ is a consistent estimator of $\text{var}[s(\widehat{\theta}_H)]$.[16] Below we use this test statistic to formally assess the ability of a standard RBC model to account for various conditional moments of the data.

In sum, this section provides a rationale for a diagnostic procedure that is based on assessing whether a given economic model can account for the estimated response of the U.S. economy to an exogenous policy shock. The key step is to attribute views to agents about how fiscal policy evolves after the onset of a Ramey-Shapiro episode. These views are summarized by our estimate of the policy rule (3.6).

An important unresolved issue is what are agents' views about the law of motion for D_t. One option would be to model agents' subjective probability distribution over rare events such as the outbreaks of war and major military buildups induced by exogenous shocks. Given the difficulty of this task and the paucity of data on such events, we adopt the following simplification: we suppose that agents expect $D_t = 0$ for all t. In addition, we assume that a realization of $D_t = 1$ does not affect agents' future expectations of D_t, i.e. they continue to expect that future values of D_t will equal zero. So from their perspective, a realization of $D_t = 1$ is just like the realization of an iid exogenous shock to F_t. But once

[16] To generate an estimate of $\text{var}[s(\widehat{\theta}_H)]$ we use the same bootstrap procedure employed to compute confidence intervals for the impulse response functions estimated in the data. (See footnote 8). Specifically, let θ_{Hi} be the point estimate of the moving average coefficients of Z_t implied by the VAR coefficients generated by the ith bootstrap draw, $i = 1, ..., N$, where $N = 500$. Define $\overline{s}(\theta_{Hi}) = (1/N)\sum_{i=1}^{N} s(\theta_{Hi})$. Then

$$\widehat{\text{var}}[s(\widehat{\theta})] = \frac{1}{N-1}\sum_{i=1}^{N}[s(\theta_{Hi}) - \overline{s}(\theta_{Hi})]^2,$$

is a consistent estimate of $\text{var}[s(\widehat{\theta}_H)]$. An alternative is to use the Generalized Method of Moments strategy described in Christiano and Eichenbaum (1992).

such a shock occurs, the expected response of F_{t+j} is given by the coefficient on L^j in the polynomial $H^2(L)$.

4. A Prototypical Real Business Cycle Model

In this section we describe a prototypical RBC model and study its implications for how the economy responds to a fiscal policy shock. The section is divided into three parts. The first subsection describes our theoretical framework, the second subsection describes the way we calibrated the model's parameters and the third subsection discusses the model's quantitative properties.

4.1. Theoretical Framework

A representative household ranks alternative streams of consumption and hours worked according to

$$E_0 \sum_{t=0}^{\infty} \beta^t \left[\log C_t + \eta V(1-n_t)\right]. \tag{4.1}$$

$$V(1-n_t) = \begin{cases} \frac{1}{1-\mu}(1-n_t)^{1-\mu}, & \mu \geq 0 \\ \ln(1-n_t), & \mu = 1 \end{cases} \tag{4.2}$$

Here E_0 is the time 0 conditional expectations operator, β is a subjective discount factor between 0 and 1, while C_t and n_t denote time t consumption and the fraction of the household's time endowment devoted to work, respectively. Given (4.2), the representative household's Frisch elasticity of labor supply, evaluated at the steady state level of hours, n, is equal to $(1-n)/(n\mu)$.[17]

The household owns the stock of capital, whose value at the beginning of time t we denote by K_t. Capital evolves according to:

[17] For $\mu = 0$, this elasticity must be interpreted with some care. Hansen (1985) and Rogerson (1988) describe model economies in which the competitive equilibrium allocation is given by the solution to a social planning problem in which leisure enters into the planner's objective function in a linear manner ($\mu = 0$). This is true even though leisure need not enter individual agents' objective function linearly. So in their model, there is no link between individuals' Frisch elasticity of labor supply and the corresponding elasticity implied by the planner's preferences.

$$K_{t+1} = (1-\delta)K_t + I_t, \quad 0 < \delta < 1, \tag{4.3}$$

where I_t denotes time t investment in capital.

The household rents out capital and supplies labor in perfectly competitive spot factor markets. We denote the real wage rate per unit of labor by w_t and the real rental rate on capital by r_t. The government taxes both rental income net of depreciation, and wage income at the rate τ_t, so that the after-tax real wage and rental rate on capital are given by $(1-\tau_t)\, W_t$ and $(1-\tau_t)r_t + \delta\tau_t$, respectively. Therefore, the household's time t budget constraint is given by

$$C_t + I_t \leq (1-\tau_t)\, W_t n_t + (1-\tau_t) r_t K_t + \delta \tau_t K_t - \Phi_t$$

where Φ_t denotes lump sum taxes paid by the household.

A perfectly competitive firm produces output, Y_t, according to

$$Y_t \leq K_t^{\alpha}(X_t n_t)^{1-\alpha}, \quad 0 < \alpha < 1, \tag{4.4}$$

where X_t represents the time t state of technology. The firm sells its output in a perfectly competitive goods market and rents labor and capital in perfectly competitive spot markets. Technology evolves in the following deterministic fashion,

$$X_t = \gamma^t,\ \gamma \geq 1. \tag{4.5}$$

The government purchases G_t units of output at time t. For simplicity we assume the government balances its budget every period. Government purchases are financed entirely via the income tax, τ_t, and lump sum taxes, Φ_t. Consequently the government's budget constraint is given by

$$G_t = \tau_t W_t n_t + \tau_t (r_t - \delta) K_t + \Phi_t.$$

Given our assumptions, Ricardian equivalence holds with respect to the timing of lump sum taxes.[18] So we could allow the government to borrow part of the

[18] This assumes the absence of distortionary taxes on government debt.

difference between its expenditures and revenues raised from distortionary taxes, subject to its intertemporal budget constraint, and it would not affect our results.

Government purchases evolve according to

$$G_t = X_t g_t. \tag{4.6}$$

We assume that $\log(g_t)$ and $\log(\tau_t)$ have finite ordered $MA(q)$ representations:

$$\log(g_t) = \log(\overline{g}) + h_1(L)u_t \tag{4.7}$$

and

$$\log(\tau_t) = \log(\overline{\tau}) + h_2(L)u_t, \tag{4.8}$$

where $h_1(L)$ and $h_2(L)$ are finite ordered polynomials in nonnegative powers of the lag operator L, and $\overline{g}$ and $\overline{\tau}$ correspond to steady state government spending and taxes, respectively.[19] Note that u_t is common to both government spending and taxes. This formalizes the notion that government spending and taxes respond simultaneously to a common fiscal shock. The innovation is iid and is orthogonal to all model variables dated time $t-1$ and earlier.

It is convenient to define:

$$c_t = \tfrac{C_t}{X_t}, \quad k_t = \tfrac{K_t}{X_t} \quad w_t = \tfrac{W_t}{X_t} \quad \phi_t = \tfrac{\Phi_t}{X_t} \ , \tag{4.9}$$

as well as

$$u(c_t, 1-n_t) = \log c_t + \eta V(1 - H_t). \tag{4.10}$$

The problem of the representative household is to maximize

$$E_0 \sum_{t=0}^{\infty} \beta^t u(c_t, 1-n_t) \tag{4.11}$$

[19] We allow for the trend in G_t to ensure that a balanced growth path for the model economy exists.

subject to

$$\begin{aligned} & c_t + \gamma k_{t+1} - (1-\delta)k_t \\ = \; & (1-\tau_t)w_t n_t + (1-\tau_t)r_t k_t + \delta\tau_t k_t - \phi_t, \quad t \geq 0, \end{aligned}$$

(4.7), (4.8) and a given stochastic process for wage and rental rates. The maximization is by choice of contingency plans for $\{c_t, k_{t+1,} n_t\}$ over the elements of the household's time t information set that includes all model variables dated time t and earlier.

The firm maximizes its time t profits and its first order conditions imply

$$\begin{aligned} w_t &= (1-\alpha)\left(k_t/n_t\right)^{\alpha}, \text{ and} \\ r_t &= \alpha\left(n_t/k_t\right)^{1-\alpha}. \end{aligned} \tag{4.12}$$

Finally, the government's budget constraint can be written as

$$g_t = \tau_t w_t n_t + \tau_t(r_t - \delta)k_t + \phi_t.$$

We use the log-linearization procedure described by Christiano (1998) to solve for the competitive equilibrium of this economy. Given this solution, we can obtain the competitive equilibrium allocations for Y_t, C_t, and K_t as well as W_t and Φ_t using (4.4), (4.5), (4.9) and (4.12).

4.2. Model Calibration

In this subsection we briefly describe how we calibrated the model's parameter values. We assume that a time period in the model corresponds to one quarter and set $\beta = 1.03^{-1/4}$ and $\gamma = 1.004$. To evaluate the dependence of the model's implications on the assumed Frisch elasticity of labor supply we consider three values for μ. The first, $\mu = 0$, corresponds to the Hansen-Rogerson infinite elasticity case. The second, $\mu = 1$, implies the utility function for leisure is logarithmic, which is a common assumption in RBC studies. Combined with our assumption that the representative agent spends 24 percent of his time endowment working (see, for example, Christiano and Eichenbaum (1992)), this value corresponds to a Frisch elasticity of 3.16. Finally, we consider $\mu = 10$, which corresponds to a steady state Frisch elasticity of 0.33. The value of the Frisch elasticity of labor supply for males is estimated in the labor literature to be close to zero (see Card (1991), Killingsworth (1983) and Pencavel (1986)). Estimates of the Frisch labor supply elasticity for females typically falls in the range 0.5 to 1.5 (see for example, Killingsworth and Heckman (1986)). The parameter η was set to imply that in nonstochastic steady state the representative consumer spends 24% of his time endowment working. The rate of depreciation on capital δ was set 0.021 while α was set to 0.34 (see Christiano and Eichenbaum 1992).

Our specification for the j^{th} coefficient in the expansion of $h_1(L)$ and $h_2(L)$ is given by:

$$\begin{aligned} &h_{1,j}: \text{ estimated response of real government purchases} \\ &\text{at } t+j \text{ to the onset of a Ramey-Shapiro episode at time } t \\ &h_{2,j}: \text{ estimated response of average marginal income taxes} \\ &\text{at } t+j \text{ to the onset of a Ramey-Shapiro episode at time } t. \end{aligned} \tag{4.13}$$

Here $h_{i,j}$ denotes the coefficient on L^j in $h_i(L)$, $i = 1, 2$ and $j = 1, 2, \ldots, 50$. We refer to this as the distortionary tax specification. Figure 2 displays the first 16 coefficients of each impulse response function. Subject to the specification error entailed in approximating an infinite ordered polynomial with a finite number of lags, specification (4.13) ensures that the experiment being conducted in the model coincides with the experiment that we claim to have isolated in the data. Consequently, if our model has been specified correctly, the dynamic consequences of a shock to government purchases should be the same (aside from sampling uncertainty) in our model as in the data. We also consider a version of the model in which all taxes are lump sum. In this specification, $h_2(L)$ equals zero. We refer to this as the lump sum tax specification.

4.3. Quantitative Implications of the Model

Figure 3 displays the dynamic responses in the model of hours worked and real wages to a fiscal shock. Columns 1 and 2 report results generated using the lump sum and distortionary tax specifications, respectively. The solid lines in columns 1 and 2 display the estimated impulse response functions of hours worked and real wages. The measure of hours worked is private hours worked. In the lump-sum tax case, wages are before-tax real manufacturing wages. In the distortionary tax case, wages are after-tax real manufacturing wages.[20] The dotted lines correspond to model based impulse response functions for $\mu = \{0, 1, 10\}$.

4.3.1. The Response of Hours Worked

We begin by discussing the performance of the model with respect to n_t for the lump sum tax specification. First, notice that for all values of μ, n_t rises in response to a fiscal policy shock. This is because an increase in G_t raises the present value of the household's taxes and lowers its permanent income. Since leisure is a normal good, equilibrium hours worked rises.

Second, when labor supply is infinitely elastic ($\mu = 0$), the model does reasonably well at accounting for the large quantitative response of n_t. For example, the peak responses of n_t in the model and the data are 4.3% and 6.2%, respectively. The performance of the model deteriorates for lower elasticities of labor supply (higher values of μ). For example, when the labor supply elasticity equals 0.33 ($\mu = 10$), the peak rise in n_t is only 0.7%, roughly 11% of the estimated peak response of n_t in the data. The basic intuition for this result is as follows. The larger is μ the more the household wishes to smooth hours worked. Since hours worked do not change in steady state, the household finds it optimal to respond to a rise in the present value of its taxes by reducing private consumption by relatively more as μ becomes larger and varying hours worked less.

[20] The estimated impulse response functions for the distortionary tax specification are reproduced from Figure 2.

Third, for all versions of μ, the model reproduces the fact that an exogenous fiscal policy shock leads to a prolonged rise in n_t and a positive conditional correlation between n_t and G_t. In the data this correlation is 0.93. In the model this correlation is 0.76, 0.78 and 0.80 for μ equal to 0, 1 and 10, respectively.

Next we consider the effect of a fiscal policy shock on hours worked with the distortionary tax specification. The key difference relative to the lump sum tax specification pertains to the shape of the dynamic response function of n_t and the sign of the implied conditional correlation between n_t and G_t. In the lump sum tax specification, a fiscal shock leads to a long persistent rise in n_t and a positive conditional correlation between n_t and G_t. In the present case, n_t initially rises but then declines relative to its pre-shock value, around 4 periods after the shock, three periods before both G_t and τ_t peak. The fall in n_t is an increasing function of the elasticity of labor supply, going from less than 1% when $\mu = 10$ to about 8% when $\mu = 0$. But regardless of the absolute magnitude of the movements in n_t, hours worked rises when G_t is relatively low and falls when G_t is relatively high. As a consequence, for all values of μ, the model generates a sharp negative conditional correlation between n_t and G_t, equal to -0.80, -0.85 and -0.85 for $\mu = 10$, 1 and 0, respectively. This negative correlation contrasts sharply with the strong positive conditional correlation between these variables in the data. Notice that allowing for a high labor supply elasticity does not help the model's performance. At least with respect to the conditional correlation of n_t and G_t, the mismatch between theory and data gets worse as we move from low to high labor supply elasticities.

The intuition for this result can be described as follows. Other things equal, a higher value of τ_t gives rise to an intratemporal effect which induces the household to shift its period t allocation of time towards leisure. In addition the hump-shaped pattern of the rise in τ_t gives rise to an intertemporal effect which induces the household to shift n_t towards periods in which τ_t is relatively low. Since τ_t moves by relatively small amounts in the first few periods after the fiscal shock, the initial intratemporal effects of the tax rate changes are small. Given the intertemporal effect of future rises in τ_t, the initial rises in n_t are slightly larger than in the lump sum tax case. As marginal tax rates begin to rise significantly, the intratemporal effect becomes quantitatively important and the responses of n_t in the lump sum and distortionary tax rate models become quite different. In the former case, n_t stays above its pre shock level for over 12 periods. In the latter cases, n_t falls below its pre shock level after 4 periods, and remains below that level for over 6 periods.

The higher is the elasticity of labor supply (the lower is μ) the more the household is willing to intertemporally substitute n_t over time. So the swing in n_t from a large initial positive response to a large negative response in the distortionary tax case is more pronounced the higher is the elasticity of labor supply. This in turn exacerbates the counterfactual negative conditional correlation between n_t and G_t that arises when the model is confronted with empirically plausible movements in tax rates.

We conclude this subsection by reporting the results of formally testing the model's ability to account for various conditional moments of the data using the J statistic defined in (3.8). The first two moments that we consider pertain to the maximal response of hours worked in the aftermath of a Ramey-Shapiro episode: $R_1(n)$ and $R_2(n)$ denote the peak rise in n_t and the average response of n_t in periods 4 through 7 after a fiscal policy shock. The values of these moments for the model and the data, $R_i^m(n)$ and $R_i^d(n)$, $i = 1, 2$, respectively, were calculated using estimates of the relevant dynamic response functions. The third moment that we consider is the correlation between g_t and n_t, $\rho(g, n)$, induced by a fiscal policy shock. We let $\rho^m(g, n)$ and $\rho^d(g, n)$ denote the values of this moment implied by the model and the data, respectively. The final moment, σ_n, is the standard deviation of hours worked induced by the onset of a Ramey-Shapiro shock. Below, σ_n^m and σ_n^d denote the values of this moment implied by the model and the data, respectively.[21]

[21] We calculated the last two moments as follows. Let the actual and model implied dynamic response function of a variable x_t to a fiscal policy shock be given by $H_x(L)D_t$ and $\tilde{H}_x(L)D_t$,

Table 1 reports the results of testing the individual hypotheses: $R_i^d(n) - R_i^m(n) = 0$, $i = 1, 2$, $\rho^d(g, n) - \rho^m(g, n) = 0$, and $\sigma_n^d - \sigma_n^m = 0$. Consider first our results for the lump sum tax specification. Note that when $\mu = 0$ or 1 there is little evidence against any of the hypotheses being investigated. However when $\mu = 10$ the hypotheses that $R_i^d(n) - R_i^m(n) = 0$, $i = 1, 2$, and $\sigma_n^d - \sigma_n^m = 0$ can each be rejected at the 1% significance level. So with lump sum taxes and values for the elasticity of labor supply often used in the RBC literature ($\mu = 0$ or 1), our results formalize claims in the literature that standard RBC models can account for the key features of how hours worked responds to an exogenous shock in government purchases.

Next consider the model with the distortionary tax specification. For all values of μ, we can reject the hypothesis $\rho^d(g, n) - \rho^m(g, n) = 0$ at the 1% significance level. High labor supply elasticity versions of the model can account for the peak response and volatility of n_t: there is little evidence against the hypotheses $R_1^d(n) - R_1^m(n) = 0$ or $\sigma_n^d - \sigma_n^m = 0$ when μ equals 0 or 1. However the model accounts for these features of the data in a way that is inconsistent with the *timing* of the actual movements in n_t. This manifests itself in the failure of the model to reproduce the estimated sign of the conditional correlation between G_t and n_t. A related failing is the model's inability to reproduce the sign of the average response of n_t during periods 4 through 7 after the shock. In the data, n_t *rises* by an average 5.4% over these time periods. In contrast, for all values of μ, the model implies that n_t *falls* on average over these time periods. Not surprisingly, we can reject the hypothesis that $R_2^d(n) - R_2^m(n) = 0$ at the 1% significance level for all versions of the model. As we increase the elasticity of labor supply, this mismatch between the model and the data gets worse. As μ falls from 10 to 0,

respectively, $x_t = \{n_t, g_t\}$. The value of σ_x implied by the model and in the data is given by $\sigma_n^m = \{\sum_{i=0}^{\infty}[\tilde{H}_n(i)]^2\}^{1/2}$ and $\sigma_x^d = \{\sum_{i=0}^{\infty}[H_x(i)]^2\}^{1/2}$, respectively. Here $H_x(i)$ and $\tilde{H}_x(i)$ denote the i^{th} coefficients in the polynomial lag operators $H_x(L)$ and $\tilde{H}_x(L)$. The values of $\rho(g_t, n_t)$ implied by the model and in the data are given by $\rho^m(g, n) = [\sum_{i=0}^{\infty}\tilde{H}_n(i)H_g(i)]/\sigma_n^m\sigma_g^d$ and $\rho^d(g, n) = [\sum_{i=0}^{\infty}H_n(i)H_g(i)]/\sigma_n^d\sigma_g^d$, respectively. Note that the value of σ_g in the model is equal to σ_g^d by construction. In practice we calculated σ_n^m, σ_n^d, $\rho^m(g, n)$, $\rho^d(g, n)$ and σ_g^d using the first 12 coefficients of the relevant dynamic response functions.

the average response of n_t during periods 4 through 7 goes from -0.5% to -5.4%. As above, we conclude that the traditional salve of technology shock driven RBC models–a high elasticity of labor supply–exacerbates the failure of the model to reproduce the timing of the response of n_t to a fiscal policy shock.

4.3.2. The Response of Real Wages

We now turn to a brief discussion of the model's implications for real wages. Consider first the model under the lump sum tax specification.[22] Notice that real wages fall for all values of μ. This reflects the fact that hours worked rises and the marginal product of labor is a decreasing function of n_t. Since the rise in n_t is an increasing function of the elasticity of labor supply, real wages fall by more the higher is that elasticity. However, regardless of which value of μ we assume, the model substantially understates the decline in real wages.

This basic pattern of results carries over to the distortionary tax specification with one interesting difference. Up to around period 4, after-tax real wages fall by more the higher is the elasticity of labor supply. But thereafter real wages fall by more the lower is the elasticity of labor supply. To understand this result recall first that, unlike the lump sum tax case, the impulse response functions being graphed pertain to after-tax real wages, $(1 - \tau_t)w_t$.[23] Next recall that τ_t does not move by very much in the first few periods after a fiscal shock. So in those periods the model responds in roughly the same way as it does in the lump sum tax case: the initial rise in n_t and the decline in real wages is an increasing function of the elasticity of labor supply.

As the rise in τ_t becomes more pronounced, real wages respond quite differently in the lump sum and distortionary tax rate cases. This reflects the different

[22] Recall that the solid line in Figure 3 pertaining to this case displays the estimated response of before tax real wages to a fiscal policy shock.

[23] The estimated declines in after tax real wages are larger than those of before tax real wages. This is because the former reflect both rises in τ_t and declines in w_t. See Edelberg, Eichenbaum and Fisher (1999) for a discussion of the response of different measures of before tax real wages to the onset of a Ramey - Shapiro episode, estimated under the equal intensity assumption.

patterns in the response of hours worked in the two cases. When taxes are lump sum n_t rises in response to the shock and declines to its unchanged steady state value from above. In the distortionary tax case, when μ is equal to zero or one, n_t falls below its initial pre shock level around the time of the peak rise in τ_t and approaches its steady state value from below. The ongoing declines in n_t are larger the higher is the elasticity of labor supply and exert correspondingly stronger upward pressure on $(1-\tau_t)w_t$. But the rises in τ_t, which are common across all values of μ, exert direct negative pressure on $(1-\tau_t)w_t$. The net result is that after period 4, the declines in after-tax wages are largest for the low labor supply elasticity versions of the model. So the low elasticity version of the model actually does a better job of accounting for the steep estimated decline in after-tax wages after period 4. But it does so only because it counterfactually predicts a very small response of hours worked.

Table 2 summarizes the results of formally testing the analog hypotheses to those reported in Table 1. The first two moments considered pertain to the maximal response of real wages in the aftermath of a Ramey-Shapiro episode: $R_1(w)$and $R_2(w)$ denote the maximal declines in real wages and the average response of real wages in periods 4 through 7 after a fiscal policy shock. The third moment which we consider is the correlation between g_t and w_t, $\rho(g,w)$, induced by a fiscal policy shock. The final moment, σ_w, is the standard deviation of real wages induced by the onset of a Ramey-Shapiro episode. In the lump sum and distortionary tax cases, the real wage measure pertains to before and after tax real wages, respectively.

Two features of Table 2 stand out. First, consistent with the data, all of the models generate a negative correlation between real wages and government purchases. In the lump sum tax specification there is very little evidence against the hypothesis that $\rho^d(g,w) - \rho^m(g,w) = 0$ when μ is equal to 0 or 1. In the distortionary tax specification there is little evidence against this hypothesis when μ equals 1 or 10. Note that all versions of the model do relatively well with respect to the correlation between G_t and w_t. But they also all do a poor job of accounting for the conditional volatility of real wages and their peak declines after a fiscal policy shock. For both the lump sum and distortionary tax specifications, we can reject the hypotheses that $R_1^d(w) - R_1^m(w) = 0$ and $\sigma_w^d - \sigma_w^m = 0$ at the 1% significance level. This is true regardless of which value of μ we consider. However the hypothesis that $\sigma_w^d - \sigma_w^m = 0$ is only marginally rejected at the 1% significance level for the high elasticity version ($\mu = 0$) of the model with the lump sum tax specification.

Based on the results of this section, we conclude that there is relatively little evidence against very high elasticity versions of the standard RBC model if we assume that tax rates are unaffected by fiscal policy shocks. But once we abandon this counterfactual assumption, there is overwhelming evidence against the notion that standard RBC models can account for the effects of a fiscal policy shock, regardless of what value we assume for the elasticity of labor supply.

5. Sensitivity Analysis

Above we argued that a key failure of standard RBC models is their counterfactual prediction that a fiscal policy shock induces a negative correlation between n_t and G_t. The mismatch between theory and data is worse the higher is the assumed elasticity of labor supply, i.e. the lower is μ. This is because the lower μ is, the more willing the household is to reduce private consumption and lower hours worked in periods when the tax rate on labor income is high.

This suggests that the model's quantitative performance would be improved if, for a given labor supply elasticity, the household was less willing to intertemporally substitute consumption. To investigate this possibility we redo our analysis in a version of the model in which (4.1) is replaced by,

$$E_0 \sum_{t=0}^{\infty} \beta^t \left[\frac{1}{1-\sigma}(C_t^{1-\sigma} - 1) + \frac{\eta}{1-\mu}(1-n_t)^{1-\mu} \right]. \tag{5.1}$$

for various values of σ. [24] For $\sigma = 1$, this specification coincides with the bench-

[24] To accomodate balanced growth, η must grow at the rate $\gamma^{1-\sigma}$.

mark model. We confine our analysis here to the distortionary tax specification.

Tables 3 and 4 report results analogous to those presented in Tables 1 and 2 for $\sigma = 5$.[25] Comparing Table 1 and Table 3 we see that, as expected, setting σ to 5 magnifies the peak response and volatility of hours worked. Nevertheless we still reject hypotheses the $R_1^d(n) - R_1^m(n) = 0$ and $\sigma_n^d - \sigma_n^m = 0$ at roughly the 1% significance level for the low labor supply elasticity version of the model.

Setting $\sigma = 5$ succeeds in mitigating the decline in average hours worked in periods 4 through 7 after the shock. It also raises the conditional correlation between G_t and n_t. Nevertheless for the higher labor elasticity versions of the model ($\mu = 0$ and 1) it is still the case that (i) n_t falls in periods 4 through 7, (ii) the conditional correlation between G_t and n_t is negative and (iii) hypothesis $\rho^d(g, w) - \rho^m(g, w) = 0$ is rejected at the 1% significance level.

Turning to the low labor supply elasticity version of the model ($\mu = 10$), we see that with $\sigma = 5$, the average response of n_t in periods 4 through 7 is positive (although small) and the conditional correlation between G_t and n_t is positive. In fact we can no longer reject the hypothesis $\rho^d(g, w) - \rho^m(g, w) = 0$ at conventional significance levels. Nevertheless we can still reject hypothesis $R_1^d(n) - R_1^m(n) = 0$ at the 1% significance level.

Finally, turning to the model's implications for real wages, we see from Table 4, that all versions of the model continue to substantially understate the peak response and overall volatility of real wages. We conclude that lowering the degree of intertemporal substitution in consumption does not cure the model's shortcomings.[26]

One cannot rule out the possibility that there exists some specification of

[25] We experimented with lower values of σ and found that the lower σ is the more similar our results are to those obtained with the benchmark model.

[26] As it turns out the modest improvement in the model's performance when $\sigma = 5$ comes at a large cost: the model predicts that hours worked *declines* after a positive technology shock. The ability of RBC models to account for the observed volatility of hours worked and output rests critically on the assumption that business cycles are driven primarily by exogenous technology shocks. Unfortunately, in technology shock driven RBC models, high values of σ give rise to the grossly counterfactual implication that hours worked are negatively correlated with output.

preferences that would render the model consistent with the data. While general statements are not possible, we did explore a variety of alternative specifications: (i) we modified (5.1) to allow for habit formation in consumption, and (ii) we analyzed a version of the model in which the household's period utility function is CES in consumption and leisure. We found that with the first modification and CES specifications in which consumption and leisure are highly complementary, the model's performance was very similar to the $\sigma = 5$ case discussed above. Evidently simple respecifications of the household's preferences are not sufficient to correct the model's shortcomings.

What of simple changes to the production technology or market structure? Again definitive statements are not possible. But we did implement a version of our model which allowed for externalities in production. In particular, we modified (4.4) to be of the form

$$Y_t \leq K_t^{\alpha}(X_t n_t)^{1-\alpha}\bar{Y}_t^{\kappa}, \quad 0 < \alpha < 1, \kappa > 0$$

where $\bar{Y}_t$ denotes the economy wide level of output. This type of model has been considered by Baxter and King (1991) and is known to be isomorphic to a setup of the type considered by Devereux, Head and Lapham (1996) in which imperfectly competitive firms face increasing returns to scale at the firm level. We considered various values of κ. Tables 3 and 4 reports results for $\kappa = 0.3$. Note that the main effects of this perturbation are to raise the peak response and overall volatility of hours work while exacerbating the failure of the standard model to match the decline in real wages. It does little to affect the negative conditional correlation between G_t and n_t. So this variant of the model also fails to account for the facts.

Finally we modified (4.4) to be of the form

$$Y_t \leq K_t^{\alpha}(X_t n_t)^{1-\alpha}\bar{Y}_{t-1}^{\kappa}, \quad 0 < \alpha < 1, \kappa > 0$$

so that there are dynamic externalities in production. With this specification, an increase in hours worked raises future returns to working thus rendering it more likely that a fiscal policy shock generates a hump shaped increase in n_t.

The results for this model are also reported in Tables 3 and 4, for a value of $\kappa = 0.5$. This was the lowest value of κ for which one version of the model was consistent with the observed behavior of hours worked. Note that with $\mu = 0$, none of the hypotheses regarding the model's implications for hours worked can be rejected at conventional significance levels. In this sense, this model perturbation is successful. But it is difficult to claim success. First, we know of no evidence to support such a high value of κ. Second, it is still the case that the model generates implications for real wages that are easily rejected (see Table 4). We conclude that simple perturbations of technology and market structure of the sort that we investigated do not render the model consistent with the data.

6. Conclusion

This paper investigated the effects of a fiscal policy shock on hours worked and real wages. A key feature of our analysis is that we explicitly allow for movements in average marginal tax rates as well as government purchases. An important finding of the paper is that movements in tax rates, hours worked and government purchases track each other: all display persistent hump shaped patterns in the wake of a Ramey-Shapiro episode. This pattern of comovement substantially affects inference about the ability of standard RBC models to account for the effects of a fiscal policy shock. Once tax effects are taken into account the model counterfactually implies that, after a fiscal policy shock, hours worked are negatively correlated with government purchases. No doubt the models' implications would be improved if tax rates initially rose and then fell after a fiscal policy shock. But this is not what we observe in the data.

While we focused our analysis on one sector versions of the RBC model, our methodology is applicable to a much broader class of models. Ramey and Shapiro (1998) show that various two sector versions of the RBC model generate predictions for aggregate hours worked and real wages that are very similar to those of the one sector model. So presumably these models too would fail our diagnostic test. Burnside, Eichenbaum and Fisher (1999) show that a variant of Alexopolous' (1998) efficiency wage model also fails our test. Rotemberg and Woodford (1992) and Devereux, Head and Lapham (1996) study the effects of changes in government purchases in stochastic general equilibrium models with increasing returns and oligopolistic pricing. Their models predict that real wages rise after an exogenous increase in government purchases. Since real wages actually fall after such a shock, our results suggest that these types of models would also do poorly with respect to our test.

We are left with the question: what causes agents to work hard when real wages are low and both government purchases and tax rates are high?

7. Data Appendix

Our data are from four main sources. Below we list the series which correspond to each of these sources. All series are seasonally adjusted except for interest rates and taxes. Most of these series were obtained by us from the Federal Reserve Board's macroeconomic database. Where possible we provide the mnemonic for the same series from the commercially available DRI BASIC Economics Database.

1. Bureau of Economic Analysis. GDP (GDPQ), Defense spending (GGFEQ), Government purchases (Defense spending plus Federal, State and Local consumption expenditures) (GGFEQ+GGOCEQ+GGSCPQ). These series are all in units of 1992 chain-weighted dollars.

2. Bureau of Labor Statistics. Index of hours of all persons in the business sector (LBMN), Manufacturing wages (LEHM), Consumer price index for all urban consumers (PUNEW), Producer price index for crude fuel in manufacturing industries (PW1310).

3. Board of Governors of the Federal Reserve System. Net 3 month Treasury Bill secondary market interest rate (FYGM3).

4. Stephenson (1998), Table 1, pp. 391-392. Updated version of Barro and Sahasakul's (1983) income-weighted measure of average marginal statutory income tax rate. Annual data linearly interpolated to quarterly frequency.

References

[1] Alexopoulos, M. (1998) "Efficiency Wages, Unemployment and the Business Cycle," manuscript, Northwestern University.

[2] Barro, R.J. and C. Sahasakul. "Measuring the Average Marginal Tax Rate from the Individual Income Tax." *Journal of Business*, October 1983, *56*(4), pp. 419–52.

[3] Baxter, M. and R.G. King. "Fiscal Policy in General Equilibrium." *American Economic Review*, 1993, *83*, pp. 315–34.

[4] Baxter, M. and R.G. King. "Productive Externalities and Business Cycles." Discussion Paper 53, Federal Reserve Bank of Minneapolis, 1991.

[5] Blanchard, O. and R. Perotti. "An Empirical Characterization of the Dynamic Effects of Changes in Government Spending and Taxes on Output." Working paper, MIT, 1998.

[6] Burnside, C. and M. Eichenbaum. "Factor Hoarding and the Propagation of Business Cycle Shocks." *American Economic Review*, December 1996, *86*(5), pp. 1154–74.

[7] Burnside, C., M. Eichenbaum and J. Fisher. "Fiscal Shocks in an Efficiency Wage Model." Working paper, Northwestern University, 1999.

[8] Burnside, C., M. Eichenbaum and S. Rebelo. "Labor Hoarding and the Business Cycle." *Journal of Political Economy*, April 1993, *101*(2), pp. 245–73.

[9] Card, D. "Intertemporal Labor Supply: An Assessment." NBER Working Paper No. 3602, 1991.

[10] Christiano, L.J. "Solving Dynamic Equilibrium Models by a Method of Undetermined Coefficients." NBER Technical Working Paper, No. 225, 1998.

[11] Christiano, L.J. and M. Eichenbaum. "Current Real Business Cycle Theories and Aggregate Labor Market Fluctuations." *American Economic Review*, June 1992, *82*(3), pp. 430–50.

[12] Christiano, L.J., M. Eichenbaum and C. Evans. "Sticky Price and Limited Participation Models of Money: A Comparison." *European Economic Review*, June 1997, *41*(6), pp. 1201–49.

[13] Christiano, L.J., M. Eichenbaum and C. Evans. "Modelling Money." NBER Working Paper No. 6371, 1998.

[14] Christiano, L.J., M. Eichenbaum and C. Evans. "Monetary Policy Shocks: What Have We Learned and to What End?" Forthcoming in M. Woodford and J. Taylor, eds. *Handbook of Monetary Economics*, 1999.

[15] Devereux, M.B., A.C. Head and B.M. Lapham. "Monopolistic Competition, Increasing Returns, and the Effects of Government Spending." *Journal of Money, Credit and Banking*, May 1996, *28*(2), pp. 233–54.

[16] Edelberg, W., M. Eichenbaum and J. Fisher. "Understanding the Effects of Shocks to Government Purchases." *Review of Economics Dynamics*, 1999, pp. 166–206.

[17] Eichenbaum, M., L.P. Hansen and K.J. Singleton. "A Time Series Analysis of Representative Agent Models of Consumption and Leisure Under Uncertainty." *Quarterly Journal of Economics,* 1988, pp. 51–78.

[18] Hansen, G. "Indivisible Labor and the Business Cycle." *Journal of Monetary Economics*, November 1985, *16*(3), pp. 309–28.

[19] Heckman, J. and M. Killingsworth "Female Labor Supply: A Survey." In Ashenfelter, O. and R. Layard, eds. *Handbook of Labor Economics.* Volume 1. Amsterdam: North-Holland, 1986, pp. 103–204.

[20] Killingsworth, M. *Labor Supply.* Cambridge: Cambridge University Press, 1993.

[21] Mulligan, C. B. "Pecuniary Incentives to Work in the United States during World War II." *Journal of Political Economy*, October 1998, *106*(5), pp. 1033–77.

[22] Newey, W. and K. West. "A Simple, Positive Semi-Definite Heteroskedasticity and Autocorrelation Consistent Covariance Matrix." *Econometrica*, May 1987, *55*, pp. 703–8.

[23] Ohanian, L.E. "The Macroeconomic Effects of War Finance in the United States: World War II and the Korean War." *American Economic Review*, March 1997, *87*(1), pp. 23–40.

[24] Pencavel, J. "Labor Supply of Men: A Survey." Ashenfelter, O. and R. Layard, eds. *Handbook of Labor Economics.* Volume 1. Amsterdam: North-Holland, 1986, pp. 3–102.

[25] Ramey,V. and M.D. Shapiro. "Costly Capital Reallocation and the Effects of Government Spending." *Carnegie Rochester Conference Series on Public Policy*, June 1998, Volume *48* pp. 145–94.

[26] Rogerson, R. "Indivisible Labor, Lotteries and Equilibrium." *Journal of Monetary Economics*, January 1998, *21*(1), pp. 3–16.

[27] Rotemberg, J. and M. Woodford. "Oligopolistic Pricing and the Effects of Aggregate Demand on Economic Activity." *Journal of Political Economy,* 1992, *100*, pp. 1153–297.

[28] Seater, J. "On the Construction of Marginal Federal Personal and Social Security Tax Rates in the U.S." *Journal of Monetary Economics*, January 1985, *15*(1), pp. 121–35.

[29] Sims, C. "Macroeconomics and Reality." *Econometrica*, January 1980, *48*(1), pp. 1–48.

[30] Stephenson, E.F. "Average Marginal Tax Rates Revisited." *Journal of Monetary Economics*, April 1998, *41*(2), pp. 389–409.

Table 1. Goodness-of-fit Tests – Response of Hours Worked in the Standard RBC Model

Moment	Lump sum tax specification			Distortionary tax specification		
	$\mu = 0$	$\mu = 1$	$\mu = 10$	$\mu = 0$	$\mu = 1$	$\mu = 10$
Peak						
Data	6.18	6.18	6.18	6.18	6.18	6.18
Model	4.33	2.75	0.68	5.02	2.71	0.59
J-statistic	0.67	2.97	9.20	0.25	3.01	9.37
P-value	0.41	0.09	0.002	0.62	0.08	0.002
Average of periods 4, 5, 6, 7						
Data	5.40	5.40	5.40	5.40	5.40	5.40
Model	3.78	2.45	0.63	-5.36	-2.86	-0.48
J-statistic	0.54	2.30	7.35	5.95	6.78	9.50
P-value	0.46	0.13	0.007	0.01	0.009	0.002
Correlation with government purchases						
Data	0.93	0.93	0.93	0.93	0.93	0.93
Model	0.76	0.78	0.80	-0.85	-0.85	-0.80
J-statistic	0.31	0.22	0.14	29.9	20.9	11.8
P-value	0.57	0.64	0.71	<0.001	<0.001	<0.001
Standard deviation						
Data	13.1	13.1	13.1	13.1	13.1	13.1
Model	12.3	8.03	2.06	20.2	10.5	1.87
J-statistic	0.02	1.15	7.74	0.42	0.12	6.72
P-value	0.89	0.28	0.005	0.52	0.73	0.009

Table 2. Goodness-of-fit Tests – Response of Real Wages in the Standard RBC Model

Moment	Lump sum tax specification $\mu = 0$	Lump sum tax specification $\mu = 1$	Lump sum tax specification $\mu = 10$	Distortionary tax specification $\mu = 0$	Distortionary tax specification $\mu = 1$	Distortionary tax specification $\mu = 10$
Peak						
Data	-5.53	-5.53	-5.53	-10.1	-10.1	-10.1
Model	-1.35	-0.87	-0.70	-2.72	-3.66	-4.78
J-statistic	13.1	17.4	18.2	31.0	29.7	19.9
P-value	$<$0.001	$<$0.001	$<$0.001	$<$0.001	$<$0.001	$<$0.001
Average of periods 4, 5, 6, 7						
Data	-4.94	-4.94	-4.94	-8.87	-8.87	-8.87
Model	-0.98	-0.67	-0.22	-2.01	-2.96	-3.82
J-statistic	13.3	16.0	19.2	28.0	30.3	23.5
P-value	$<$0.001	$<$0.001	$<$0.001	$<$0.001	$<$0.001	$<$0.001
Correlation with government purchases						
Data	-0.94	-0.94	-0.94	-0.97	-0.97	-0.97
Model	-0.74	-0.81	-0.81	-0.88	-0.96	-0.98
J-statistic	1.18	1.92	6.49	11.6	0.16	0.02
P-value	0.28	0.17	0.01	$<$0.001	0.69	0.89
Standard deviation						
Data	15.8	15.8	15.8	26.9	26.9	26.9
Model	3.54	2.52	1.30	7.42	9.65	11.9
J-statistic	7.02	19.4	22.8	26.3	30.5	24.9
P-value	0.008	$<$0.001	$<$0.001	$<$0.001	$<$0.001	$<$0.001

Table 3. Goodness-of-fit Tests – Response of Hours Worked with Alternative Model Specifications

	$\sigma = 5$			Static Increasing Returns			Dynamic Increasing Returns		
Moment	$\mu = 0$	$\mu = 1$	$\mu = 10$	$\mu = 0$	$\mu = 1$	$\mu = 10$	$\mu = 0$	$\mu = 1$	$\mu = 10$
Peak									
Data	6.18	6.18	6.18	6.18	6.18	6.18	6.18	6.18	6.18
Model	6.78	4.07	1.33	10.1	3.71	0.63	7.37	4.37	0.69
J-statistic	0.05	1.01	6.89	1.16	1.43	9.20	0.17	0.69	11.3
P-value	0.82	0.31	0.009	0.28	0.23	0.002	0.68	0.41	<0.001
Average of periods 4, 5, 6, 7									
Data	5.40	5.40	5.40	5.40	5.40	5.40	5.40	5.40	5.40
Model	-3.80	-1.63	0.23	-6.69	-3.37	-0.48	5.20	1.03	-0.32
J-statistic	5.35	5.49	7.08	5.09	6.76	9.44	0.004	2.66	9.02
P-value	0.02	0.02	0.008	0.02	0.009	0.002	0.95	0.10	0.003
Correlation with government purchases									
Data	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93
Model	-0.70	-0.59	0.22	-0.74	-0.81	-0.77	0.47	-0.22	-0.65
J-statistic	21.3	11.4	1.30	38.9	25.8	11.8	1.76	7.45	8.42
P-value	<0.001	<0.001	0.25	<0.001	<0.001	<0.001	0.18	0.006	0.004
Standard deviation									
Data	13.1	13.1	13.1	13.1	13.1	13.1	13.1	13.1	13.1
Model	18.5	9.49	2.46	33.8	13.9	1.98	15.9	11.9	1.98
J-statistic	0.36	0.35	6.49	2.13	0.01	6.58	0.15	0.03	4.60
P-value	0.55	0.56	0.01	0.14	0.92	0.01	0.70	0.86	0.03

Table 4. Goodness-of-fit Tests – Response of Real Wages with Alternative Model Specifications

	$\sigma = 5$			Static Increasing Returns			Dynamic Increasing Returns		
Moment	$\mu = 0$	$\mu = 1$	$\mu = 10$	$\mu = 0$	$\mu = 1$	$\mu = 10$	$\mu = 0$	$\mu = 1$	$\mu = 10$
Peak									
Data	-10.1	-10.1	-10.1	-10.1	-10.1	-10.1	-10.1	-10.1	-10.1
Model	-3.07	-4.02	-5.07	-3.35	-4.26	-4.94	-0.94	-3.66	-4.94
J-statistic	30.3	26.6	18.2	34.5	24.2	16.6	30.2	21.3	11.6
P-value	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
Average of periods 4, 5, 6, 7									
Data	-8.87	-8.87	-8.87	-8.87	-8.87	-8.87	-8.87	-8.87	-8.87
Model	-2.47	-3.35	-4.10	-1.85	-3.16	-3.88	-0.20	-1.18	-3.52
J-statistic	27.0	26.2	21.1	39.2	31.8	21.4	36.4	37.2	21.2
P-value	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
Correlation with government purchases									
Data	-0.97	-0.97	-0.97	-0.97	-0.97	-0.97	-0.97	-0.97	-0.97
Model	-0.87	-0.95	-0.98	-0.87	-0.96	-0.97	-0.35	-0.82	-0.96
J-statistic	17.5	0.72	0.04	0.50	0.02	0.004	1.14	0.19	0.005
P-value	<0.001	0.39	0.83	0.48	0.88	0.95	0.29	0.66	0.94
Standard deviation									
Data	26.9	26.9	26.9	26.9	26.9	26.9	26.9	26.9	26.9
Model	8.97	10.9	12.8	7.75	10.8	12.3	1.68	6.96	12.2
J-statistic	25.4	26.1	22.3	35.9	30.3	21.6	15.5	14.1	11.0
P-value	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001

Figure 1: Fiscal Variables and the Ramey-Shapiro Dates, 1947-1994

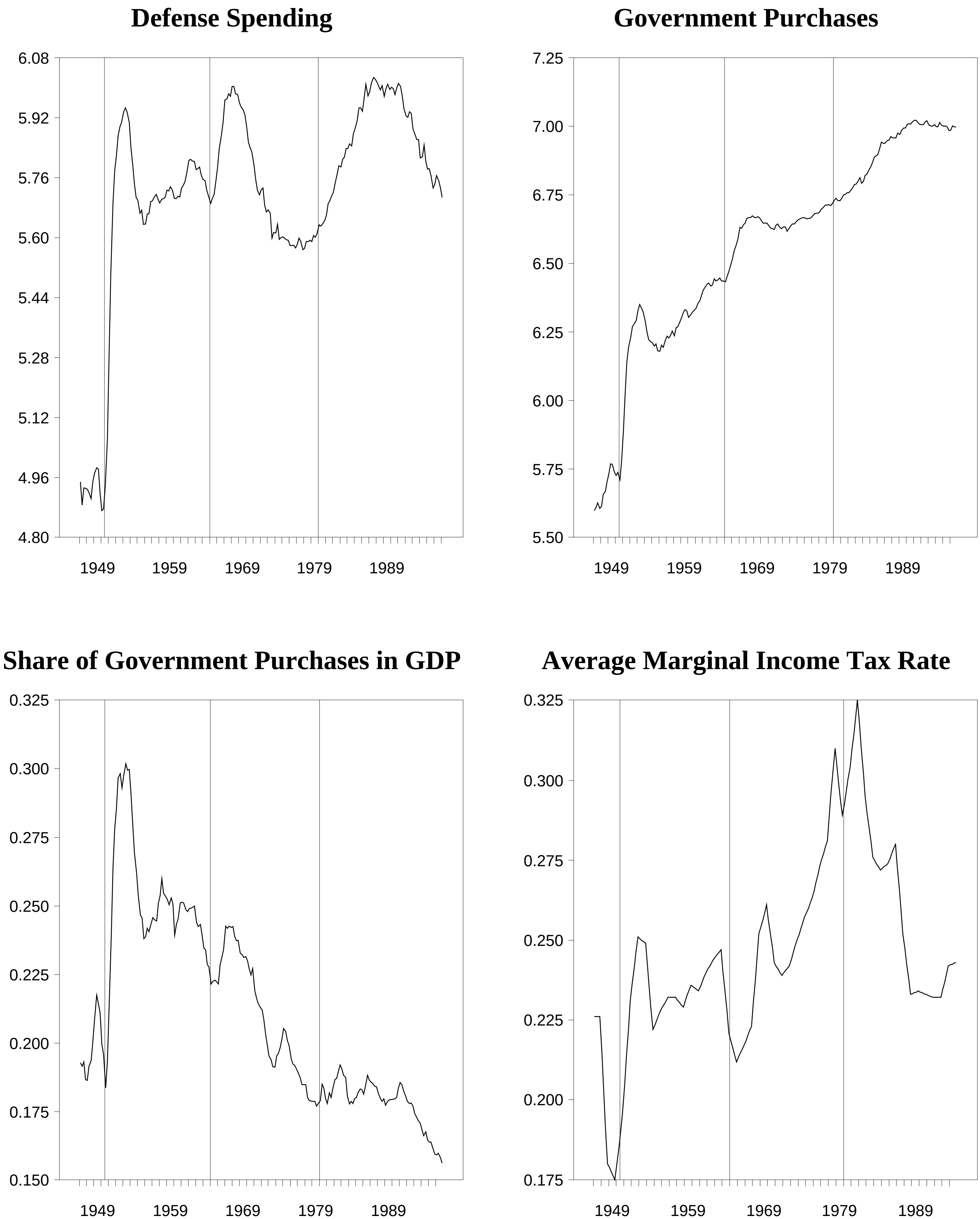

Figure 2: Estimated Responses of Fiscal and Labor Market Variables After a Ramey-Shapiro Episode

Figure 3: Labor Market Dynamics After a Ramey-Shapiro Episode in the Standard Model

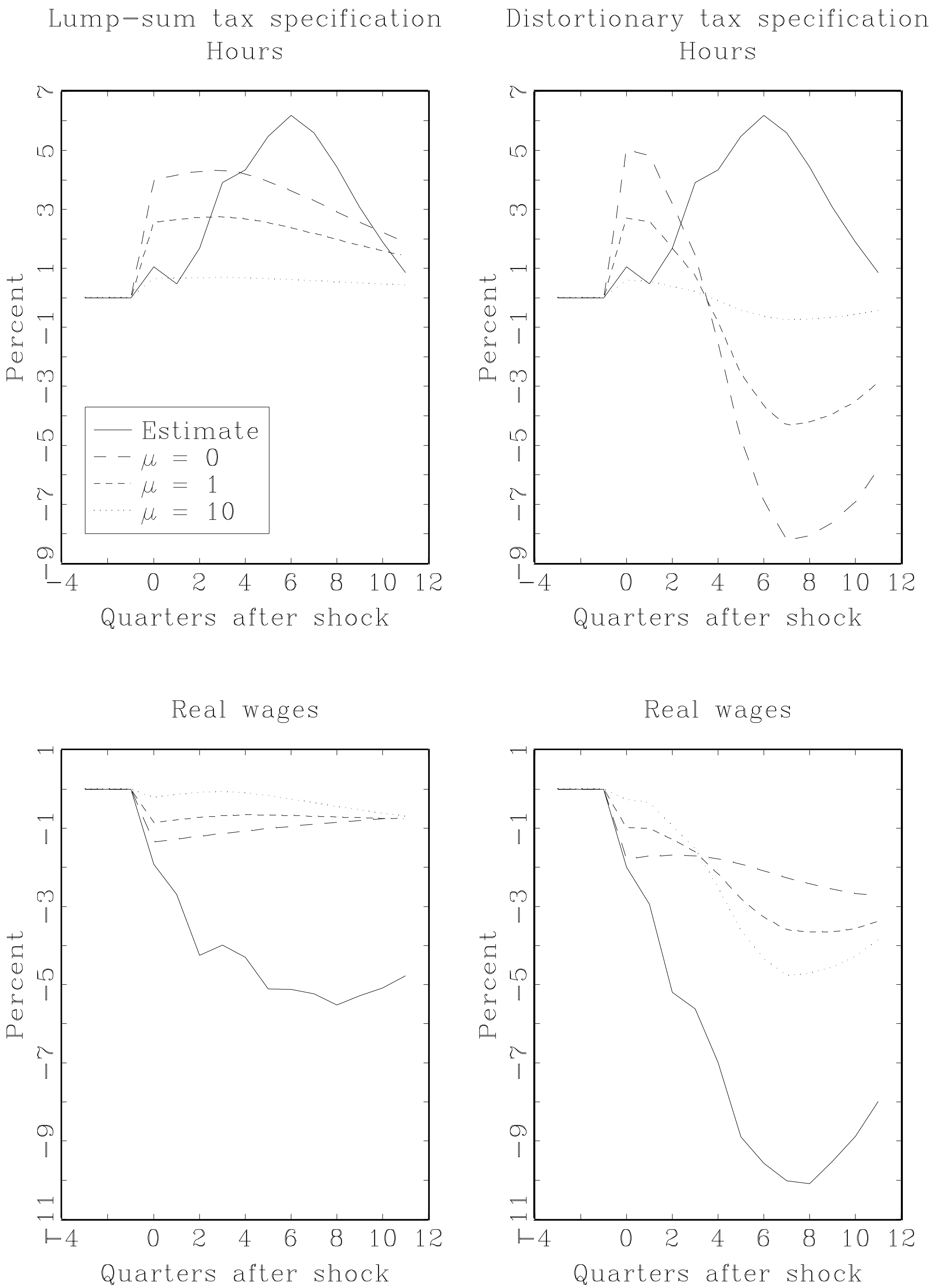

Figure 4: Labor Market Dynamics After a Ramey-Shapiro Episode in the Alternative Models