Model Uncertainty and Intertemporal Tax Smoothing

Yulei Luo, Jun Nie and Eric R. Young

Model Uncertainty and Intertemporal Tax Smoothing[*]

Yulei Luo[†]

The University of Hong Kong

Jun Nie[‡]

Federal Reserve Bank of Kansas City

Eric R. Young[§]

University of Virginia

April 3, 2012

Abstract

In this paper we examine how model uncertainty due to the preference for robustness (RB) affects optimal taxation and the evolution of debt in the Barro tax-smoothing model (1979). We first study how the government spending shocks are absorbed in the short run by varying taxes or through debt under RB. Furthermore, we show that introducing RB improves the model's predictions by generating (i) the observed relative volatility of the changes in tax rates to government spending, (ii) the observed comovement between government deficits and spending, and (iii) more consistent behavior of government budget deficits in the US economy.

Keywords: Robustness, Model Uncertainty, Taxation Smoothing, Taxation Tilting.

JEL Classification Numbers: D83, E21, F41, G15.

1 Introduction

Figure 1 presents the post-war behavior of US government spending, average tax rates, and government debt. In this paper we ask how to rationalize the joint behavior of these fiscal variables; specifically, we ask whether they are jointly consistent with the idea that governments seek to use debt to smooth the taxes needed to finance an exogenous stream of government spending.

Barro (1979) proposed a simple full-information rational expectations (RE) tax-smoothing model with only uncontingent debt in which the government spreads the burden of raising distortionary income taxes over time in order to minimize their welfare losses to address these questions. Specifically, the model predicts that the government should issue debts in order to spread the increases in tax rates over longer periods and minimize the welfare losses when facing a positive government spending shock.[1] Furthermore, the model also predicts that changes in the tax rate are unpredictable, i.e., the tax rate should follow a random walk.[2] It is worth noting that the tax-smoothing hypothesis (TSH) model is an analogy with the permanent income hypothesis (PIH) model in which consumers smooth consumption over time; tax rates respond to permanent changes in the public budgetary burden rather than transitory ones.[3]

Huang and Lin (1993) and Ghosh (1995) went beyond the random-walk tests and use the analysis of saving in Campbell (1987) to test the restrictions on the joint behavior of budget deficits and government spending implied by the tax-smoothing hypothesis.[4] They found mixed empirical evidence for the TSH. For example, Huang and Lin (1993) applied a log-linearized TSH model to the U.S. data from $1929-1988$, and found that the TSH is rejected for the full sample period, but it is not rejected for the sub-sample period from $1947-1988$. Ghosh (1995) used the US data from $1961-1988$ and Canadian data from $1962-1988$ and found that the TSH model cannot be rejected for either country. Adler (2006) tested the TSH using the Swedish central government data and found that it is not possible to statistically reject the TSH for the full period $1952-1999$, but the TSH is rejected using the sub-sample period from $1970-1996$.

---

[1]The tax-smoothing model is widely used in the literature to address various fiscal policy issues; see Sahasakul (1986), Trehan and Walsh (1988), Bohn (1990), Ghosh (1995), Angeletos (2002), and Lloyd-Ellis, Zhan, and Zhu (2002). For the tax-smoothing setting with state-contingent debt, see Lucas and Stokey (1983), Aiyagari *et al.* (2002), and Karantounias, Hansen, and Sargent (2009).

[2]Although Barro (1979, 1981) report that in the US data the prediction that the tax rate is a random walk is difficult to reject statistically, Sahasakul (1986) is able to reject it.

[3]For a statement of this equivalence, see Sargent (2001).

[4]It is well-known that it is often difficult to reject the null hypothesis of a random walk for many macroeconomic time series given the length of the data. Moreover, as argued in Ghosh (1995), the TSH might be only one of many potential explanations for unpredictable tax rate changes.

However, as shown in Table 3 in the next section, the standard full-information TS model cannot generate two key stochastic properties of the joint behavior of government deficits, tax rates, and spending, (i) the relative volatility of changes in tax rates and spending and (ii) the contemporaneous correlation between government deficits and spending, in the U.S. economy. Specifically, the model generates too low relative volatility and too high contemporaneous correlation.[5] As a result, we also find that the TSH is rejected using the US data from both $1951-2010$ and $1960-2010$.

In this paper, we show that model uncertainty due to the preference for robustness (RB) significantly improves the model's ability to fit the data discussed above. Hansen, Sargent, and Tallarini (1999) and Hansen and Sargent (2007) introduced the preference for robustness (a concern for model misspecification) into economic models. In robust control problems, agents are concerned about the possibility that their model is misspecified in a manner that is difficult to detect statistically; consequently, they choose their decisions as if the subjective distribution over shocks was chosen by an evil agent in order to minimize their utility (that is, the solution to a robust decision-maker's problem is the equilibrium of a max-min game between the decision-maker and the evil agent).[6] In the TSH setting, specifically, the preference for RB interacts with the *fundamental uncertainty* (the government spending and output shocks in the TSH model) and gives rise to a type of *induced uncertainty*: model uncertainty. This type of induced uncertainty can affect the model's dynamics even within the linear-quadratic (LQ) framework.[7] We adopt the LQ-TSH setting in this paper because the main purpose of this paper is to inspect the mechanisms through which the induced uncertainty affects the joint dynamics of government deficits, tax rates, and spending, and it is much more difficult to study this informational friction in non-LQ frameworks.[8] After solving the model explicitly, we then examine how model uncertainty due

---

[5]Aguiar and Gopinath (2007) showed that the otherwise standard small-open economy RBC model cannot generate realistic joint behavior of the current account, consumption, and output. In particular, they found that the model failed to generate the following two observations on the unconditional moments: (1) the correlation between the current account and output is more negative in emerging economies (i.e., the current account is more countercyclical in emerging countries) and (2) the relative volatility of consumption to income is larger in emerging countries than in developed countries. Luo, Nie, and Young (2012) showed that a simple intertemporal current account model with informational frictions can generate realistic unconditional moments of the joint dynamics of the current account, consumption, and income.

[6]As shown in Luo and Young (2010), consumption models can still produce precautionary savings but remain within the class of LQ-Gaussian models, which leads to analytical simplicity.

[7]Note that in the traditional linear-quadratic, linearized, or log-linearized models, uncertainty measured by the variance of the fundamental shock does not affect the model dynamics.

[8]As argued in Hansen and Sargent (2007), if the objective function is not quadratic or the state transition equation is not linear, worst possible distributions due to RB are generally non-Gaussian, which significantly

2

to RB can improve the model's predictions on these important dimensions of the joint dynamics of government deficits, tax rates, and spending we discussed above.

We show that the RB model has the potential to lead to the observed different joint behavior of government deficits, tax rates, and spending in the US, depending on the parameter governing the concern for model misspecification. Specifically, we find that incorporating RB can improve the model along the two key dimensions: generating higher relative volatility of changes in tax rates to government spending and lower contemporaneous correlation between government deficits and government spending.[9] In addition, after calibrating the RB parameter using the detection error probability, we find that RB can quantitatively help generate the empirical stochastic properties of the joint dynamics of government deficits, tax rates, and spending.

In addition to improving the basic model's ability to match the unconditional moments of fiscal policy, we are also able to show that the RB model improves the match of conditional moments; specifically, the RB model can pass the bivariate VAR test used in Huang and Lin (1993), Ghosh (1995), Adler (2006), and others, provided the desire for robustness is strong enough. We discipline this desire using detection error probabilities, as advocated by Hansen and Sargent (2007) and used in Luo, Nie, and Young (2012); the result is that the model passes the VAR test for reasonable error probabilities, particularly if we exclude the Korean War from our data.

The remainder of the paper is organized as follows. Section 2 reviews the standard full-information rational expectations taxation smoothing model proposed by Barro (1979). Section 3 introduces robustness into the standard TSH model and examines the implications of model uncertainty due to RB for the joint dynamics of optimal taxation and government debt. Section 4 discusses how to use detection error probabilities to calibrate the robustness parameter and presents the calibration results. Section 5 concludes.

2   The Model

In this section, we propose a robustness (RB) version of Barro (1979)'s tax-smoothing model and explore how RB affects optimal taxation and government debt. We assume that the only asset that is traded internationally is a riskfree bond. In choosing the optimal tax rate, the government

---

complicates the computational task. See Hansen and Sargent (2007) for detailed discussions on the difficulties in solving the non-LQ RB models, and Bidder and Smith (2011) and Young (2012) for nonlinear approaches that characterize the worst-case distributions.

[9]In an extension, we show that the impacts of RB on these two key moments still hold even if we allow for multiple shocks in the government spending process.

3

is concerned about model misspecification.

2.1 Taxation Smoothing Hypothesis Problem

In this subsection we present optimal paths of the tax rate and the budget surplus when the government chooses to minimize the present value of excess burden from taxation given the stochastic processes of government spending and real GDP. In the absence of a first-best system of lump-sum taxes, the government must seek to minimize the welfare losses arising from its choice of tax rates. These losses are assumed to be an increasing, convex and time invariant function of the average tax rate. Specifically, following Barro (1979), Roubini (1988), Bohn (1990), Ghosh (1995), and Lloyd-Ellis, Zhan, and Zhu (2005), we assume that the excess burden of taxation is summarized by a quadratic loss function $c(\tau_t) = \frac{1}{2}(\tau_t + \varphi)^2$, which measures the value of real income "wasted" when taxes are τ_t.[10] The optimization problem of the government can thus be written as

$$\max_{\{\tau_t, b_{t+1}\}} E_0 \left[-\frac{1}{2} \sum_{t=0}^{\infty} \beta^t (\tau_t + \varphi)^2 \right], \tag{1}$$

subject to its flow budget constraint

$$B_{t+1} = RB_t + G_t - \tau_t Y_t, \tag{2}$$

where $E_0[\cdot]$ is the government's expectation conditional on its information set at time 0, β is the government's subjective discount factor, τ_t is the tax rate, B_t is the amount of government debt, G_t is government spending, Y_t is real GDP, and $R \geq 1$ is the gross interest rate.

Suppose that real GDP is growing at a constant rate, n. Then (2) can be rewritten as

$$(1+n) b_{t+1} = Rb_t + g_t - \tau_t, \tag{3}$$

where lowercase letters denote corresponding variables expressed as a fraction of real GDP, Y_t. Solving the government's optimization problem (1) subject to (3) and a no-Ponzi condition yields the following optimal path for the tax rate:

$$\tau_t = \left(R - \frac{(1+n)^2}{\beta R} \right) E_t \left[b_t + \frac{1}{R} \sum_{j=0}^{\infty} \left(\frac{1+n}{R} \right)^j g_{t+j} \right], \tag{4}$$

---

[10] Following Barro (1979), Sargent (1987), Bohn (1989), and Huang and Lin (1993), we impose the restrictions $c'(\tau) > 0$ and $c''(\tau) > 0$, on the loss function, $c(\tau)$, by choice of φ.

which is just a permanent government expenditures theory proposed by Barro (1979).[11] Denoting by $\widetilde{R} = R/(1+n)$ the effective interest rate faced by the government, (4) can be written as

$$\tau_t = (1+n)\left(\widetilde{R} - \frac{1}{\beta\widetilde{R}}\right)h_t, \tag{5}$$

where

$$h_t = E_t\left[b_t + \frac{1}{(1+n)\widetilde{R}}\sum_{j=0}^{\infty}\left(\frac{1}{\widetilde{R}}\right)^j g_{t+j}\right] \tag{6}$$

is permanent government spending, equal to the sum of government debt and the present discounted value of expected government expenditure. This expression is analogous to permanent income in a standard consumption-savings problem.

Define $\gamma \equiv \frac{1-1/(\beta\widetilde{R}^2)}{1-1/\widetilde{R}}$. It is clear that $\gamma \lessgtr 1$ iff $\beta\widetilde{R} \lessgtr 1$. In the literature, the case in which $\beta\widetilde{R} < 1$ (i.e., $\gamma < 1$) is called tax tilting.[12] In the case in which $n = 0$ and $\beta R = 1$, $\gamma = 1$. Tax smoothing and tax tilting are two main motivations for a government to run budget deficits. Note that even if government spending (as a fraction of real GDP) is constant over business cycles such that there is no need for tax-smoothing, budget deficits may still arise due to the tax-tilting motivation to shift (tilt) taxes across time. Specifically, if $\beta\widetilde{R} < 1$, i.e., the subjective discount rate of the government is high, the government will choose to have low tax rates early and increase tax rates over time to finance its accumulated debts. By contrast, if $\beta\widetilde{R} > 1$, the government will choose to have relatively high tax rates early and then use accumulated assets (i.e., negative debts) to reduce tax rates over time. Finally, if $\beta\widetilde{R} = 1$, then there is no tax-tilting motive for budget deficits but there would be a strong tax-smoothing motive. Following the literature, we will focus only on the tax-smoothing motive (i.e., the $\beta\widetilde{R} = 1$ case) when we examine the stochastic properties of the joint behavior of government taxes, deficits, and spending. In this case, the tax function can be written as

$$\tau_t = (1+n)\left(\widetilde{R} - 1\right)h_t. \tag{7}$$

Note that if $n = 0$, $\tau_t = (R-1)h_t$ and $h_t = E_t\left[b_t + \sum_{j=0}^{\infty}\left(\frac{1}{R}\right)^j g_{t+j}\right]$.

As shown in Luo and Young (2010) and Luo, Nie, and Young (2012), in order to facilitate the introduction of robustness we reduce the above multivariate model with a general exogenous process to a univariate model with iid innovations that can be solved in closed-form. Specifically, in the above TSH model, if h_t is defined as a new state variable, combining (6) with the original

---

[11]See Appendix 6.1 for the derivation.
[12]See Ghosh (1995) for a discussion on this topic.

government budget constraint, we obtain the following new budget constraint:

$$h_{t+1} = \widetilde{R}h_t - \frac{1}{1+n}\tau_t + \frac{1}{1+n}\zeta_{t+1} \tag{8}$$

where

$$\zeta_{t+1} = \sum_{j=0}^{\infty}\left(\frac{1}{\widetilde{R}}\right)^{j+1}(E_{t+1} - E_t)[g_{t+1+j}] \tag{9}$$

is the time $(t+1)$ innovation to permanent spending. We can thus reformulate the above model as

$$v(h_0) = \max_{\{\tau_t\}_{t=0}^{\infty}}\left\{E_0\left[\sum_{t=0}^{\infty}\beta^t c(\tau_t)\right]\right\} \tag{10}$$

subject to (8) and $v(h_0)$ is the government's value function under RE.[13]

To close the model, we assume that government expenditures evolve according to the following AR(1) process:

$$g_{t+1} = (1 - \rho_g)\overline{g} + \rho_g g_t + \epsilon_t, \tag{11}$$

where $\rho_g \in [0,1]$ and ϵ_t is an iid shock with mean zero and variance ω^2. Substituting (11) into (6) and (9) yields:

$$h_t = b_t + \frac{1}{(1+n)\left(\widetilde{R} - \rho_g\right)}g_t + \frac{(1 - \rho_g)\overline{g}}{(1+n)\left(\widetilde{R} - \rho_g\right)\left(\widetilde{R} - 1\right)} \tag{12}$$

$$\zeta_t = \frac{1}{\widetilde{R} - \rho_g}\epsilon_t. \tag{13}$$

Therefore, under the full-information RE hypothesis, the change in optimal taxation can be written as

$$\Delta\tau_t = \frac{\widetilde{R} - 1}{\widetilde{R}}(E_t - E_{t-1})\left[\sum_{j=0}^{\infty}\left(\frac{1}{\widetilde{R}}\right)^j g_{t+j}\right] \tag{14}$$

$$= \frac{\widetilde{R} - 1}{\widetilde{R} - \rho_g}\epsilon_t,$$

which relates the innovation to taxation to government spending shocks. In other words, the optimal tax rate follows a random walk in the standard rational expectations model. Although the optimal plan here is to smooth tax rates so that $\tau_t = E_t[\tau_{t+1}]$, it does not mean that τ_t is

---

[13]In the next section, we will introduce robustness directly into this 'reduced' TSH model; in the Appendix we show that the general robust multivariate version delivers the same decision rules because debt evolves deterministically, so that the evil nature cannot distort its law of motion.

unaffected by the government spending shock; rather, the optimal tax behavior is to make any changes in taxes unpredictable.

The government budget deficit can be written as

$$
\begin{aligned}
def_t &= b_{t+1} - b_t \\
&= \left(\widetilde{R} - 1 \right) b_t + \frac{1}{1+n} g_t - \frac{1}{1+n} \tau_t \\
&= - \sum_{j=1}^{\infty} \widetilde{R}^{-j} E_t \left[\Delta g_{t+j} \right],
\end{aligned}
\tag{15}
$$

which means that the government sets the budget deficit to be equal to expected changes in government spending. Using (11), we obtain that $def_t = \frac{1}{1+n} \frac{1-\rho_g}{\widetilde{R}-\rho_g} (g_t - \overline{g})$. That is, the government deficit is positively correlated with government spending. We now consider two special cases: (i) when the government spending shock is permanent, i.e., $\rho_g = 1$ (e.g., expenditures on education and health care), and (ii) when the shock is purely transitory, i.e., $\rho_g = 0$ (e.g., expenditures on unemployment benefit and natural disaster). From (14) and (15), it is clear that for the first case, it is optimal to absorb a permanent expenditure shock *entirely* by taxes because

$$
\Delta \tau_t = \epsilon_t,
$$
$$
def_t = 0.
$$

In contrast, for the second case,

$$
\Delta \tau_t = \frac{\widetilde{R} - 1}{\widetilde{R}} \epsilon_t
$$
$$
def_t = \frac{1}{1+n} \frac{1}{\widetilde{R}} \epsilon_t,
$$

which means that for a positive government spending shock, τ_t must increase by $\frac{\widetilde{R}-1}{\widetilde{R}} \epsilon_t$ and the debt must be raised by $\frac{1}{1+n} \frac{1}{\widetilde{R}} \epsilon_t$ to guarantee that the GBC is satisfied. In other words, a proportion of the positive shock, $\frac{\widetilde{R}-1}{\widetilde{R}}$, is absorbed by taxation and the rest $\frac{1}{\widetilde{R}}$ is absorbed by debt. Since $\widetilde{R}$ is close to 1, most of the shock is absorbed by debt.

2.2   Model's Predictions

Using (14), it is straightforward to compute the relative volatility of the change in tax rates to government spending, which we denote by μ:

$$
\mu = \frac{\text{sd} \left(\Delta \tau_t \right)}{\text{sd} \left(\Delta g_t \right)} = \frac{\widetilde{R} - 1}{\widetilde{R} - \rho_g} \sqrt{\frac{1 + \rho_g}{2}},
\tag{16}
$$

where sd denotes standard deviation and we use the fact that $\Delta g_t = \frac{(\rho_g - 1)\epsilon_{t-1}}{1 - \rho_g L} + \epsilon_t$. Using the estimated ρ_g reported in Table 2 and assuming that $\widetilde{R} = 1.01$, the RE model predicts that $\mu = 0.05$. However, in the data as reported in Table 1, $\mu = 1.21$, so the model is not even in the ballpark.[14] Given the estimated processes of government spending, the model predicts that changes in tax rates (relative to changes in spending) should be much smaller than they are in the data.

Using (15), the model predicts that the contemporaneous correlation between the government debt and government spending, corr (def_t, g_t), is 1. This model prediction contradicts the empirical evidence: in the data the empirical correlation between them is well below 1. As reported in Table 1, corr $(def_t, g_t) = 0.72$ (for HP filter) and 0.57 (for linear filter).

In the literature, there are two other common tests on the time series properties of average tax rates and deficits. The first is to test whether tax rates follow a random walk. For example, Barro (1981) found that the average tax rate in the U.S. between 1884 and 1979 follows a random walk.[15] However, as is well known, this test is not conclusive for the tax-smoothing hypothesis (TSH) because tax rates could also follow a random walk if the rates are determined by a random political process. The second is to test whether the budget balance is informative about future changes in government spending using the VAR techniques. For example, Huang and Lin (1993) and Ghosh (1995) applied Campbell's (1987) VAR approach and found that increases in the budget surplus signal future increases in government expenditure, which supports the tax-smoothing hypothesis.[16] This VAR approach is to formulate the TSH as an expression for the budget surplus (or deficits) because it takes into account the entire structure of the TSH model and then use a bivariate VAR for government spending and the budget surplus to evaluate the restrictions imposed by the TSH model. The basic idea of this approach is that using the bivariate VAR, it is possible to compute the predicted time path of the budget deficits that is optimal for the government under the TSH and then compare this optimal path to the actual time path of the budget deficits. Specifically, we compare the optimal path of deficits, $def_t = -\sum_{j=1}^{\infty} \widetilde{R}^{-j} E_t [\Delta g_{t+j}]$, to the actual

---

[14]Given the estimated ρ_g using the linear filter reported in Table 3, the RE model predicts that $\mu = 0.016$, which is even smaller relative to the data.

[15]Other studies include Trechan and Walsh (1988) for the U.S., Ghosh (1995) for the U.S. and Canada, and Strazicich (1997) for the sub-national governments of Canada and the U.S..

[16]Adler (2006) tested the TSH using Swedish data from the period $1952 - 1999$, and found that the tax smoothing behavior of the government can explain a significant fraction of the variability in the Swedish government budget deficits.

8

one, $def_t = (R-1)b_t + g_t - \tau_t$, using the following bivariate VAR:[17]

$$\begin{bmatrix} \Delta g_{t+1} \\ def_{t+1} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} \Delta g_t \\ def_t \end{bmatrix} + \begin{bmatrix} e_{1,t+1} \\ e_{2,t+1} \end{bmatrix}, \tag{17}$$

or $z_{t+1} = Az_t + e_{t+1}$, where $e_{1,t+1}$ and $e_{2,t+1}$ are the VAR innovations. Denote $e_{t+1} = \begin{bmatrix} e_{1,t+1} & e_{2,t+1} \end{bmatrix}$ and assume that $\Sigma_e = E\left[e_{t+1}e'_{t+1}\right]$ is the 2×2 variance-covariance matrix of the reduced-form VAR innovations. The k-step ahead forecast from the VAR is

$$E_t\left[z_{t+k}\right] = A^k z_t$$

and (15), $def_t = -\sum_{j=1}^{\infty} \widetilde{R}^{-j} E_t\left[\Delta g_{t+j}\right]$, can be rewritten as

$$\begin{bmatrix} 0 & 1 \end{bmatrix} z_t = \sum_{j=1}^{\infty} \widetilde{R}^{-j} A^j \begin{bmatrix} 1 & 0 \end{bmatrix} z_t.$$

Given that both Δg_t and def_t are stationary, the infinite sum in the above equation converges to

$$def_t^* = -\begin{bmatrix} 1 & 0 \end{bmatrix} A\widetilde{R}^{-1}\left(I - A\widetilde{R}^{-1}\right)^{-1} z_t \equiv \Lambda z_t = \lambda_1 \Delta g_t + \lambda_2 def_t. \tag{18}$$

Therefore, if the null hypothesis that the government smooths taxes over time is correct, the recovered coefficients from the estimated VAR, λ_1 and λ_2, should be 0 and 1, respectively. We will use this test in Section 4; to satisfy the reader's curiosity for the moment, we note here that a joint test of these restrictions under rational expectations has a Wald statistic of 20.04 on data from 1951-2010, which is about as strong a rejection as one is likely to see (see Table 8).[18]

3 Incorporating Robustness

3.1 The RB Version of the Tax Smoothing Hypothesis Model

Robust control emerged in the engineering literature in the 1970s and was introduced into economics and further developed by Hansen, Sargent, and others. A simple version of robust optimal

---

[17]As argued in Campbell (1987), Huang and Lin (1993), Ghosh (1995), and others, although the information set of an econometrician is only a subset of the government's information set, the econometrician can still compute the predicted path of the budget deficit because the deficit itself contains all of the information about future changes in government spending. That is, the budget deficit should Granger-cause changes in government spending.

[18]As shown in Table 9, a sample from 1960-2010 does better (Wald statistic of only 8.04), but that is still a very strong rejection

control considers the question of how to make decisions when the agent does not know the probability model that generates the data. Specifically, an agent with a preference for robustness considers a range of models surrounding the given approximating model, (8):

$$h_{t+1} = \widetilde{R}h_t - \frac{1}{1+n}\tau_t + \widetilde{\zeta}_{t+1} + \omega_{\widetilde{\zeta}}\nu_t, \tag{19}$$

where $\widetilde{\zeta}_{t+1} = \frac{1}{1+n}\zeta_{t+1}$ and $\omega_{\widetilde{\zeta}} = \frac{1}{1+n}\omega_{\zeta}$, and makes decisions that maximize lifetime expected utility given this worst possible model (i.e., the distorted model). To make that model (8) is a good approximation when (19) generates the data, we constrain the approximation errors by an upper bound η_0:

$$E_0\left[\sum_{t=0}^{\infty}\beta^{t+1}\nu_t^2\right] \leq \eta_0, \tag{20}$$

where $E_0[\cdot]$ denotes conditional expectations evaluated with model, and the left side of this inequality is a statistical measure of the discrepancy between the distorted and approximating models. Note that the standard full-information RE case corresponds to $\eta_0 = 0$. In the general case in which $\eta_0 > 0$, the evil agent is given an intertemporal entropy budget $\eta_0 > 0$ which defines the set of models that the agent is considering. Following Hansen and Sargent (2007), we compute robust decision rules by solving the following two-player zero-sum game: a minimizing decision maker chooses optimal taxation $\{\tau_t\}$ and a maximizing evil agent chooses model distortions $\{\nu_t\}$. Specifically, the robustness version of the Barro model proposed in Section 2 can be written as

$$v(h_t) = \max_{\tau_t}\min_{\nu_t}\left\{-\frac{1}{2}(\tau_t+\varphi)^2 + \beta\left[\vartheta\nu_t^2 + E_t\left[v(h_{t+1})\right]\right]\right\} \tag{21}$$

subject to the distorted transition equation (i.e., the worst-case model), (19), where ν_t distorts the mean of the innovation, and $\vartheta > 0$ is the Lagrange multiplier on the constraint specified in (20) and controls how bad the error can be.[19]

When the ratio of government spending to real GDP follows the AR(1) process, (11), solving this robust control problem yields the following proposition:

Proposition 1 *Under RB, the optimal tax function is*

$$\tau_t = \frac{(1+n)\left(\widetilde{R}-1\right)}{1-\Sigma}h_t + \frac{\Sigma\varphi}{1-\Sigma}, \tag{22}$$

---

[19]Formally, this setup is a game between the decision-maker and a malevolent nature that chooses the distortion process ν_t. $\vartheta \geq 0$ is a penalty parameter that restricts attention to a limited class of distortion processes; it can be mapped into an entropy condition that implies agents choose rules that are robust against processes which are close to the trusted one. In a later section we will apply an error detection approach to calibrate ϑ in the U.S. economy.

the mean of the worst-case shock is

$$\nu_t \omega_{\widetilde{\zeta}} = \frac{\left(\widetilde{R}-1\right)\Sigma}{1-\Sigma} h_t + \frac{1}{1+n}\frac{\Sigma}{1-\Sigma}\varphi, \tag{23}$$

the government deficit can be written as

$$def_t = b_{t+1} - b_t = \frac{1}{1+n}\frac{1-\rho_g}{\widetilde{R}-\rho_g}\left(g_t - \overline{g}\right) - \left(1-\rho_h\right)h_t - \frac{1}{1+n}\frac{\Sigma\varphi}{1-\Sigma}, \tag{24}$$

and $h_t\left(= b_t + \frac{1}{(1+n)\left(\widetilde{R}-\rho_g\right)}g_t + \frac{(1-\rho_g)\overline{g}}{(1+n)\left(\widetilde{R}-\rho_g\right)\left(\widetilde{R}-1\right)}\right)$ is governed by

$$h_{t+1} = \rho_h h_t - \frac{1}{1+n}\frac{\Sigma\varphi}{1-\Sigma} + \widetilde{\zeta}_{t+1}, \tag{25}$$

where $\zeta_{t+1} = \frac{1}{\widetilde{R}-\rho_g}\epsilon_{t+1}$, $\widetilde{\zeta}_{t+1} = \frac{1}{1+n}\zeta_{t+1}$,

$$\Sigma = (1+n)^2\,\widetilde{R}\omega_{\widetilde{\zeta}}^2/\left(2\vartheta\right) = \widetilde{R}\omega_{\zeta}^2/\left(2\vartheta\right) > 0 \tag{26}$$

measures the degree of the preference of robustness, and $\rho_h = \frac{1-\widetilde{R}\Sigma}{1-\Sigma} \in (0,1)$.

Proof. See Appendix 6.2. ∎

The univariate RB model with unique state variable h leads to the same taxation function as the corresponding multivariate RB model in which the state variables are b_t and g_t. The key difference between these two models is that in the univariate RB model the evil agent distorts the transition equation of permanent spending h_t, whereas in the multivariate RB model the evil agent distorts the spending-output ratio process g_t. Theoretically, introducing RB into the multivariate model affects both the coefficients attached to b_t and g_t in the taxation function. That is, in the multivariate model RB may affect the relative importance of the two state variables in the taxation function, whereas in the univariate model the relative importance of the two effects are *fixed* by reducing the (b,g) state space to the univariate state h. However, after solving the two-state model numerically using the standard procedure proposed in Hansen and Sargent (2007), we can see that the two models lead to the same decision rule (see Appendix 6.3); the intuition is that the debt evolves deterministically from t to $t+1$, so that the evil agent cannot influence it. The main reason that we adopt the univariate RB model here is to obtain the explicit expressions for optimal taxation and government debts which can be easily used to compute the model's stochastic properties (e.g., the relative volatility of taxation and debts to government spending and the correlation between government debts and spending).

The effect of the preference for robustness, Σ, is jointly determined by the RB parameter, ϑ, and the volatility of the permanent spending shock, $\omega_{\widetilde{\zeta}}$. This interaction provides a novel

11

channel that the government spending shock affects optimal taxation and debts. Specifically, optimal taxation under RB, (22), shows that the preference for robustness, ϑ, affects the optimal response of optimal taxation to the change in permanent government expenditures, $(1+n)\frac{\widetilde{R}-1}{1-\Sigma}$. The smaller the value of ϑ the larger the response of optimal taxation; that is, under RB optimal taxation is more sensitive to unanticipated government spending shocks. This response is referred to as "making hay while the sun shines" in van der Ploeg (1993) to explain the overreaction of consumption to unanticipated income shocks, and reflects the precautionary aspect of these preferences.

Proposition 2 $\Sigma < 1$.

Proof. The second-order condition for a minimization by nature is

$$A = \frac{1}{2}\frac{(1+n)^2\,\widetilde{R}\left(\widetilde{R}-1\right)}{1-(1+n)^2\,\widetilde{R}\omega_\zeta^2/\left(2\vartheta\right)} > 0,$$

which can be rearranged into

$$\vartheta > (1+n)^2\,\widetilde{R}\omega_\zeta^2. \tag{27}$$

Using the definition of Σ we obtain

$$1 > \Sigma.$$

■

It is straightforward to show that the robust policy rule, (22), can also be obtained by solving the following risk-sensitive TSH model:

$$v\left(h_t\right) = \max_{\tau_t}\left\{-\frac{1}{2}\left(\tau_t+\varphi\right)^2 + \beta\mathcal{R}_t\left[v\left(h_{t+1}\right)\right]\right\} \tag{28}$$

subject to (8), and the distorted expectation operator $\mathcal{R}_t$ is defined by

$$\mathcal{R}_t\left[v\left(h_{t+1}\right)\right] = -\frac{1}{\alpha}\log E_t\left[\exp\left(-\alpha v\left(h_{t+1}\right)\right)\right], \tag{29}$$

where $\alpha > 0$ measures higher risk aversion of the government *vis a vis* the von Neumann-Morgenstern specification.[20] Risk-sensitivity (RS) was first introduced into the LQG framework by Jacobson (1973) and extended by Whittle (1981). van der Ploeg (1993) applied this preference

---

[20]The detailed proof is available from the authors by request. The observational equivalence between the risk-sensitive and robust LQG models has been well established in the literature. See Hansen, Sargent, and Tallarni (1999), Backus, Routledge, and Zin (2004), and Luo and Young (2010).

to examine its implication for precautionary savings; Hansen and Sargent (1995) introduced discounting into the RS specification and showed that the resulting decision rules are time-invariant; Hansen, Sargent, and Tallarini (1999) also explored its implications for precautionary savings and asset prices; and Luo and Young (2010) examined its implications for consumption and precautionary savings when consumers are inattentive. In the risk-sensitive TSH specified in (28), the government is prudent in the sense that they minimizes the expected value of an exponential transformation of a quadratic welfare loss function and adjust optimal taxation policy more aggressively to changes in government expenditures.

3.2 Implications of RB for Tax Smoothing and Government Debt

Given the expression for optimal taxation (22), the dynamics of tax can be written as

$$\tau_{t+1} = \rho_h \tau_t + \frac{(1+n)\left(\widetilde{R}-1\right)}{1-\Sigma}\widetilde{\zeta}_{t+1} - \left(\widetilde{R}-1\right)\frac{\Sigma\varphi}{1-\Sigma}, \text{ or } \tau_{t+1} = \rho_h \tau_t + \frac{\widetilde{R}-1}{1-\Sigma}\zeta_{t+1} - \left(\widetilde{R}-1\right)\frac{\Sigma\varphi}{1-\Sigma} \tag{30}$$

and the government deficit can be rewritten as the following AR(1) process:

$$def_{t+1} = \rho_h def_t + \frac{1}{1+n}\frac{(1-\rho_g)(\rho_g-\rho_h)}{\widetilde{R}-\rho_g}(g_t-\overline{g}) + (\rho_h-\rho_g)\widetilde{\zeta}_{t+1} \tag{31}$$

It is worth noting that even if the government's subjective discount factor is assumed to equal the inverse of the effective interest rate, i.e., $\beta\widetilde{R} = 1$, the tax-tilting aspect of government budget deficits still exists under RB. The intuition is very simple: The effect of RB on optimal taxation and deficits are the same as that of the discount factor within the LQ setting.[21] We also note that the tax rate is a stationary process even if g is a random walk, in contrast to the RE model.

3.2.1 The Relative Volatility of Taxes and Government Debts to Government Spending

The relative volatility of the change in tax rates to government spending, μ_1, can be written as

$$\mu_1 = \frac{\text{sd}(\Delta\tau_t)}{\text{sd}(\Delta g_t)} = \frac{\widetilde{R}-1}{(1-\Sigma)\left(\widetilde{R}-\rho_g\right)}\sqrt{\frac{1+\rho_g}{1+\rho_h}}, \tag{32}$$

---

[21] See Hansen, Sargent, and Tallarini (1999) and Luo and Young (2010) for discussions on the observational equivalence between RB and the discount factor in the LQ permanent income models.

where we use the fact that $\Delta g_t = \frac{(\rho_g - 1)\epsilon_{t-1}}{1 - \rho_g L} + \epsilon_t.$[22] (See Appendix 6.3.) It is clear from (32) that RB increases the relative volatility via two channels: first, it strengthens the marginal propensity to taxation out of permanent government spending $\left(\frac{\widetilde{R}-1}{1-\Sigma}\right)$; and second, it increases the volatility of the change in the tax rate by reducing the persistence of permanent government spending measured by ρ_h: $\frac{\partial \rho_h}{\partial \Sigma} < 0$. Figure 2 illustrates how this relative volatility varies with the degree of model uncertainty, Σ. It clearly shows that this ratio is increasing with the degree of RB, Σ.

3.2.2 The Correlation between the Government Debt and Government Spending

A description of the comovement of the government debt and government spending is the contemporaneous correlation between them, corr (def_t, g_t). Under RB, the correlation can be written as:

$$\text{corr}\,(def_t, g_t) = \left[\left(\widetilde{R} - \rho_g\right)\left(\frac{1}{1+\rho_g} - \frac{1-\rho_h}{1-\rho_g\rho_h}\right)\right] \Big/ \left[\sqrt{\frac{\left(\widetilde{R}-\rho_g\right)^2}{1-\rho_g^2}}\sqrt{\frac{1-\rho_g}{1+\rho_g} + \frac{1-\rho_h}{1+\rho_h} - 2\frac{(1-\rho_h)(1-\rho_g)}{1-\rho_g\rho_h}}\right],$$

(33)

which converges to 1 as Σ converges to 0. (See Appendix 6.3.) Figure 3 illustrates that how RB affects the correlation between the government debt and government spending for different values of ρ. It is clear that corr (def_t, g_t) is decreasing with Σ. RB thus aligns the model and the data more closely along this dimension.

3.3 Two Special Cases

3.3.1 A Permanent Increase in Government Spending

Under RB, when $\rho_g = 1$, we have

$$\Delta\tau_t = \frac{\rho_h - 1}{1 - \Sigma}\frac{1}{1 - \rho_h L}\epsilon_{t-1} + \frac{1}{1 - \Sigma}\epsilon_t, \tag{34}$$

$$def_t = -\frac{1}{1+n}\frac{\Sigma}{1-\Sigma}\frac{1}{1-\rho_h L}\epsilon_t = -\frac{1}{1+n}\frac{\Sigma}{1-\Sigma}\frac{1}{1-\rho_h L}\epsilon_{t-1} - \frac{1}{1+n}\frac{\Sigma}{1-\Sigma}\epsilon_t. \tag{35}$$

It is clear that in this permanent income case, a proportion $\frac{1}{1-\Sigma}$ of a newly-arrived shock (ϵ_t) is absorbed by τ_t and $-\frac{1}{1+n}\frac{\Sigma}{1-\Sigma}$ of the shock is absorbed by debt. In other words, taxation is more sensitive to the permanent shock under RB and thus the government debt can be reduced even if

---

[22]We use the relative volatility of the change in the tax rate to the change in government spending instead of that of the tax rate to government spending to compare the implications of RE and RB models, as the tax rate follows a random walk under RE and the volatility of the tax rate is not well defined in this model.

the spending shock is positive. In addition, both taxes and debts react to the lagged innovations to government spending. This prediction is clearly different from that we obtained in the RE model: the permanent expenditure shock is absorbed *entirely* by taxes.

3.3.2 A Temporary Increase in Government Spending

Under RB, when $\rho_g = 0$, we have

$$\Delta\tau_t = \frac{(\rho_h - 1)\left(\widetilde{R} - 1\right)}{(1 - \Sigma)\widetilde{R}} \frac{\epsilon_{t-1}}{1 - \rho_h L} + \frac{\widetilde{R} - 1}{(1 - \Sigma)\widetilde{R}}\epsilon_t, \tag{36}$$

$$def_t = \frac{1}{1+n}\left[\frac{1}{\widetilde{R}}\epsilon_t - \frac{\Sigma\left(\widetilde{R} - 1\right)}{(1 - \Sigma)\widetilde{R}}\frac{\epsilon_t}{1 - \rho_h L}\right] = \frac{1}{1+n}\left[-\frac{\Sigma\left(\widetilde{R} - 1\right)}{(1 - \Sigma)\widetilde{R}}\frac{\rho_h}{1 - \rho_h L}\epsilon_{t-1} + \frac{1 - \Sigma\widetilde{R}}{(1 - \Sigma)\widetilde{R}}\epsilon_t\right], \tag{37}$$

which means that for a positive government spending shock (ϵ_t), τ_t must increase by $\frac{\widetilde{R}-1}{(1-\Sigma)\widetilde{R}}\epsilon_t$ and the debt must be reduced by $\frac{1-\Sigma\widetilde{R}}{(1-\Sigma)\widetilde{R}}\epsilon_t$. In other words, a proportion of the positive shock, $\frac{\widetilde{R}-1}{(1-\Sigma)\widetilde{R}}$, is absorbed by taxation and the rest, $\frac{1-\Sigma\widetilde{R}}{(1-\Sigma)\widetilde{R}}$, is absorbed by debt. Note that in the RE model a proportion of the shock, $\frac{\widetilde{R}-1}{\widetilde{R}}$, is absorbed by taxation and the rest, $\frac{1}{\widetilde{R}}$, is absorbed by debt. Since $\Sigma > 0$, it is clear that RB strengthens the relative importance of taxation to government debts in absorbing the government spending shock.

3.4 Extension: Multiple Government Spending Shocks

In the benchmark model presented above, we assume that there is only a single shock to government spending. Since government spending in the model, g, is a ratio of the amount of government expenditures to real GDP, G/Y, there might exist muliple shocks to g as real GDP can also be modelled with multiple shocks. In this section, we consider both permanent and transitory shocks to government spending simultanenously and show that the incorporating multiple shocks does not affect our results on how RB affects the two key unconditional moments we discussed in the previous subsection.

Specifically, we now assume that the governement spending-real GDP ratio g_t can be expressed as the sum of permanent and transitory components:

$$g_{t+1} = g_{t+1}^p + g_{t+1}^t, \tag{38}$$

where the superscripts p and t denote permanent and transitory, respectively. Each of these

components follows its own stochastic process; g_t^p follows a random walk:

$$g_{t+1}^p = g_t^p + \varepsilon_{t+1}, \tag{39}$$

and the transitory component follows an stationary AR process:

$$g_{t+1}^t = \overline{y} + \rho_g \left(g_t^t - \overline{y} \right) + \epsilon_{t+1}, \tag{40}$$

where $\rho_g \in [0, 1)$, and all innovations are assumed to have zero mean, be uncorrelated over time and with each other, and the variance of ε_t and ϵ_t are ω_ε^2 and ω_ϵ^2, respectively. We assume that the decomposition of net income $(\omega_\varepsilon^2, \omega_\epsilon^2, \rho_g)$ does not affect the volatility of the change in observed g (sd $[\Delta y_t]$). For simplicity, we consider the case in which $\rho_g = 0$ and $\omega_\varepsilon^2 = m\omega_\epsilon^2$ $(m > 0)$. (Allowing $\rho_g > 0$ does not affect our conclusions that RB has the similar impacts on the joint dynamics of government taxes, deficits, and spending.)

Given the specifications, it is straightforward to show that the relative volatility of the change in tax rates to government spending, μ_1, can be written as

$$\mu_1 = \frac{\text{sd} \left(\Delta \tau_t \right)}{\text{sd} \left(\Delta g_t \right)} = \sqrt{ \frac{2 \left(\widetilde{R} - 1 \right)^2}{(1 - \Sigma) \left[2 - \Sigma \left(1 + \widetilde{R} \right) \right]} \Bigg/ \left[\frac{2 + m}{m / \left(\widetilde{R} - 1 \right)^2 + 1 / \widetilde{R}^2} \right] }, \tag{41}$$

where we use the facts that $\Delta g_{t+1} = \varepsilon_{t+1} + \epsilon_{t+1} - \epsilon_t$, $\Delta \tau_{t+1} = (\rho_h - 1) \tau_t + \frac{\widetilde{R}-1}{1-\Sigma} \zeta_{t+1} - \left(\widetilde{R} - 1 \right) \frac{\Sigma \varphi}{1-\Sigma}$, and $\zeta_{t+1} = \frac{\varepsilon_{t+1}}{R-1} + \frac{\epsilon_{t+1}}{R}$. Given that $\widetilde{R} = 1.01$, the full-information RE model predicts that $\mu_1 = 0.05$, which contradicts the empirical counterpart reported in Table 1, $\mu = 1.21$. Figure 4 illustrates how this relative volatility varies with the degree of model uncertainty, Σ, for different values of m that measure the relative importanace of the permanent shock (ε) to the transitory shock (ϵ). It clearly shows that this ratio is consistently increasing with Σ, which means that incorporating RB into the multiple-shock case has the potential to help improve the model's predictions. For example, when $m = 1$ and $\Sigma = 0.52$, $\mu_1 = 1.21$, which is just its empirical counterpart.

The correlation between the deficits and the change in government spending can be written as:[23]

$$\text{corr} \left(def_t, \Delta g_t \right) = \frac{-m \left(1 - \rho_h \right) / \left[(1 + n) \left(\widetilde{R} - 1 \right) \right] + \rho_h \left(2 - \rho_h \right) / R}{\sqrt{2 + m} \sqrt{ \left(\frac{1}{R} \right)^2 + \left[\frac{2}{\left(\widetilde{R} - 1 \right)^2} + \frac{1}{\widetilde{R}^2} \right] \left(\frac{1 - \rho_h}{1 + \rho_h} \right) / (1 + n)^2 - \frac{2(1 - \rho_h)}{R^2} }}, \tag{42}$$

---

[23]Note that here we use corr $\left(def_t, \Delta g_t \right)$ instead of corr $\left(def_t, g_t \right)$ because in this case the government spending process is non-stationary and thus corr $\left(def_t, g_t \right)$ is not well-defined.

which converges to 1 as Σ converges to 0. Figure 5 illustrates that how RB affects the correlation between the government debt and government spending for different values of m. It is clear that $\text{corr}\,(def_t, \Delta g_t)$ is consistently decreasing with Σ in the multiple-shock case, which means that RB can also have the potential to help explain this dimension. For example, when $m = 1$ and $\Sigma = 0.1$, $\text{corr}\,(def_t, \Delta g_t) = 0.2$, which is much closer to its empirical counterpart. (Note that in the full-information RE case $\text{corr}\,(def_t, \Delta g_t) = 0.6$.)

4   Calibration and Main Findings

In the previous section we have seen that the model uncertainty due to the preference for robustness can help better explain the observed relative volatility of the tax rate and the correlation between government budget deficit and government spending. Specifically, we have shown that an increase of Σ (which measures the RB effect) increases the relative volatility of the tax rate and reduces the correlation between government budget deficit and government spending. But careful readers may still have questions on how to further interpret this parameter. For example, two very reasonable questions are: What does Σ really measure? How should we link the value of Σ to the degree of model uncertainty? Thus, in this section, we will provide one useful method to quantify the RB effect Σ which intuitively links it to the degree of model uncertainty.

4.1   Calibrating the RB Parameter

We follow Hansen, Sargent, and Wang (2002) and Hansen and Sargent (2007) to calibrate the RB parameter (ϑ and Σ). Specifically, we calibrate the model by using the model detection error probability that is based on a statistical theory of model selection (the approach will be precisely defined below). We can then infer what values of the RB parameter ϑ imply reasonable fears of model misspecification for empirically-plausible approximating models. In other words, the model detection error probability is a measure of how far the distorted model can deviate from the approximating model without being discarded; standard significance levels for testing are then used to determine what reasonable fears entail.

4.1.1   The Definition of the Model Detection Error Probability

Let model A denote the approximating model and model B be the distorted model. Define p_A as

$$p_A = \text{Prob}\left(\log\left(\frac{L_A}{L_B}\right) < 0 \,\Big|\, A\right), \tag{43}$$

where $\log\left(\frac{L_A}{L_B}\right)$ is the log-likelihood ratio. When model A generates the data, p_A measures the probability that a likelihood ratio test selects model B. In this case, we call p_A the probability of the model detection error. Similarly, when model B generates the data, we can define p_B as

$$p_B = \text{Prob}\left(\left.\log\left(\frac{L_A}{L_B}\right) > 0\right| B\right). \tag{44}$$

Following Hansen, Sargent, and Wang (2002) and Hansen and Sargent (2007), the detection error probability, p, is defined as the average of p_A and p_B:

$$p\left(\vartheta\right) = \frac{1}{2}\left(p_A + p_B\right), \tag{45}$$

where ϑ is the robustness parameter used to generate model B. Given this definition, we can see that $1 - p$ measures the probability that econometricians can distinguish the approximating model from the distorted model. Now we show how to compute the model detection error probability in the RB model.

4.1.2 Calibrating the RB Parameter in the Tax-smoothing Model

Let's first consider the model with the robustness preference. Under RB, assuming that the approximating model generates the data, the state, h_t, evolves according to the transition law

$$\begin{aligned} h_{t+1} &= \widetilde{R}h_t - \frac{1}{1+n}\tau_t + \widetilde{\zeta}_{t+1}, \\ &= \frac{1 - \widetilde{R}\Sigma}{1 - \Sigma}h_t - \frac{1}{1+n}\frac{\Sigma\varphi}{1 - \Sigma} + \widetilde{\zeta}_{t+1}. \end{aligned} \tag{46}$$

In contrast, assuming that the distorted model generates the data, h_t evolves according to

$$\begin{aligned} h_{t+1} &= \widetilde{R}h_t - \frac{1}{1+n}\tau_t + \widetilde{\zeta}_{t+1} + \omega_{\widetilde{\zeta}}\nu_t, \\ &= h_t + \widetilde{\zeta}_{t+1}. \end{aligned} \tag{47}$$

In order to compute p_A and p_B, we use the following procedure:

1. Simulate $\{h_t\}_{t=0}^{T}$ using (46) and (47) a large number of times. The number of periods used in the simulation, T, is set to be the actual length of the data for each individual country.

2. Count the number of times that $\left.\log\left(\frac{L_A}{L_B}\right) < 0\right| A$ and $\left.\log\left(\frac{L_A}{L_B}\right) > 0\right| B$ are each satisfied.

3. Determine p_A and p_B as the fractions of realizations for which $\left.\log\left(\frac{L_A}{L_B}\right) < 0\right| A$ and $\left.\log\left(\frac{L_A}{L_B}\right) > 0\right| B$, respectively.

18

In practice, given Σ, to simulate the $\{h_t\}_{t=0}^{T}$ we need to know the volatility of ζ_t in (46) and (47). For a), we can compute it from $\operatorname{sd}(\zeta) = \frac{\sqrt{1-\rho_g^2}}{R-\rho_g}\operatorname{sd}(g)$, where $\operatorname{sd}(g)$ is the standard deviation of government spending. The gross interest rate R is set to be 1.04 in all calibrations.

4.2   Data

To implement the calibration described in the previous section, we need to have the measure of GDP, government spending (G_t), government debt (B_t), government budge deficit (def_t), tax rate (τ_t), and risk-free interest rate. We follow Huang and Lin (1993) and Lloyd-Ellis et al. (2005) to construct these variables and use the same data sources as described in their papers. Specifically, government expenditures are constructed by using federal expenditures minus net federal interest payments. The data on both variables are taken from *National Income and Product Accounts (NIPA)*. Government budget deficit is defined as total outlays minus total receipts.[24] Government debt is defined as the federal debt (end of period) held by the public. Tax rate is defined as tax revenues divided by GDP. The data on these variables are taken from *The Economic Report of President, 2011*. The data covers the period $1950-2010$. Finally, we need to know the annual growth rate of real GDP, n, and the risk-free interest rate, R. Using the real GDP data from *NIPA*, the former is calculated as the average annual growth rate of real GDP in the period of $1950-2010$, which is 3.3%. We use the risk-free rate reported in Mehra and Prescott (1985) so $R = 1.008$.

For variables as a ratio of GDP, we use either a linear filter or the Hodrick-Prescott (HP) filter (with a smoothing parameter of 100) to detrend the data.[25] For variables expressed in the form of differences (with a symbol Δ) the unfiltered series are used.

4.3   Main Results

The calibration results are reported in Table 2. We set $p = 0.1$ which means that there exists a 10-percent chance a likelihood ratio test fails to separate the approximating model and the distorted model. The resulted value of the RB parameter Σ (which measures the effect of RB) is 0.95. In the same table, we also report the persistence of the government spending (as a ratio of GDP), g, and the measured coefficient of variation for the processes of g and ζ (the innovation to

[24]In Lloyd-Ellis, Zhan, and Zhu (2005), the authors use the government budget surplus, which is defined as total receipts minus total outlays.

[25]We report the values of the key moments using both filters in Table 1. Since our permanent income model is stationary, we need to remove the low frequency component from the data. Thus in this paper we focus primarily on the linear filter when we calibrate the parameters and compare models with data.

the h_t process defined in budget constraint equation (8)).

As shown in the previous section, introducing model uncertainty (concern for robustness) qualitatively improves the model performance in explaining the two unconditional moments: (1) the relative standard deviation of the tax rate, $\frac{sd(\Delta\tau)}{sd(\Delta g)}$; and (2) the correlation between government budget deficit and government spending, $corr(def, g)$. Specifically, we have shown that an increase of Σ can increase the relative standard deviation of the tax rate while reduces the correlation of the government budget deficit and government spending. Since we have a calibrated value of Σ now, we can quantitatively show these improvements more clearly by comparing the model predicted values from the RE model and the RB model. The results are reported in Table 3.

The first column of Table 3 lists the values of the two moments in the data. The second column reports predicted in the model with no model uncertainty (RE model). As it shows, the relative volatility of the tax rate predicted in the RE model is only 0.04 while it is 1.21 in the data. On the other hand, the RE model overpredicts the correlation between government budget deficit and government spending. As a comparison, the RB (at $p = 0.1$) can predict values that are much closer to the data, as shown in the third column. These results just confirm our theoretic findings in Section 3. In that section, we have shown that as the RB effect (measured by Σ) increases, the model can predict more reasonable relative volatility of the tax rate and the correlation between deficit and spending. The results in Table 3 present one numerical example how much these improvements can be. Of course, the improvements depend on the used detection error probability.

As we have explained before, using the model detection error probability (p) is one intuitive way to measure the model uncertainty. In this example, we set $p = 0.1$ so that agents are considering a range of models which cannot be distinguished by a likelihood ratio test up to a probability of 0.1. Alternatively speaking, this means that agents have doubts that the model is misspecified to a certain degree that a likelihood ratio test cannot separate the (unknown) true model with other similar models by a 10-percent chance. To see how the degree of model uncertainty affects the model prediction and provide more robust check, we vary the detection error probability p (or the degree of model uncertainty) and report the corresponding results in Table 5. The corresponding calibrated values of Σ are reported in Table 4. Remember that as the detection error probability (p) increases, it's more difficult (using a likelihood-ratio test) to distinguish the set of considered models. Thus the range of models considered have to be smaller. In other words, there is less model uncertainty or agents have less concern about the model misspecifcations. As Table 5 shows, as the degree of the model uncertainty decreases, the relative volatility of the tax rate declines and the correlation between deficit and government spending

increases. These findings again confirm the theoretic results in the previous sections.

We note here that our results are sensitive to the sample period. We recalculated our Σ values using data covering 1960-2010, which excludes the very high volatility period surrounding the Korean War, and find two things (reported in Tables 6 and 7). First, the relative volatility of taxes to spending is significantly lower, while the correlation between the deficit and spending is roughly unchanged; given that spending volatility drops quite a bit as a result, the relative drop is driven by a large decline in the volatility of tax changes. Second, the ability of the RB model to match the relative volatility is impaired substantially; the RB model explains roughly two-thirds of the volatility in the long sample and only one-half in the short one. Exactly why this occurs is unclear.

4.4   Implications of RB for the Bivariate VAR Test on the TSH

In this subsection, we examine the implications of RB for the bivariate VAR test on the TSH. Using the optimal taxation function under RB, $\tau_t = \frac{(1+n)(\widetilde{R}-1)}{1-\Sigma}h_t + \frac{\Sigma\varphi}{1-\Sigma}$, we obtained in Section 3.1., the resulting deficit equation can be written as

$$def_t = -\sum_{j=0}^{\infty} \widetilde{R}^{-j} E_t\left[\Delta g_{t+j}\right] - \Sigma\tau_t \text{ or } \widetilde{def}_t = -\sum_{j=0}^{\infty} \widetilde{R}^{-j} E_t\left[\Delta g_{t+j}\right]$$

where $\widetilde{def}_t = def_t + \Sigma\tau_t$. Following the same procedure we discussed in Section 2.2, the VAR can be formulated

as

$$\left[\begin{array}{c} \Delta g_{t+1} \\ \widetilde{def}_{t+1} \end{array}\right] = \left[\begin{array}{cc} a_{11} & a_{12} \\ a_{21} & a_{22} \end{array}\right]\left[\begin{array}{c} \Delta g_t \\ \widetilde{def}_t \end{array}\right] + \left[\begin{array}{c} e_{1,t+1} \\ e_{2,t+1} \end{array}\right]. \tag{48}$$

Table 8 reports the recovered λ_1 and λ_2 from the full-information RE and RB models using the data for the period from $1951-2010$. To investigate how model uncertainty affects the test results, we report nine different RB models in which Σ varies from 0.1 to 0.9. Remember that a larger Σ means more model uncertainty.[26] The full-information RE model can be considered a model with zero model uncertainty. Thus, in total, Table 8 reports the recovered values for λ_1 and λ_2 for ten models differing only in the degree of model uncertainty, which increases from zero (the RE model) to a significantly large degree (measured by Σ). The second and third columns show that both λ_1 and λ_2 increase with the degree of model uncertainty. More importantly, the increase

---

[26]As explained in the previous section, a larger Σ corresponds to a smaller the detection error probability which means that the difference between the distorted model and the approximating model is larger. In other words, the range of the models (around the approximating model) is larger, or, there is more model uncertainty faced by the decision maker.

of model uncertainty leads λ_1 to increase from a negative level to a level closer to 0 (although it remains negative) and λ_2 to rise significantly toward 1. In other words, RB can help make the model better fit the data because the recovered coefficients from the estimated VAR, λ_1 and λ_2, should be 0 and 1, respectively, if the TSH holds. More formally, the last two columns report the Wald statistics and the associated p-values for the joint hypothesis of $\lambda_1 = 0$ and $\lambda_2 = 1$. As they show, the Wald statistic decreases and the associated p-value increases as the degree of model uncertainty increases. Of particular interest is that the values of Σ consistent with reasonable detection error probabilities – namely, those above 0.9 – are ones which potentially could pass the VAR tests. Thus, these findings suggest that a model which incorporates some positive degree of model uncertainty is more likely to pass the VAR test regarding the tax smoothing motivation.

Table 9 shows that the shorter sample from $1960 - 2010$ results in smaller coefficients for λ_1 when Σ is small and larger ones when Σ is large, along with larger values for λ_2. Combined with the results from the previous section, it appears that excluding the Korean War period helps the model generate more positive results on λ_1 and λ_2.

5 Conclusions

This paper has reconsidered the tax-smoothing hypothesis model of Barro (1979) under the assumption that the government faces model uncertainty regarding the stochastic process for required government spending. Our key finding is that many of the aspects of the tax-smoothing model that are inconsistent with the data under rational expectations – that is, under the assumption of no model uncertainty – are potentially consistent with the data when decisions by the government are made with an eye on robustness. In particular, we can increase the volatility of tax changes relative to changes in spending, decrease the correlation between spending and deficits, and pass VAR tests regarding the joint dynamics of changes in spending and deficits.

6 Appendix

6.1 Solving the Tax-Smoothing Model

For the following optimization problem:

$$\max_{\{\tau_t, b_{t+1}\}} E_0 \left[-\sum_{t=0}^{\infty} \beta^t \left(\frac{1}{2} (\tau_t + \varphi)^2 \right) \right],$$

subject to $(1+n) b_{t+1} = Rb_t + g_t - \tau_t$, the Lagrangian function can be written as

$$L = E_0 \left\{ \sum_{t=0}^{\infty} \beta^t \left[-\frac{1}{2} (\tau_t + \varphi)^2 + \lambda_t (Rb_t + g_t - \tau_t - (1+n) b_{t+1}) \right] \right\}$$

The FOC w.r.t. τ_t gives

$$\tau_t + \varphi = \lambda_t$$

$$- (1+n) \lambda_t + \beta R E_t [\lambda_{t+1}] = 0$$

which means that

$$\tau_t = \frac{\beta R}{1+n} E_t [\tau_{t+1}] + \varphi \left(\frac{\beta R}{1+n} - 1 \right). \tag{49}$$

When we impose the condition that $\frac{\beta R}{1+n} = 1$, (49) leads to the well-known random walk result of tax rates, $\tau_t = E_t [\tau_{t+1}]$.

The intertemporal budget constraint can be written as

$$\sum_{j=0}^{\infty} \left(\frac{1+n}{R} \right)^j g_{t+j} + Rb_t = \sum_{j=0}^{\infty} \left(\frac{1+n}{R} \right)^j \tau_{t+j}$$

Taking conditional expectations on both sides gives

$$E_t \left[Rb_t + \sum_{j=0}^{\infty} \left(\frac{1+n}{R} \right)^j g_{t+j} \right] = \sum_{j=0}^{\infty} \left(\frac{1+n}{R} \right)^j E_t [\tau_{t+j}]$$

$$= \left[\sum_{j=0}^{\infty} \left(\frac{1+n}{R} \right)^j \left(\frac{1+n}{\beta R} \right)^j \right] \tau_t$$

$$= \left[\sum_{j=0}^{\infty} \left(\frac{(1+n)^2}{\beta R^2} \right)^j \right] \tau_t$$

$$= \frac{1}{1 - (1+n)^2 / (\beta R^2)} \tau_t$$

and the optimal tax rate can be written as

$$\tau_t = \left(1 - \frac{(1+n)^2}{\beta R^2} \right) E_t \left[Rb_t + \sum_{j=0}^{\infty} \left(\frac{1+n}{R} \right)^j g_{t+j} \right]. \tag{50}$$

In the case in which $n = 1$ and $\beta R = 1$,

$$\tau_t = (R - 1) E_t \left[b_t + \frac{1}{R} \sum_{j=0}^{\infty} \left(\frac{1}{R} \right)^j g_{t+j} \right]$$

23

After defining

$$h_t = E_t \left[b_t + \frac{1}{R} \sum_{j=0}^{\infty} \left(\frac{1+n}{R} \right)^j g_{t+j} \right],$$

and combining this expression with the original budget constraint, we obtain the following budget constraint:

$$(1+n) h_{t+1} = Rh_t - \tau_t + \zeta_{t+1} \text{ or } h_{t+1} = \frac{R}{1+n} h_t - \frac{1}{1+n} \tau_t + \frac{1}{1+n} \zeta_{t+1}$$

where $\zeta_{t+1} = \sum_{j=0}^{\infty} \left(\frac{1+n}{R} \right)^{j+1} (E_{t+1} - E_t) [g_{t+1+j}]$.

6.2 Solving the Robust Model

To solve the Bellman equation (21), we conjecture that

$$v(h_t) = -Ah_t^2 - Bh_t - C,$$

where A, B, and C are undetermined coefficients. Substituting this guessed value function into the Bellman equation gives

$$-Ah_t^2 - Bh_t - C = \max_{\tau_t} \min_{\nu_t} \left\{ -\frac{1}{2} (\tau_t + \varphi)^2 + \beta E_t \left[\vartheta \nu_t^2 - Ah_{t+1}^2 - Bh_{t+1} - C \right] \right\}, \qquad (51)$$

where $h_{t+1} = \widetilde{R} h_t - \frac{1}{1+n} \tau_t + \widetilde{\zeta}_{t+1} + \omega_{\widetilde{\zeta}} \nu_t$. We can do the min and max operations in any order, so we choose to do the minimization first. The first-order condition for ν_t is

$$2\vartheta \nu_t - 2AE_t \left(\widetilde{R} h_t - \frac{1}{1+n} \tau_t + \omega_{\widetilde{\zeta}} \nu_t \right) \omega_{\widetilde{\zeta}} - B\omega_{\widetilde{\zeta}} = 0,$$

which means that

$$\nu_t = \frac{B + 2A \left(\widetilde{R} h_t - \frac{1}{1+n} \tau_t \right)}{2 \left(\vartheta - A\omega_{\widetilde{\zeta}}^2 \right)} \omega_{\widetilde{\zeta}}. \qquad (52)$$

Substituting (52) back into (51) gives

$$-Ah_t^2 - Bh_t - C = \max_{\tau_t} \left\{ -\frac{1}{2} (\tau_t + \varphi)^2 + \beta E_t \left(\vartheta \left[\frac{B + 2A \left(\widetilde{R} h_t - \frac{1}{1+n} \tau_t \right)}{2 \left(\vartheta - A\omega_{\widetilde{\zeta}}^2 \right)} \omega_{\widetilde{\zeta}} \right]^2 - Ah_{t+1}^2 - Bh_{t+1} - C \right) \right\},$$

The first-order condition for τ_t is

$$- (\tau_t + \varphi) - 2\beta\vartheta \frac{1}{1+n} \frac{A\omega_{\widetilde{\zeta}}}{\vartheta - A\omega_{\widetilde{\zeta}}^2} \nu_t + 2\beta A \frac{1}{1+n} \left(1 + \frac{A\omega_{\widetilde{\zeta}}^2}{\vartheta - A\omega_{\widetilde{\zeta}}^2} \right) \left(\widetilde{R} h_t - \frac{1}{1+n} \tau_t + \omega_{\widetilde{\zeta}} \nu_t \right)$$

$$+ \beta B \frac{1}{1+n} \left(1 + \frac{A\omega_{\widetilde{\zeta}}^2}{\vartheta - A\omega_{\widetilde{\zeta}}^2} \right) = 0.$$

24

Using the solution for ν_t the solution for taxation is

$$\tau_t = \frac{2\beta\widetilde{R}A}{(1+n) - (1+n)\,A\omega_\zeta^2/\vartheta + 2\beta A/\,(1+n)}h_t + \frac{-(1+n)\left(1 - A\omega_\zeta^2/\vartheta\right)\varphi + \beta B}{(1+n) - (1+n)\,A\omega_\zeta^2/\vartheta + 2\beta A/\,(1+n)}. \quad (53)$$

Substituting the above expressions into the Bellman equation gives

$$-Ah_t^2 - Bh_t - C$$

$$= -\frac{1}{2}\left[\frac{2\beta\widetilde{R}A}{(1+n) - (1+n)\,A\omega_\zeta^2/\vartheta + 2\beta A/\,(1+n)}h_t + \frac{\beta B + 2\beta\varphi A/\,(1+n)}{(1+n) - (1+n)\,A\omega_\zeta^2/\vartheta + 2\beta A/\,(1+n)}\right]^2$$

$$+ \frac{\beta\vartheta\omega_\zeta^2}{\left[2\left(\vartheta - A\omega_\zeta^2\right)\right]^2}\left\{\begin{array}{l}2A\widetilde{R}\left[\frac{(1+n)^2 - (1+n)^2 A\omega_\zeta^2/\vartheta}{(1+n)^2 - (1+n)^2 A\omega_\zeta^2/\vartheta + 2\beta A}\right]h_t \\ + \frac{\left((1+n)^2 B + 2A(1+n)\varphi\right)\left(1 - A\omega_\zeta^2/\vartheta\right)}{(1+n)^2 - (1+n)^2 A\omega_\zeta^2/\vartheta + 2\beta A}\end{array}\right\}^2$$

$$- \beta A\left\{\begin{array}{l}\left[\frac{(1+n)^2\widetilde{R}}{(1+n)^2 - (1+n)^2 A\omega_\zeta^2/\vartheta + 2\beta A}\right]^2 h_t^2 - \frac{(1+n)^2\widetilde{R}\left[-2(1+n)\varphi + 2\beta B - B\omega_\zeta^2/\vartheta(1+n)^2\right]}{\left[(1+n)^2 - (1+n)^2 A\omega_\zeta^2/\vartheta + 2\beta A\right]^2}h_t \\ + \frac{-2(1+n)\varphi + 2\beta B - B\omega_\zeta^2/\vartheta(1+n)^2}{2\left[(1+n)^2 - (1+n)^2 A\omega_\zeta^2/\vartheta + 2\beta A\right]} + \omega_\zeta^2\end{array}\right\}$$

$$- \beta B\left[\frac{(1+n)^2\,\widetilde{R}}{(1+n)^2 - (1+n)^2\,A\omega_\zeta^2/\vartheta + 2\beta A}h_t - \frac{-2\,(1+n)\,\varphi + 2\beta B - B\omega_\zeta^2/\vartheta\,(1+n)^2}{2\left((1+n)^2 - (1+n)^2\,A\omega_\zeta^2/\vartheta + 2\beta A\right)}\right] - \beta C.$$

Given $\beta\widetilde{R} = 1$, collecting and matching terms, the constant coefficients turn out to be

$$A = \frac{(1+n)^2\,\widetilde{R}\left(\widetilde{R} - 1\right)}{2 - (1+n)^2\,\widetilde{R}\omega_\zeta^2/\vartheta}, \quad (54)$$

$$B = \frac{(1+n)\,\varphi\widetilde{R}}{1 - (1+n)^2\,\widetilde{R}\omega_\zeta^2/\,(2\vartheta)}. \quad (55)$$

Substituting (54) and (55) into (53) yields the taxation function (22) in the text.

We impose parameter restrictions so that $A > 0$, implying the value function is concave; these restrictions amount to requiring that ϑ not be too small and are shown in the text to imply $\Sigma < 1$.

6.3 Deriving the Stochastic Properties of Optimal Taxation and Government Debts under RB

6.3.1 Deriving the Volatility of the Change in Taxes

Given (30), $\tau_{t+1} = \rho_h\tau_t + \frac{\widetilde{R}-1}{1-\Sigma}\zeta_{t+1} - \left(\widetilde{R} - 1\right)\frac{\Sigma\varphi}{1-\Sigma}$, the variance of $\Delta\tau_{t+1}$ can be written as:

$$\operatorname{var}\left(\Delta\tau_{t+1}\right) = \operatorname{var}\left(\left(\rho_h - 1\right)\tau_t + \frac{\widetilde{R}-1}{1-\Sigma}\zeta_{t+1}\right)$$

$$= \left(1-\rho_h\right)^2 \left(\frac{\widetilde{R}-1}{1-\Sigma}\right)^2 \frac{\omega_\zeta^2}{1-\rho_h^2} + \left(\frac{\widetilde{R}-1}{1-\Sigma}\right)^2 \omega_\zeta^2$$

$$= \left[1 + \frac{\left(1-\rho_h\right)^2}{1-\rho_h^2}\right]\left(\frac{\widetilde{R}-1}{1-\Sigma}\right)^2 \omega_\zeta^2$$

$$= \frac{2}{1+\rho_h}\left(\frac{\widetilde{R}-1}{1-\Sigma}\right)^2 \omega_\zeta^2.$$

Using the definition of the relative volatility of the change in taxes and government spending, we can obtain (32) in the text.

6.3.2 Deriving the Volatility of the Government Debt

Given (24), $def_t = \frac{1-\rho_g}{(1+n)\left(\widetilde{R}-\rho_g\right)}\left(g_t - \overline{g}\right) - \left(1-\rho_h\right)h_t - \frac{1}{1+n}\frac{\Sigma\varphi}{1-\Sigma}$,

$$\operatorname{var}\left(def_t\right) = \operatorname{var}\left(\frac{1}{1+n}\frac{1-\rho_g}{\widetilde{R}-\rho_g}g_t - \left(1-\rho_h\right)h_t\right)$$

$$= \left(\frac{1-\rho_g}{(1+n)\left(\widetilde{R}-\rho_g\right)}\right)^2 \operatorname{var}\left(g_t\right) + \left(1-\rho_h\right)^2 \operatorname{var}\left(h_t\right) - 2\frac{\left(1-\rho_g\right)\left(1-\rho_h\right)}{(1+n)\left(\widetilde{R}-\rho_g\right)}\operatorname{cov}\left(g_t, h_t\right)$$

$$= \left(\frac{1-\rho_g}{\widetilde{R}-\rho_g}\right)^2 \frac{\left(\widetilde{R}-\rho_g\right)^2 \omega_{\widetilde{\zeta}}^2}{1-\rho_g^2} + \left(1-\rho_h\right)^2 \frac{\omega_{\widetilde{\zeta}}^2}{1-\rho_h^2} - 2\frac{\left(1-\rho_g\right)\left(1-\rho_h\right)}{\widetilde{R}-\rho_g}\frac{\left(\widetilde{R}-\rho_g\right)\omega_{\widetilde{\zeta}}^2}{1-\rho_g\rho_h}$$

$$= \left[\frac{1-\rho_g}{1+\rho_g} + \frac{1-\rho_h}{1+\rho_h} - 2\frac{\left(1-\rho_h\right)\left(1-\rho_g\right)}{1-\rho_g\rho_h}\right]\omega_{\widetilde{\zeta}}^2, \tag{56}$$

where we use the facts that $\rho_h = 1 - \frac{\Sigma\left(\widetilde{R}-1\right)}{1-\Sigma}$, $\operatorname{var}\left(g_t\right) = \frac{\omega^2}{1-\rho_g^2} = \frac{\left(\widetilde{R}-\rho_g\right)^2\omega_{\widetilde{\zeta}}^2}{1-\rho_g^2}$, $\operatorname{var}\left(h_t\right) = \frac{\omega_{\widetilde{\zeta}}^2}{1-\rho_h^2}$, and $\operatorname{cov}\left(g_t, h_t\right) = \frac{(1+n)\left(\widetilde{R}-\rho_g\right)\omega_{\widetilde{\zeta}}^2}{1-\rho_g\rho_h}$. Using the definition of the relative volatility of the government deficit and government spending, we can obtain (??) in the text.

6.3.3 Deriving the Correlation between Government Debts and Spending

Given (24) and (11), the covariance between the government debt and spending is

$$
\begin{aligned}
\text{cov}\left(def_t, g_t\right) &= \text{cov}\left(\frac{1-\rho_g}{(1+n)\left(\widetilde{R}-\rho_g\right)} g_t - (1-\rho_h) h_t, g_t\right) \\
&= \frac{1-\rho_g}{(1+n)\left(\widetilde{R}-\rho_g\right)} \, \text{var}\left(g_t\right) - (1-\rho_h)\,\text{cov}\left(h_t, g_t\right) \\
&= \frac{1-\rho_g}{(1+n)\left(\widetilde{R}-\rho_g\right)} \frac{\left(\widetilde{R}-\rho_g\right)^2}{1-\rho_g^2}\omega_\zeta^2 - \frac{(1+n)\left(\widetilde{R}-\rho_g\right)(1-\rho_h)}{1-\rho_g\rho_h}\omega_\zeta^2 \\
&= (1+n)\left(\widetilde{R}-\rho_g\right)\left(\frac{1}{1+\rho_g} - \frac{1-\rho_h}{1-\rho_g\rho_h}\right)\omega_\zeta^2.
\end{aligned}
\tag{57}
$$

Using (56) and (57), the correlation between the current account and net income can be written as (33) in the text.

6.4 The Equivalence between the Univariate and Multivariate RB Models

The solution methods used to solve the univariate model and the standard multivariate model are different in the sense that the expectation operator applies to different objects. In the univariate model, the evil agent distorts the transition equation of h_t which by itself includes the expectation on future income, whereas in the multivariate model the evil agent distorts the government spending process (g_t). Note that in this case the evil agent does not distort the law of motion for b_t because it is a predetermined variable with no uncertainty given τ_t. Specifically, the following compact matrix equation can be used to characterize the dynamics of (b, g):

$$
\begin{bmatrix} b_{t+1} \\ g_{t+1} \end{bmatrix} = \begin{bmatrix} \frac{R}{1+n} & \frac{1}{1+n} \\ 0 & \rho_g \end{bmatrix} \begin{bmatrix} b_t \\ g_t \end{bmatrix} - \begin{bmatrix} \frac{1}{1+n} \\ 0 \end{bmatrix} \tau_t + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \epsilon_{t+1},
\tag{58}
$$

where we ignore the constant term as it does not affect the stochastic properties of the model. Assume that (58) is the approximation model. The corresponding distortion model is

$$
\begin{bmatrix} b_{t+1} \\ g_{t+1} \end{bmatrix} = \begin{bmatrix} \frac{R}{1+n} & \frac{1}{1+n} \\ 0 & \rho_g \end{bmatrix} \begin{bmatrix} b_t \\ g_t \end{bmatrix} - \begin{bmatrix} \frac{1}{1+n} \\ 0 \end{bmatrix} \tau_t + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \left(\epsilon_{t+1} + \omega\nu_t\right).
\tag{59}
$$

The robust control problem can thus be written as follows:

$$
v\left(b_t, g_t\right) = \min_{\tau_t} \max_{\nu_t} \left\{ -\frac{1}{2}\tau_t^2 + \beta\left[\vartheta\nu_t^2 + E_t\left[v\left(b_{t+1}, g_{t+1}\right)\right]\right] \right\},
\tag{60}
$$

subject to (59). Since there is no closed-form solution for this problem, we solve this model numerically. Following the standard procedure proposed in Hansen and Sargent (Chapters 2 and 10, 2007), we can solve this robust linear-quadratic regulator numerically and obtain a linear function of τ_t in terms of (b_t, g_t):

$$\tau_t = \text{MPT}\,(b_t + \lambda g_t), \tag{61}$$

where MPT is the marginal propensity to tax, λ measures the relative importance of government spending (g) and government debts (b) in determining optimal taxation. From 60), it is clear that in the multivariate model RB might affect the relative importance of the two state variables on the taxation function, while in the univariate model the relative importance of the two effects are *fixed* in order to reduce the state space. (i.e., $h_t = b_t + \frac{1}{R-\rho_g}g_t + \frac{(1-\rho_g)\overline{g}}{(R-\rho_g)(R-1)}$.) We now show that *numerically* the two modeling strategies can lead to the identical decision rules. Note that in our univariate model, after introducing the new state variable, the consumption function under RB can be written as

$$\tau_t = \text{MPT}\,h_t, \tag{62}$$

where we ignore the constant term; $\text{MPT} = \frac{R-1}{1-\Sigma}$, $h_t = b_t + \frac{1}{R-\rho_g}g_t$, and $\Sigma = R\omega_\zeta^2/(2\vartheta)$. Figure 6 illustrates how the marginal propensity of taxation (MPT) is affected by the degree of robustness in both the multivariate and univariate models when we set $R = 1.02$, $\rho_g = 0.6$, and $\omega^2 = 1$. The figure clearly shows that the two models deliver the identical MPT for various values of ϑ. For example, the multivariate model predicts that

$$\tau_t = 0.0353b_t + 0.0841g_t \tag{63}$$

when $1/\vartheta = 0.15$, whereas the univariate model predicts that

$$\tau_t = 0.0353h_t = 0.0353\,(b_t + 2.381g_t) = 0.0353b_t + 0.0841g_t \tag{64}$$

for the same value of ϑ (i.e., $\Sigma = 0.4337$). In addition, we find the relative importance of government spending and government debt is also identical in the two models. (See Figure 7 for an illustration.)

6.5   Deriving the Standard Errors for the VAR Test

In this appendix we derive expressions for the standard errors of the estimators for λ_1 and λ_2 used in the VAR test of Section 4.4. Let $d \equiv \frac{1}{R}$,

$$A \equiv \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix},$$

28

and I be the 2×2 identify matrix. Define

$$\Lambda \equiv \begin{bmatrix} \lambda_1 & \lambda_2 \end{bmatrix} = - \begin{bmatrix} 1 & 0 \end{bmatrix} Ad \, (I - dA)^{-1}. \tag{65}$$

Define

$$B \equiv (I - dA)^{-1} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix},$$

we have $B\,(I - dA) = I$, and we can solve for the coefficients of the B matrix:

$$b_{11} = \frac{1 - da_{22}}{(1 - da_{11})(1 - da_{22}) - d^2 a_{12} a_{21}},$$

$$b_{12} = \frac{da_{12}}{(1 - da_{11})(1 - da_{22}) - d^2 a_{12} a_{21}},$$

$$b_{21} = \frac{da_{21}}{(1 - da_{11})(1 - da_{22}) - d^2 a_{12} a_{21}},$$

$$b_{22} = \frac{1 - da_{11}}{(1 - da_{11})(1 - da_{22}) - d^2 a_{12} a_{21}}.$$

Now we substitute the B matrix into Equation (65), obtaining

$$\lambda_1 = -d \frac{a_{11}(1 - da_{22}) + da_{12}a_{21}}{(1 - da_{11})(1 - da_{22}) - d^2 a_{12} a_{21}} \equiv -d \frac{c_1}{c_2},$$

$$\lambda_2 = -\frac{da_{12}}{(1 - da_{11})(1 - da_{22}) - d^2 a_{12} a_{21}} \equiv -\frac{f_1}{f_2}.$$

Let the 2×4 matrix G denote the gradient of Λ with respect to the vectorization of VAR coefficient matrix A: $\mathrm{vec}(A) \equiv \begin{bmatrix} a_{11} & a_{21} & a_{12} & a_{22} \end{bmatrix}^T$, where $G(i,j) = \frac{\partial \lambda_i}{\partial vec(A)_j}$ $(i = 1, 2;\ j = 1, 2, 3, 4)$ and G is given by

$$G_{11} = -\frac{d}{c_2}(1 - da_{22})(1 - \lambda_1),\ G_{12} = -\frac{d^2 a_{12}}{c_2}\left(1 + d\frac{c_1}{c_2}\right),$$

$$G_{13} = -\frac{d^2 a_{21}}{c_2}\left(1 + d\frac{c_1}{c_2}\right),\ G_{14} = \frac{d^2}{c_2}\left(a_{11} - \frac{c_1}{c_2}(1 - da_{11})\right),$$

$$G_{21} = \frac{d}{f_2}\lambda_2(1 - da_{22}),\ G_{22} = \frac{d^2}{f_2}\lambda_2 a_{12},$$

$$G_{23} = -\frac{d}{f_2}(1 - \lambda_2 da_{21}),\ G_{24} = \frac{d}{f_2}\lambda_2(1 - da_{11}).$$

Finally, the covariance matrix of Λ is computed as

$$\mathrm{var}\,(\Lambda) = G\,\mathrm{var}\,(A)\,G^T,$$

where $\mathrm{var}\,(A)$ is the covariance matrix (4×4) of $\mathrm{vec}\,(A)$.

References

[1] Adler, Johan (2006), "The Tax-Smoothing Hypothesis: Evidence from Sweden, 1952-1999," *Scandinavian Journal of Economics* 108(1), 81-95.

[2] Aguiar, Mark and Gita Gopinath (2007), "Emerging Market Business Cycles: The Cycle Is the Trend," *Journal of Political Economy* 115(1), 69-102.

[3] Angeletos, George-Marios (2002), "Fiscal Policy with Non-Contingent Debt and the Optimal Maturity Structure," *Quarterly Journal of Economics* 117(3), 1105-1131.

[4] Aiyagari, S. Rao, Albert Marcet, Thomas J. Sargent, and Juha Seppala (2002), "Optimal Taxation without State-contingent Debt," *Journal of Political Economy* 110, 1220–1254.

[5] Barro, Robert J. (1979), "On the Determination of the Public Debt," *Journal of Political Economy* 87, 940-971.

[6] Barro, Robert J. (1981), "On the Predictability of Tax-rate Changes," NBER Working Paper 636; reprinted in Barro, R.J. (ed.) (1990), Macroeconomic policy, Harvard University Press, Cambridge, MA.

[7] Bidder, Rhys and Matthew E. Smith (2011), "Robust Control in a Nonlinear DSGE Model," manuscript.

[8] Bohn, Henning (1990), "Tax Smoothing with Financial Instruments," *American Economic Review* 80(5), 1217–1230.

[9] Bohn, Henning (1998), "The Behavior of US Public Debt and Deficits," *Quarterly Journal of Economics* 113, 949-963.

[10] Campbell, John Y. (1987), "Does Saving Anticipate Declining Labor Income? An Alternative Test of the Permanent Income Hypothesis," *Econometrica* 55, 1249-1273.

[11] Chari, V. V., Lawrence J. Christiano, and Patrick J. Kehoe (1994), "Optimal Fiscal Policy in a Business Cycle Model," *Journal of Political Economy*, CII, 617-652.

[12] Ghosh, Atish R. (1995), "Intertemporal Tax-smoothing and the Government Budget Surplus: Canada and the United States," *Journal of Money, Credit and Banking* 27, 1033–1045.

[13] Hansen, Lars Peter and Thomas J. Sargent (1995), "Discounted Linear Exponential Quadratic Gaussian Control," *IEEE Transactions on Automatic Control* 40, 968-971.

[14] Hansen, Lars Peter and Thomas J. Sargent (2007), *Robustness*, Princeton University Press.

[15] Hansen, Lars Peter, Thomas J. Sargent, and Thomas D. Tallarini, Jr. (1999), "Robust Permanent Income and Pricing," *Review of Economic Studies* 66(4), 873-907.

[16] Hansen, Lars Peter, Thomas J. Sargent, and Neng Wang (2002), "Robust Permanent Income and Pricing with Filtering," *Macroeconomic Dynamics* 6(1), 40-84.

[17] Huang, Chao-Hsi and Kenneth S. Lin (1993), "Deficits, Government Expenditures and Tax Smoothing in the United States: 1929–1988," *Journal of Monetary Economics* 31, 317–339.

[18] Jacobson, David H. (1973), "Optimal Stochastic Linear Systems with Exponential Performance Criteria and Their Relation to Deterministic Differential Games," *IEEE Transactions on Automatic Control* 18, 124-131.

[19] Karantounias, Anastasios G, Lars Peter Hansen, and Thomas J. Sargent (2009), "Managing Expectations and Fiscal Policy," forthcoming in *Theoretical Economics.*

[20] Lloyd-Ellis, Huw, Shiqiang Zhan, and Xiaodong Zhu (2002), "Tax Smoothing with Stochastic Interest Rates: A Re-assessment of Clinton's Fiscal Legacy," *Journal of Money, Credit, and Banking* 699-724.

[21] Lucas, Robert E., Jr., and Nancy L. Stokey (1983), "Optimal Fiscal and Monetary Policy in an Economy without Capital," *Journal of Monetary Economics*, XII, 55-93.

[22] Luo, Yulei (2008), "Consumption Dynamics under Information Processing Constraints," *Review of Economic Dynamics* 11(2), 366-385.

[23] Luo, Yulei and Eric R. Young (2010), "Risk-sensitive Consumption and Savings under Rational Inattention," *American Economic Journal: Macroeconomics* 2(4), 281-325.

[24] Luo, Yulei, Jun Nie, and Eric R. Young (2011), "Robustness, Information-processing Constraints, and the Current Accounts in Small Open Economies," forthcoming in *Journal of International Economics.*

[25] Roubini, Nouriel (1988), "Current Account and Budget Deficits in an Intertemporal Model of Consumption and Taxation Smoothing. A Solution to the Feldstein-Horioka Puzzle?" *NBER* Working Paper No. 2773.

[26] Sahasakul, Chaipat (1986), "The U.S. Evidence on Optimal Taxation over Time," *Journal of Monetary Economics* 18, 251-75.

[27] Sargent, Thomas J. (2001), "Comment on 'Fiscal Consequences for Mexico Adopting the Dollar' by Christopher A. Sims," *Journal of Money, Credit, and Banking* 33(2), 617-25.

[28] Trehan, Bharat and Carl E. Walsh (1988), "Common trends, the Government Budget Constraint, and Revenue Smoothing," *Journal of Economic Dynamics and Control* 12, 425-444.

[29] van der Ploeg, Frederick (1993), "A Closed-Form Solution for a Model of Precautionary Saving," *Review of Economic Studies* 60(2), 385-95.

[30] Whittle, Peter (1981), "Risk-Sensitive Linear/Quadratic/Gaussian Control," *Advances in Applied Probability* 13, 764-777.

[31] Young, Eric R. (2012), "Robust Policymaking in the Face of Sudden Stops," forthcoming, *Journal of Monetary Economics.*

[32] Zhu, Xiaodong (1992), "Optimal Fiscal Policy in a Stochastic Growth Model," *Journal of Economic Theory* 58, 250-89.
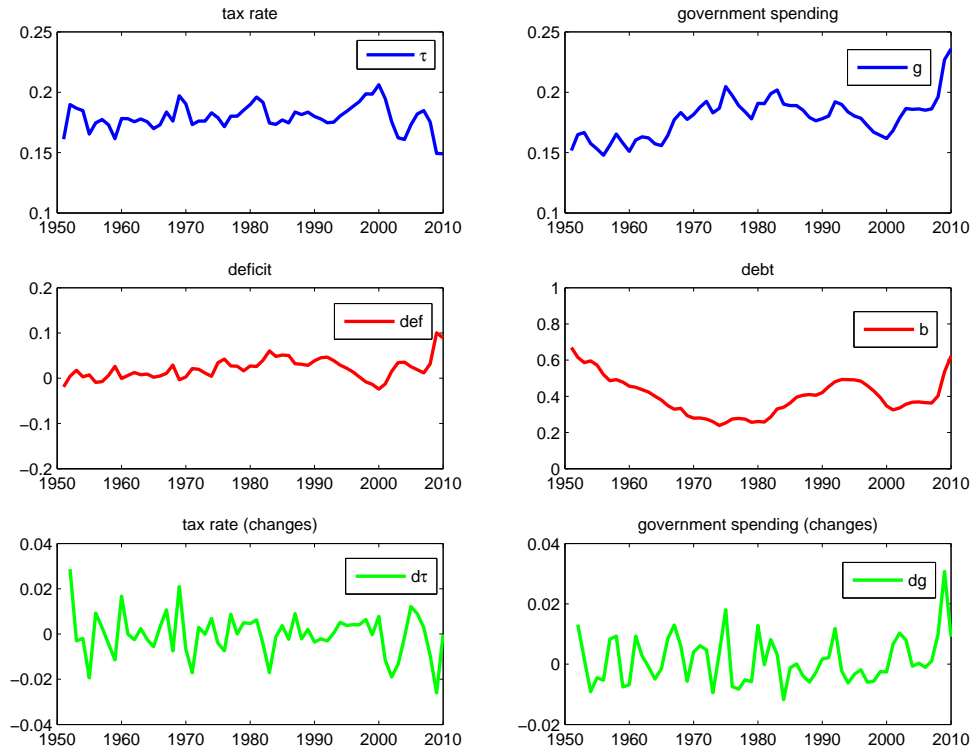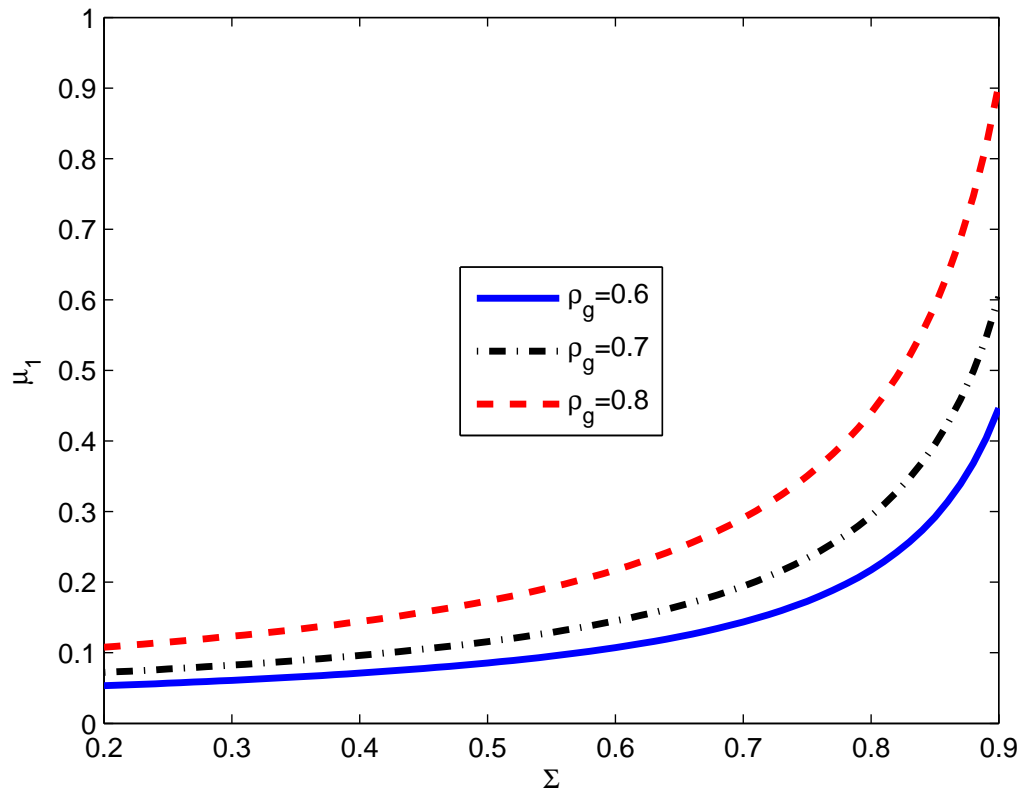
Figure 1: US Fiscal Policy

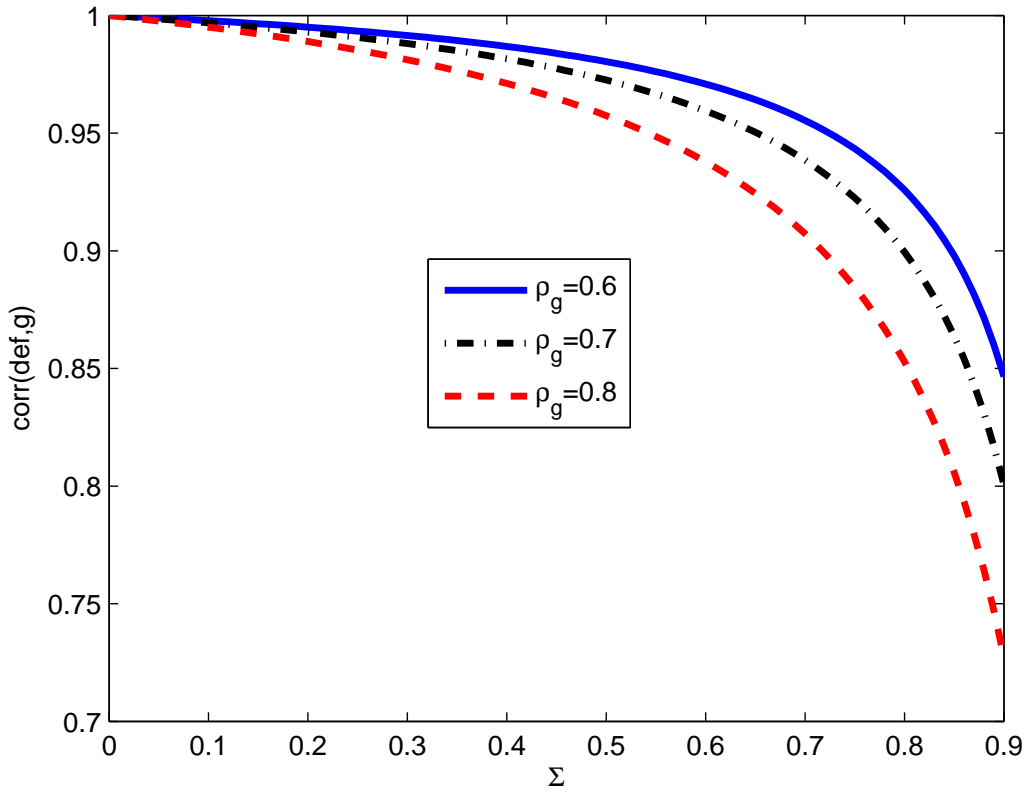Figure 2: The Relative Volatility of Tax Change to Government Spending

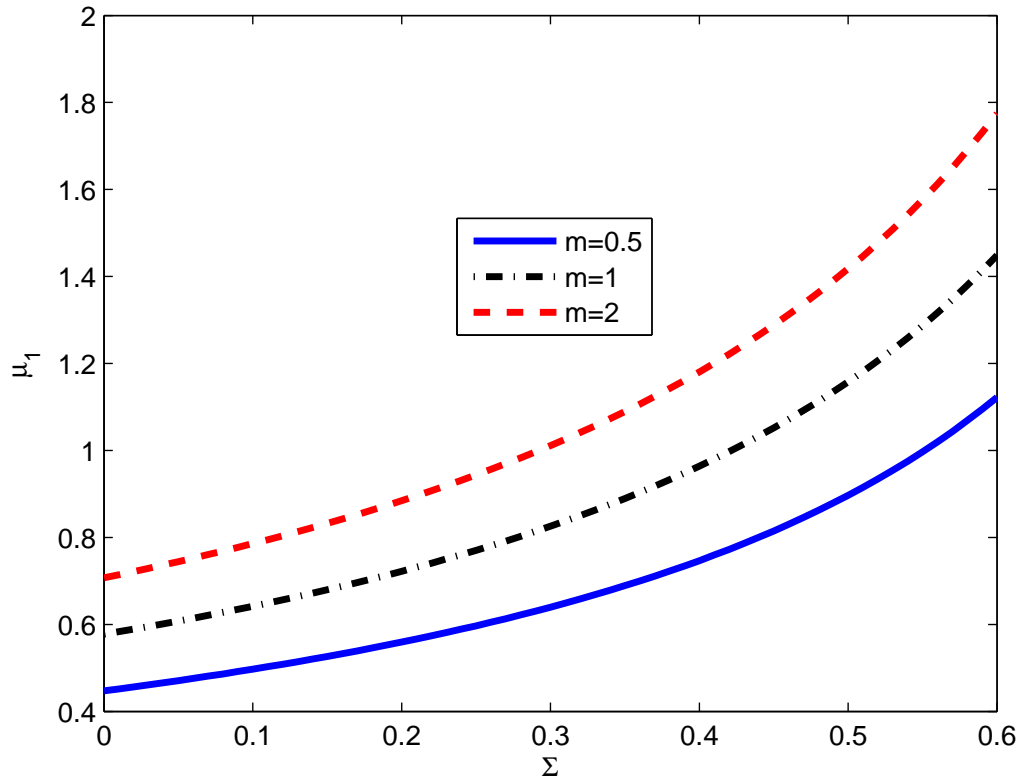Figure 3: The Correlation between the Government Deficit and Government Spending

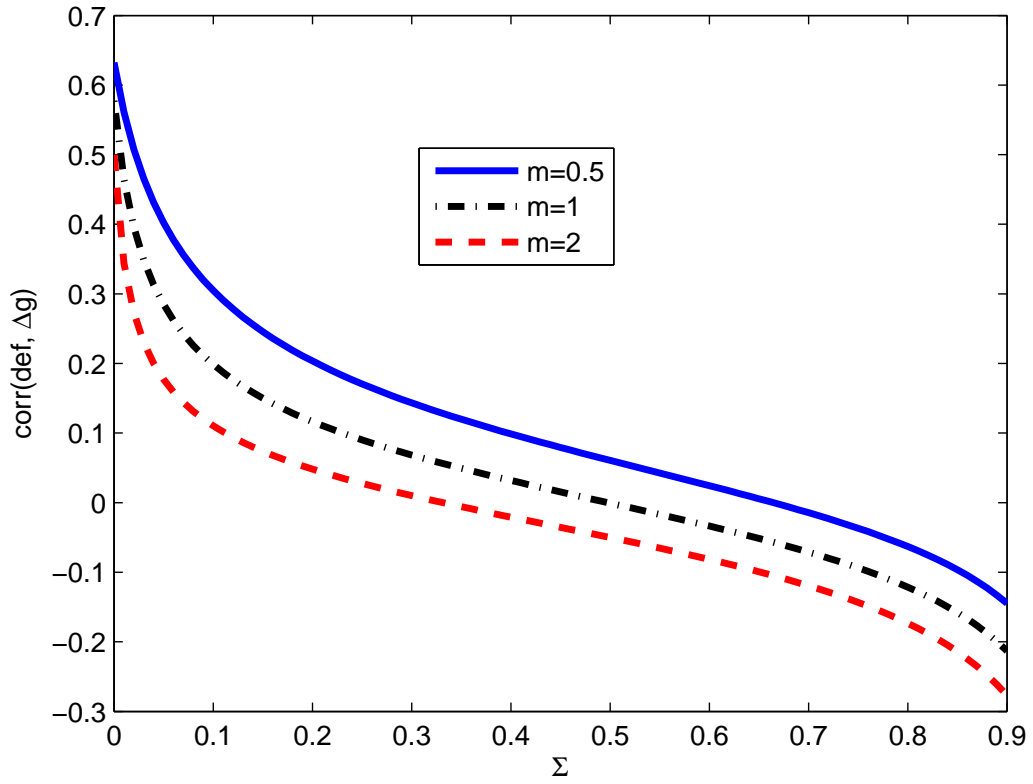Figure 4: The Relative Volatility of Tax Change to Government Spending

Figure 5: The Correlation between the Government Deficit and Government Spending
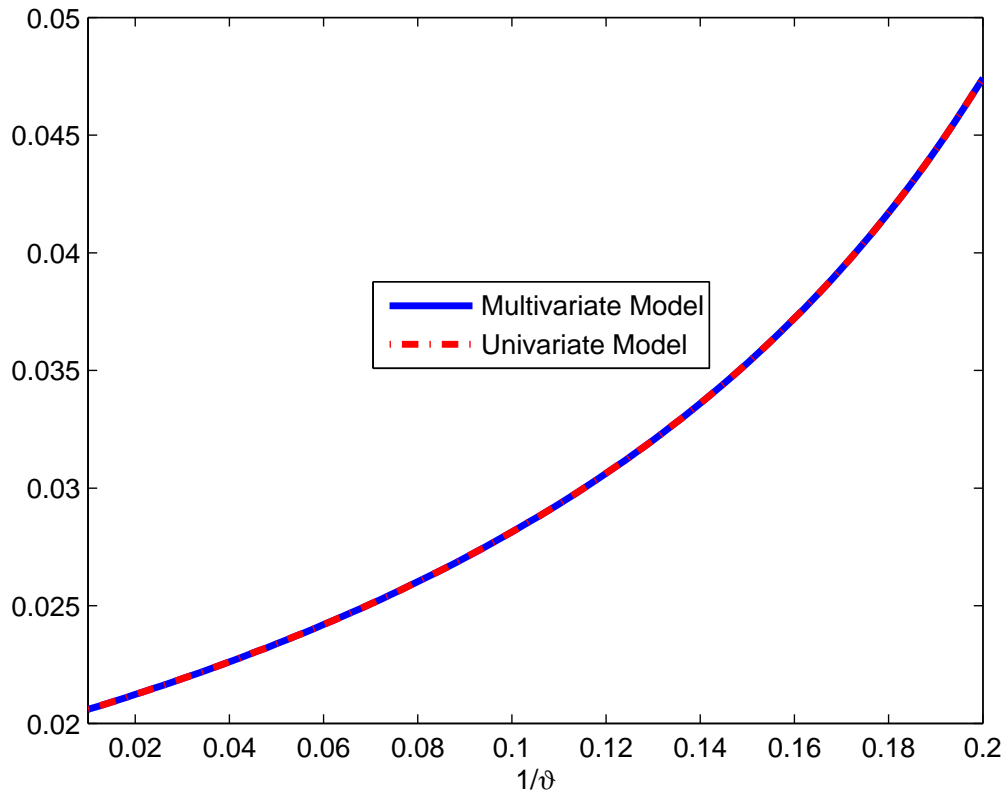
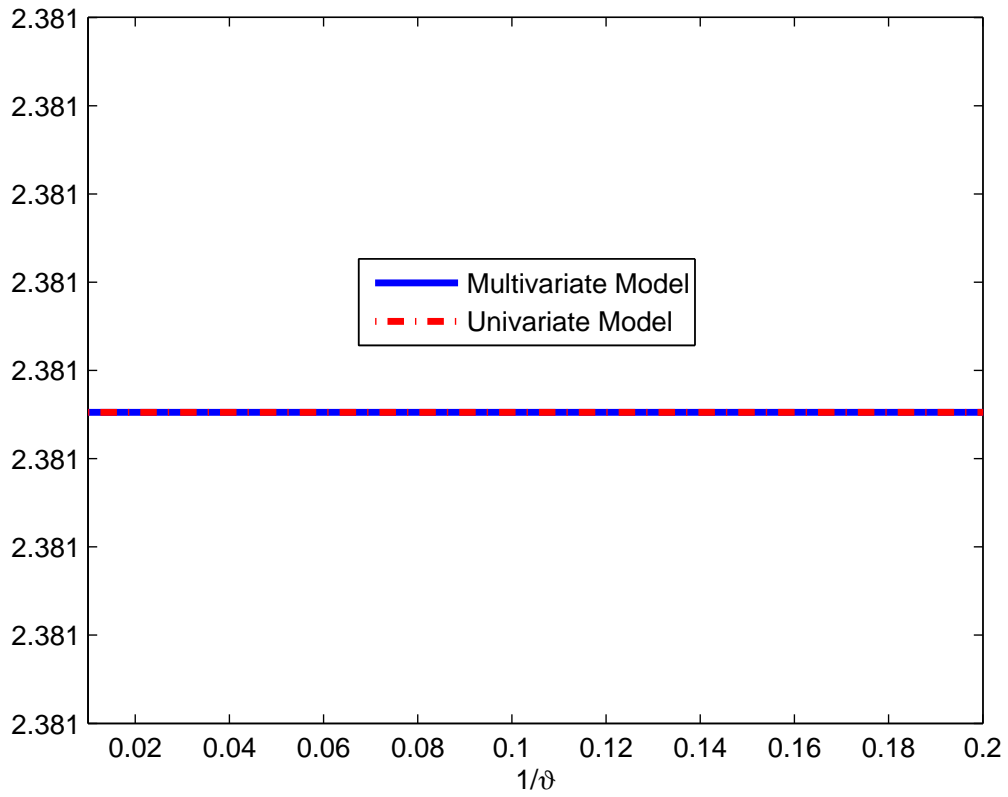Figure 6: The Equivalence between the Multivariate and Univariate Models

Figure 7: The Equivalence between the Multivariate and Univariate Models

Table 1: Key Moments on Relative Volatility and Correlation

	Linear Filter	HP Filter
$\frac{sd(\Delta\tau)}{sd(\Delta g)}$	1.21(0.15)	1.21(0.15)
$corr(def, g)$	0.72(0.10)	0.57(0.09)
$corr(def, \Delta g)$	0.31(0.20)	0.31(0.20)
ρ_g(1951-2010)	0.83(0.04)	0.46(0.07)
ρ_g(1960-2010)	0.84(0.04)	0.50(0.07)

Table 2: Calibration Results

Variable	Value
Σ	0.95
p	0.10
ρ_g	0.83
$\frac{\sigma(g)}{\mu(g)}$	0.08
$\frac{\sigma(\zeta)}{\mu(g)}$	0.24

Table 3: Model Comparison $(p = 0.1)$

	Data	RE	RB
$\frac{sd(\Delta\tau)}{sd(\Delta g)}$	1.21	0.04	0.85
$corr(def, g)$	0.72	1.00	0.74

Table 4: Calibration: Varying The Detection Error Probability

	$p = 0.09$	$p = 0.10$	$p = 0.11$	$p = 0.12$	$p = 0.13$
Σ	0.952	0.949	0.944	0.941	0.936

Table 5: Robust Check: Model Comparisons

	Data	RE	RB $(p = 0.09)$	RB $(p = 0.10)$	RB $(p = 0.11)$	RB $(p = 0.12)$	RB $(p = 0.13)$
$\frac{sd(\Delta\tau)}{sd(\Delta g)}$	1.21	0.04	0.90	0.85	0.79	0.73	0.69
$corr(def, g)$	0.72	1.00	0.73	0.74	0.75	0.76	0.78

Table 6: Key Moments on Relative Volatility and Correlation (1960-2010)

	Linear Filter	HP Filter
$\frac{sd(\Delta\tau)}{sd(\Delta g)}$	1.09(0.15)	1.09(0.15)
$corr(def, g)$	0.78(0.10)	0.62(0.09)

Table 7: Model Comparison ($p = 0.1$, 1960-2010)

	Data	RE	RB
$\frac{sd(\Delta\tau)}{sd(\Delta g)}$	1.085	0.045	0.518
$corr(def, g)$	0.779	1.000	0.821

41

Table 8: Results of The VAR Test (**1951** − 2010)

Model	λ_1	λ_2	$\chi^2_W(2)$	p-value
RE	$-0.09(0.21)$	$0.34(0.21)$	20.04	0.00
$\Sigma = 0.1$	$-0.08(0.21)$	$0.37(0.21)$	19.48	0.00
$\Sigma = 0.2$	$-0.07(0.21)$	$0.40(0.20)$	18.28	0.00
$\Sigma = 0.3$	$-0.07(0.21)$	$0.43(0.19)$	16.41	0.00
$\Sigma = 0.4$	$-0.06(0.22)$	$0.47(0.18)$	14.03	0.00
$\Sigma = 0.5$	$-0.06(0.22)$	$0.51(0.18)$	11.36	0.00
$\Sigma = 0.6$	$-0.05(0.22)$	$0.55(0.17)$	8.70	0.01
$\Sigma = 0.7$	$-0.05(0.22)$	$0.60(0.17)$	6.28	0.04
$\Sigma = 0.8$	$-0.05(0.23)$	$0.66(0.17)$	4.28	0.12
$\Sigma = 0.9$	$-0.05(0.23)$	$0.71(0.18)$	2.78	0.25
$\Sigma = 0.95$	$-0.06(0.23)$	$0.73(0.18)$	2.23	0.33

Note: $\chi^2_W(2)$ refers to the Wald statistics for the joint hypothesis test: $\lambda_1 = 0,\ \lambda_2 = 1$.

Table 9: Results of The VAR Test $(1960 - 2010)$

Model	λ_1	λ_2	$\chi^2_W(2)$	p-value
RE	$-0.17(0.26)$	$0.38(0.29)$	8.04	0.02
$\Sigma = 0.1$	$-0.16(0.26)$	$0.42(0.27)$	7.66	0.02
$\Sigma = 0.2$	$-0.15(0.26)$	$0.46(0.26)$	6.81	0.03
$\Sigma = 0.3$	$-0.14(0.26)$	$0.50(0.25)$	5.52	0.06
$\Sigma = 0.4$	$-0.13(0.27)$	$0.56(0.24)$	3.97	0.14
$\Sigma = 0.5$	$-0.12(0.27)$	$0.63(0.24)$	2.45	0.29
$\Sigma = 0.6$	$-0.10(0.28)$	$0.73(0.25)$	1.19	0.55
$\Sigma = 0.7$	$-0.08(0.28)$	$0.85(0.30)$	0.35	0.84
$\Sigma = 0.8$	$-0.05(0.29)$	$1.02(0.39)$	0.03	0.99
$\Sigma = 0.9$	$-0.00(0.29)$	$1.28(0.55)$	0.33	0.85
$\Sigma = 0.91$	$-0.00(0.29)$	$1.31(0.57)$	0.40	0.82

Note: $\chi^2_W(2)$ refers to the Wald statistics for the joint hypothesis test: $\lambda_1 = 0$, $\lambda_2 = 1$.