Identifying U.S. Fiscal Policy Behavior

Bing Li∗
Department of Economics
Indiana University
August 6, 2008

Abstract

To estimate a single-equation tax rule by Ordinary Least Squares (OLS) method has been one of the common approaches to identify U.S. fiscal policy behavior. However, OLS estimation of the tax rule is subject to simultaneity bias because the single rule is isolated from a system of equations implied by general equilibrium theory. This paper investigates the simultaneity bias problem with a simple DSGE model. The simultaneity bias is derived analytically, which turns out to be ubiquitous for both non-Ricardian and Ricardian equilibria. As a consequence, OLS-based identification of fiscal policy behavior is unreliable in small samples, which is illustrated in Monte Carlo experiments. Therefore, Vector Autoregression (VAR) subject to restrictions implied by the intertemporal government budget constraint is proposed, which avoids the simultaneity bias and provides reliable identification of fiscal policy behavior. The simultaneity bias problem associated with OLS estimation always exists in general equilibrium framework, which calls for extra caution of the empirical macroeconomists.

∗Contact Information: Department of Economics, Indiana University, Wylie Hall 344, 100 S. Woodlawn, Bloomington, IN 47405. Phone Number:(812)855-5308. Email: bl1@indiana.edu. I am especially grateful to Eric Leeper, my advisor, for his valuable advice, guidance and encouragement during the formation of this paper. I am also grateful to Hess Chung, Nora Traum, Tim Cogley, Todd Walker, Rusty Tchernis and Pei Pei for their valuable comments and suggestions. Special thanks to Nora Traum and Hess Chung for sharing their Matlab codes and unpublished papers with me. This version is incomplete. All remaining errors are my own.
1 Introduction

In the current literature, U.S. fiscal policy is usually characterized by a simple tax rule, just as the monetary policy being characterized by a Taylor rule. Generally, the tax rule links primary surplus to government debt in the following form

\[ s_t = \rho d_{t-1} + \alpha' Z_t + \epsilon_t \]  

(1)

where \( s_t \) is real primary surplus, \( d_{t-1} \) is lagged real government debt, \( Z_t \) a vector of other variables and \( \epsilon_t \) is fiscal policy shock. In empirical work, applying Ordinary Least Squares (OLS) method on the tax rule has been one of the common approaches to identify U.S. fiscal policy behavior, as in Bohn (1998). However, general equilibrium theory implies that the intertemporal government budget constraint

\[ d_{t-1} = \text{Expected Present Value of (Future Surpluses + Future Seigniorage)} \]  

(2)

always holds as an equilibrium condition regardless of the tax rule in place. So there are two channels linking government debt and future surpluses in equilibrium. If an econometrician runs OLS regression on (1) using equilibrium data without controlling for other conditions such as (2), OLS estimator of \( \rho \) will capture correlation information between \( s_t \) and \( d_{t-1} \) from both (1) and (2), which makes fiscal policy behavior implied by (1) indistinguishable. Under this circumstance, OLS estimator of \( \rho \) is subject to simultaneity bias and the corresponding identification of fiscal policy behavior may be misleading\(^1\).

The result in Bohn (1998), which is based on OLS regression, favors Ricardian interpretation of post-war U.S. data more than a non-Ricardian one\(^2\). Recently, Favero and Monacelli (2003) and Sala (2004) estimate Vector Autoregressions (VAR) with both monetary and fiscal policy rules, whose results imply that U.S. post-war fiscal

---

\(^1\)Some people have pointed out this identification problem associated with OLS regression on the tax rule. A non-exhaustive list includes Cochrane (1998), Woodford (1998) and Davig, Leeper and Chung (2007).

\(^2\)According to Leeper (1991)'s terminology, a Ricardian equilibrium is when monetary policy is active and fiscal policy is passive. A non-Ricardian equilibrium is when monetary policy is passive and fiscal policy is active.
regime was very variable, which is also confirmed by the estimation of regime-switching policy rules as in Favero and Monacelli (2005), Davig and Leeper (2006) and Davig, Leeper and Chung (2007). More precisely, these recent results imply that U.S. fiscal policy from the 1960s throughout the 1980s may be characterized as active, gradually switching to passive in the early 1990s and switching back to active in early 2001. Therefore, the result in Bohn (1998) at best characterizes the average behavior of U.S. post-war fiscal policy, or even worse, because the OLS-based identification of fiscal policy behavior is potentially misleading due to simultaneity bias.

In this paper, we investigate the simultaneity bias problem within general equilibrium framework and try to find an estimation scheme for reliable identification of fiscal policy behavior. We specify an illustrative DSGE model similar to Leeper (1991) and (2005) and treat it as data-generating process (DGP). Our first contribution is to derive the simultaneity bias analytically. Through Monte Carlo experiments, we illustrate the negative effect of simultaneity bias on identification of fiscal policy behavior. Our second contribution is to identify fiscal policy behavior in VAR framework, where a restriction implied by equilibrium condition like (2) is imposed.

A couple of results from the analysis with the simple DSGE model are worth mentioning here. Firstly, with serially correlated fiscal policy shock, the simultaneity bias is ubiquitous for various DGPs. Specifically, when DGP is non-Ricardian, the bias can be negative, positive or zero; when DGP is Ricardian, the bias is always negative. Secondly, regardless of the underlying DGP, OLS-based identification of fiscal policy behavior in small samples is unreliable due to the simultaneity bias. Thirdly, monetary and fiscal policy interaction matters for identification of fiscal policy behavior. For example, with non-Ricardian DGP, as monetary policy pays less attention to inflation, OLS estimator of the tax rule would be more biased up so that the econometrician would mistakenly conclude the DGP as Ricardian with higher probability. Fourthly, if DGP is non-Ricardian, unrestricted VAR provides more reliable inference on fiscal policy behavior. While if DGP is Ricardian, unrestricted VAR is not superior to OLS regression.

This paper is organized as follows. In section 2, we specify and solve the simple DSGE model. We also specify the state-space representation of the model, which
connects the model to data. In section 3, we derive the simultaneity bias analytically, discuss economic intuition underlying the bias and report small-sample Monte Carlo experiment results. In section 4, we derive the restrictions to be imposed on the VAR, sketch the estimation procedure and report small-sample Monte Carlo experiment results. Section 5 concludes.

2 The Illustrative Model

2.1 Model Setup

This is an endowment economy model which is similar to Leeper (1991) and (2005). In the model, there is an infinitely lived representative agent who chooses sequences \{c_t, M_t, B_t\}_{t=0}^{\infty} to solve the household problem:

\[
\max E_0 \sum_{t=0}^{\infty} \beta^t [\ln(c_t) + \delta \ln(M_t/P_t)]
\] (3)

subject to the flow budget constraint

\[
c_t + \frac{M_t + B_t}{P_t} + \tau_t = y + \frac{M_{t-1} + R_{t-1}B_{t-1}}{P_t}
\] (4)

taking the initial liabilities \( M_{-1} + R_{-1}B_{-1} > 0 \) and sequences \{y, \tau_t, R_t, P_t\}_{t=0}^{\infty} as given. In (3) and (4), \( \beta \in (0, 1) \) is discount factor, \( \delta \in (0, \infty) \) is weight on real money balance in the utility function, \( c_t \) is real consumption, \( M_t \) is nominal money balance, \( B_t \) is nominal one-period government debt with gross nominal interest rate \( R_t \), \( P_t \) is price level, \( y \) is constant endowment and \( \tau_t \) is lump-sum taxes (if positive) or transfers (if negative).

There is a government with policy sequences \{M_t, B_t, \tau_t\}_{t=0}^{\infty} subject to the government budget constraint

\[
\frac{M_t + B_t}{P_t} + \tau_t = g_t + \frac{M_{t-1} + R_{t-1}B_{t-1}}{P_t}
\] (5)

where \( g_t \) is government spending.

The resource constraint is

\[
c_t + g_t = y
\] (6)
For simplicity, we assume $g_t = 0$ for all $t$. So (6) reduces to $c_t = c = y$, which is the goods market clearing condition.

We obtain the following Fisher and money-demand relations from the standard first-order necessary conditions and the resource constraint (6):

$$\frac{1}{R_t} = \beta E_t \left[ \frac{1}{\pi_{t+1}} \right]$$

$$m_t = \delta c \left[ \frac{R_t}{R_t - 1} \right]$$

where $\pi_{t+1} \equiv P_{t+1}/P_t$ and $m_t \equiv M_t/P_t$ are inflation rate and real money balance, respectively.

The model is closed by specifying monetary and fiscal policies, which are characterized by simple rules that determine their corresponding policy instruments. Monetary policy is described by the Taylor-type interest rate rule

$$R_t = e^{\alpha_0 \pi_t} \theta_t$$

and fiscal policy is described by the tax rule

$$\tau_t = e^{\gamma_0 b_{t-1}^\gamma} \psi_t$$

where $b_{t-1} \equiv B_{t-1}/P_{t-1}$ is real government debt. We assume that the exogenous monetary and fiscal policy shocks $\theta_t$ and $\psi_t$ have unit means and their logs follow AR(1) process

$$\ln(\theta_t) = \rho_\theta \ln(\theta_{t-1}) + \varepsilon_{\theta_t}$$

$$\ln(\psi_t) = \rho_\psi \ln(\psi_{t-1}) + \varepsilon_{\psi_t}$$

where $|\rho_\theta| < 1$ and $|\rho_\psi| < 1$ for stationarity. Both $\varepsilon_{\theta_t}$ and $\varepsilon_{\psi_t}$ are random variables with zero means and bounded support. Their standard deviations are $\sigma_\theta$ and $\sigma_\psi$, respectively. We assume that $\varepsilon_\theta$ and $\varepsilon_\psi$ are both serially uncorrelated and uncorrelated between each other for all leads and lags.
2.2 Model Solution

The general equilibrium of the model is fully characterized by (5), (7)-(12). To solve the model, we firstly log-linearize the equilibrium conditions around the deterministic steady state. The linearized system is organized in the standard matrix form

\[
\Gamma_0 Y_{t+1} = \Gamma_1 Y_t + \Pi \eta_{t+1} + \Psi \varepsilon_{t+1}
\]  

(13)

where \(Y_{t+1} = \left[\hat{\pi}_{t+1}, \hat{b}_{t+1}, \hat{\theta}_{t+1}, \hat{\psi}_{t+1}\right]'\), \(\varepsilon_{t+1} = [\varepsilon_{\theta_{t+1}}, \varepsilon_{\psi_{t+1}}]'\). Any variable \(\hat{x}_t\) denotes log deviation of \(x_t\) from its corresponding steady state value \(x\), i.e. \(\hat{x}_t \equiv \ln(x_t) - \ln(x)\). We also define the one-period-ahead endogenous forecast error \(\eta_{t+1} \equiv \hat{\pi}_{t+1} - E_t \hat{\pi}_{t+1}\).

Since \(\Gamma_0\) is invertible, (13) can be expressed as

\[
Y_{t+1} = \Gamma_1^* Y_t + \Pi^* \eta_{t+1} + \Psi^* \varepsilon_{t+1}
\]  

(14)

where \(\Gamma_1^* = \Gamma_0^{-1} \Gamma_1\), \(\Pi^* = \Gamma_0^{-1} \Pi\) and \(\Psi^* = \Gamma_0^{-1} \Psi\). Determinacy of bounded equilibrium of the model hinges on the eigenvalues of \(\Gamma_1^*\), which are \([\alpha, \beta^{-1} - \gamma(\beta^{-1} - 1), \rho_\theta, \rho_\psi]\).

According to Leeper (1991) and (2005), we are able to characterize four different regions in the first quadrant of \((\alpha, \gamma)\) space, two of which indicate determinacy: (1) When \(\alpha < 1\) and \(\gamma < 1\), monetary policy is passive and fiscal policy is active. The equilibrium is determinate and non-Ricardian; (2) When \(\alpha > 1\) and \(\gamma > 1\), monetary policy is active and fiscal policy is passive. The equilibrium is determinate and Ricardian; (3) When \(\alpha < 1\) and \(\gamma > 1\), monetary and fiscal policies are both passive and bounded equilibrium is indeterminate; (4) When \(\alpha > 1\) and \(\gamma < 1\), monetary and fiscal policies are both active and no bounded equilibrium exists. Figure 1 demonstrates these four regions, where determinacy regions are superimposed by red lines.

We apply Jordan decomposition to obtain analytical solution of the model\(^4\). If the bounded equilibrium is determinate, the following conditions that suppress the unstable root of the system must hold for all \(t\):

\[
P^\pi Y_t = 0
\]  

(15)

\[
P^\pi \Pi^* \eta_{t+1} + P^\pi \Psi^* \varepsilon_{t+1} = 0
\]  

(16)

\(^3\)See Appendix A.

\(^4\)See Appendix A.
where $P^i$ is the $i$th row of $P^{-1}$ and the $i$th eigenvalue in $\Lambda$ is unstable. Both $P^{-1}$ and $\Lambda$ are defined in Appendix A. Specifically to this model, $i = 1$ indicates Ricardian equilibrium and 2 indicates non-Ricardian equilibrium. In both cases, $\eta_{t+1}$ is uniquely determined by (16). If the bounded equilibrium is indeterminate, there is no unique mapping between $\eta_{t+1}$ and $\varepsilon_{t+1}$. For the remaining part of the paper, we will focus on the determinacy regions. The solution outlined here will be used later to derive the analytical form of the simultaneity bias associated with surplus-debt regression.

### 2.3 State-Space Representation

In order to implement Monte Carlo experiments, we define the state-space representation of the model. The state equation, given by (17), can be obtained with Sims’ (2001) method and gensys algorithm.

$$Y_{t+1} = G_1 Y_t + M \varepsilon_{t+1}$$

Matrices $G_1$ and $M$ in (17) are outputs from gensys algorithm. Since both the model and the data are quarterly, we define the observation equation as

$$\begin{bmatrix} \pi_{t+1}^o \\ b_{t+1}^o \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \hat{\pi}_{t+1} \\ \hat{b}_{t+1} \\ \hat{\theta}_{t+1} \\ \psi_{t+1} \end{bmatrix}$$

where superscript “o” denotes observable variables.

Based on (17) and (18), we are able to simulate data sets, which are employed in the Monte Carlo experiments.

### 3 Statistical Inference I: Surplus-Debt Regression

Let us suppose the DSGE model specified in the previous section is the underlying DGP and an econometrician wants to draw statistical inference on fiscal policy behavior. We assume the data the econometrician uses is in logarithm, which is denoted by $\tilde{x}_t \equiv \ln(x_t)$ for any variable $x_t$. Suppose the econometrician chooses to regress
primary surplus $\tilde{\tau}_t^o$ on lagged government debt $\tilde{b}_{t-1}^o$, a commonly-used approach in the literature. Because observable variables are equivalent to model variables according to (18), we will suppress the superscript “o” henceforth.

Taking log on both sides of (10), we get

$$\tilde{\tau}_t = \gamma_0 + \gamma \tilde{b}_{t-1} + \tilde{\psi}_t$$ (19)

For simplicity, we assume the econometrician just estimates the correctly specified tax rule (19) and tries to draw inference on fiscal behavior based on $\hat{\gamma}_{OLS}$, the OLS estimator of $\gamma$.

Mathematically, the probability limit of $\hat{\gamma}_{OLS}$ can be shown as

$$\text{plim} \hat{\gamma}_{OLS} = \gamma + \frac{\text{cov}(\tilde{b}_{t-1}, \tilde{\psi}_t)}{\text{var}(\tilde{b}_{t-1})}$$ (20)

where $\text{cov}(\cdot, \cdot)$ and $\text{var}(\cdot)$ are covariance and variance operators, respectively. It is clear in (20) that the probability limit of $\hat{\gamma}_{OLS}$ contains an additional term besides $\gamma$. Later we will illustrate that the additional term is nonzero ubiquitously in the parameter space, only with exceptions in some special cases. When the econometrician runs OLS regression on (19), he actually has isolated (19) from the general equilibrium structure of the model, which is characterized by a system of equations. So $\hat{\gamma}_{OLS}$ is intrinsically subject to simultaneity bias, which, asymptotically, is the additional term in (20).

In this section, we investigate the simultaneity bias in four cases. DGP of the first three cases is non-Ricardian and that of the last case is Ricardian. For each case, we also discuss some economic intuition underlying the bias. At the end of this section, we report small-sample Monte Carlo simulation results which show that for a wide range of parameter values, simultaneity bias associated with surplus-debt regression makes it very difficult to correctly identify fiscal policy behavior.

3.1 Case I (Non-Ricardian): $\alpha = \gamma = \rho_\theta = \rho_\psi = 0$

In this case, (1) nominal interest rate is pegged; (2) primary surplus does not respond to government debt; (3) both monetary and fiscal policy shocks are serially uncorre-

---

$^5$See Appendix B.
lated. Algebra is greatly simplified under these assumptions. Firstly, condition (15) reduces to
\[ \tilde{b}_t = \ln(b) + \beta(\varphi_1 + \varphi_4)\varepsilon_{\vartheta_t} \]  
(21)
Since \( \tilde{\psi}_t \) is i.i.d., we have \( \tilde{\psi}_t = \varepsilon_{\psi_t} \). Then the covariance term in (20) becomes
\[ \text{cov}(\tilde{b}_{t-1}, \tilde{\psi}_t) = \text{cov} \left[ \ln(b) + \beta(\varphi_1 + \varphi_4)\varepsilon_{\vartheta_{t-1}}, \varepsilon_{\psi_{t}} \right] = 0 \]  
(22)
where the covariance is 0 because we have assumed \( \varepsilon_{\vartheta} \) and \( \varepsilon_{\psi} \) are uncorrelated between each other for all leads and lags. Consequently, \( \text{plim} \hat{\gamma}_{OLS} = \gamma \) and OLS estimator of \( \gamma \) is free of simultaneity bias, which guarantees reliable identification of fiscal policy behavior. However, to assume fiscal policy shock as serially uncorrelated is unrealistic, considering the empirical evidence on U.S. fiscal policy.

3.2 Case II (Non-Ricardian): \( \alpha = \gamma = \rho_{\theta} = 0, \rho_{\psi} \neq 0 \)

In this case, we keep the assumptions of case I except that fiscal policy shock \( \tilde{\psi}_t \) becomes serially correlated. Firstly, condition (15) reduces to
\[ \tilde{b}_t = \ln(b) + \beta(\varphi_1 + \varphi_4)\tilde{\theta}_t + \left[ \frac{(\beta^{-1} - 1)\rho_{\psi}}{\beta^{-1} - \rho_{\psi}} \right] \tilde{\psi}_t \]
\[ = \ln(b) + \beta(\varphi_1 + \varphi_4)\varepsilon_{\vartheta_t} + \left[ \frac{(\beta^{-1} - 1)\rho_{\psi}}{\beta^{-1} - \rho_{\psi}} \right] \frac{\varepsilon_{\psi_t}}{(1 - \rho_{\psi}L)} \]  
(23)
where we rewrite \( \tilde{\psi}_t \) as \( \varepsilon_{\psi_t}/(1 - \rho_{\psi}L) \) by introducing lag operator. Then the covariance term in (20) becomes
\[ \text{cov}(\tilde{b}_{t-1}, \tilde{\psi}_t) = \text{cov} \left[ \ln(b) + \beta(\varphi_1 + \varphi_4)\varepsilon_{\vartheta_t} + \frac{(\beta^{-1} - 1)\rho_{\psi}}{\beta^{-1} - \rho_{\psi}} \frac{\varepsilon_{\psi_{t-1}}}{(1 - \rho_{\psi}L)}, \frac{\varepsilon_{\psi_{t}}}{(1 - \rho_{\psi}L)} \right] \]
\[ = \left[ \frac{(\beta^{-1} - 1)\rho_{\psi}}{\beta^{-1} - \rho_{\psi}} \right] \text{cov} \left[ \frac{\varepsilon_{\psi_{t-1}}}{(1 - \rho_{\psi}L)}, \frac{\varepsilon_{\psi_{t}}}{(1 - \rho_{\psi}L)} \right] \]
\[ = \left[ \frac{(\beta^{-1} - 1)\sigma_{\psi}^2}{\beta^{-1} - \rho_{\psi}} \right] \sum_{i=1}^{\infty} \rho_{\psi}^i \]
\[ = \frac{(\beta^{-1} - 1)\rho_{\psi}^2\sigma_{\psi}^2}{(\beta^{-1} - \rho_{\psi})(1 - \rho_{\psi}^2)} \]  
(24)
Similarly, the variance term in (20) can be shown as

\[
\mathrm{var}(\tilde{b}_{t-1}) = \mathrm{cov}(\tilde{b}_{t-1}, \tilde{b}_{t-1}) = \beta^2 (\varphi_1 + \varphi_4)^2 \sigma_\theta^2 + \frac{(\beta^{-1} - 1)^2 \rho_\psi^2 \sigma_\psi^2}{(\beta^{-1} - \rho_\psi)^2 (1 - \rho_\psi^2)}
\]  

(25)

Since \( \mathrm{cov}(\tilde{b}_{t-1}, \tilde{\psi}_t) > 0 \), the simultaneity bias \( \mathrm{cov}(\tilde{b}_{t-1}, \tilde{\psi}_t)/\mathrm{var}(\tilde{b}_{t-1}) \) is always positive, which implies that \( \mathrm{plim} \hat{\gamma}_{OLS} > \gamma \) and OLS estimator of \( \gamma \) is inconsistent. In small samples, \( \hat{\gamma}_{OLS} \) can be significantly larger than \( \gamma \) so that the econometrician may mistakenly identify the underlying DGP as Ricardian.

### 3.3 Case III (Non-Ricardian):

\( \gamma = \rho_\theta = 0, \alpha \neq 0, \rho_\psi \neq 0 \)

In this case, we keep the assumptions of case II except that monetary authority starts to respond to inflation when setting the nominal interest rate. In general form, condition (15) is

\[
a_{21}\hat{\pi}_t + \hat{b}_t + a_{23}\hat{\theta}_t + a_{24}\hat{\psi}_t = 0
\]  

(26)

where \( a_{ij} \) is the \( ij \)th entry of \( P^{-1} \). Derivation of \( \mathrm{cov}(\tilde{b}_{t-1}, \tilde{\psi}_t) \) and \( \mathrm{var}(\tilde{b}_{t-1}) \) is shown in Appendix C.

From (64), it is hard to determine the sign of the simultaneity bias. As a solution, we evaluate the bias numerically over a grid in the space of \( (\alpha, \rho_\psi) \), where both parameters range from 0.01 to 0.99 with increment 0.01. As assumed, \( \gamma \) and \( \rho_\theta \) are both zeros. For the other parameters, we normalize the economy by setting \( y = 1 \) and calibrate \( \beta \) to be 0.99. Then we calibrate \( \delta = 0.001 \) to match steady state normalized real money balance in the U.S. data, which is 0.1. \( \gamma_0 \) is implicitly calibrated to match steady state normalized real government debt in the U.S. data, which is 0.8. At each grid point, \( \sigma_\theta \) and \( \sigma_\psi \) are calibrated to match standard deviations of \( \hat{\pi}_t^o \) and \( \hat{b}_t^o \) in the U.S. data, which are 0.007 and 0.3, respectively. Here, \( \hat{\pi}_t^o \) and \( \hat{b}_t^o \) are observed percentage deviations of inflation and real government debt from their corresponding steady state values.

Figure 2 is the plot of the simultaneity bias over the grid. Conditional on the calibrated parameter values, the bias can be either positive or negative, depending on
\(\alpha\) and \(\rho_\psi\). The bias function is continuous at zero, which means that \(\rho_\psi = 0\) is only sufficient but not necessary for zero bias in the non-Ricardian case. Obviously, with highly persistent fiscal policy shocks \((\rho_\psi > 0.8)\) and relatively passive monetary policy \((\alpha < 0.5)\), the bias can be very big. Sometimes, the bias even exceeds 1, the critical value between non-Ricardian and Ricardian regions. In such cases, as in case II, the econometrician may mistakenly identify the underlying DGP as Ricardian. Keeping \(\rho_\psi\) at high levels, increasing \(\alpha\) reverses the sign of the bias, which is also possible to disturb the econometrician’s inference.

### 3.4 Case IV (Ricardian)

In this case, monetary policy is active and fiscal policy is passive. Condition (15) reduces to

\[
\tilde{\pi}_t = -\frac{1}{\alpha - \rho_\theta} \tilde{\theta}_t \tag{27}
\]

From condition (16), we can solve for the unique \(\eta_t\)

\[
\eta_t = -\frac{1}{\alpha - \rho_\theta} \varepsilon_{\theta_t} \tag{28}
\]

Since the DGP is Ricardian, the second row of system (62) is a stable first-order difference equation, from which we can solve for \(\tilde{b}_t\) as a function of \(\varepsilon_{\theta_t}\) and \(\varepsilon_{\psi_t}\). Eventually, Appendix C shows that

\[
\text{cov}(\tilde{b}_{t-1}, \tilde{\psi}_t) = -\frac{\rho_\psi \sigma_{\tilde{\psi}}^2}{[1 - (\beta^{-1} - \gamma(\beta^{-1} - 1))\rho_\psi](1 - \rho_\psi^2)} \tag{29}
\]

which is always negative unless \(\rho_\psi = 0\). This implies that serially correlated fiscal policy shock is sufficient and necessary for negative simultaneity bias in the case of Ricardian equilibrium. Hence, \(\text{plim} \hat{\gamma}_{OLS} < \gamma\) and OLS estimator of \(\gamma\) is inconsistent. In small samples, \(\hat{\gamma}_{OLS}\) can be significantly smaller than \(\gamma\) so that the econometrician may mistakenly identify the underlying DGP as non-Ricardian, which is symmetric to case II.
3.5 Some Economics

By iterating (5) forward over $B/P$ and taking expectation conditional on information at period $t - 1$, we get the expected present-value government budget constraint:

$$
\frac{B_{t-1}}{P_{t-1}} = E_{t-1} \sum_{i=0}^{\infty} \left( \prod_{j=0}^{i} \pi_{t+j} R_{t+j-1}^{-1} \right) \left[ \tau_{t+i} + \frac{M_{t+i} - M_{t+i-1}}{P_{t+i}} \right]
$$

(30)

where the transversality condition for government debt has been imposed. The intertemporal government budget constraint (30) is a ubiquitous equilibrium condition, holding in both non-Ricardian and Ricardian equilibria. This condition tells us that, in equilibrium, real government debt is always equal to the sum of expected present value of future surpluses and seigniorage revenues.

Condition (30) helps to understand the economic intuition underlying the results of previous cases. For illustration, let us suppose in period $t - 1$, there is a surprise tax cut, i.e. $\varepsilon_{\psi_{t-1}} < 0$.

In case I, since $\rho_{\theta} = 0$, $\tilde{b}_{t-1}$ only responds to $\varepsilon_{\theta_{t-1}}$ as shown in (21). Because $\varepsilon_{\psi}$ and $\varepsilon_{\theta}$ are uncorrelated between each other for all leads and lags, $\varepsilon_{\psi_{t}}$ and $\varepsilon_{\theta_{t-1}}$ are uncorrelated. Also because $\rho_{\psi} = 0$, $\tilde{\psi}_{t} = \varepsilon_{\psi_{t}}$. Therefore, $\tilde{b}_{t-1}$ and $\tilde{\psi}_{t}$ are uncorrelated, which is consistent with (22).

In case II, since $\rho_{\psi} \neq 0$, a tax cut at period $t - 1$ projects lower path of future taxes $\tau_{t+i}$, for $i \geq 0$. Therefore, expected future surpluses on the right-hand side (RHS) of (30) get lower. Besides that, lower expected path of future taxes makes people feel wealthier, which immediately raises demand for goods. In endowment economy, higher demand leads to higher inflation. However, since $\alpha = 0$, nominal interest rate is pegged and does not respond to higher inflation. So expected future inflation and seigniorage revenues are unaffected. Consequently, the RHS of (30) gets lower, which makes $B_{t-1}/P_{t-1}$ and $\tilde{b}_{t-1}$ lower. Because $\tilde{\psi}_{t}$ is on a lower path, it tends to be lower than before. Consequently, $\tilde{b}_{t-1}$ and $\tilde{\psi}_{t}$ are positively correlated, which is consistent with the positive covariance in (24).

In case III, a tax cut at period $t - 1$ lowers the expected future surpluses on the RHS of (30), which is the same as case II. Also, lower expected path of future taxes leads to higher inflation at $t - 1$ through wealth effect. However, since $\alpha \neq 0$, generally
\( \alpha > 0 \), higher inflation leads to higher nominal interest rate through the monetary policy rule. According to the Fisher equation (51), higher interest rate raises expected inflation, which in turn raises expected seigniorage revenues. So the net effect of a tax cut on the RHS of (30) depends on the tradeoff between the effect of lower expected future surpluses and the effect of higher expected seigniorage revenues. Therefore, the correlation between the tax cut and \( \tilde{b}_{t-1} \) is ambiguous. Loosely speaking, given \( \alpha \), higher \( \rho \psi \) implies stronger expected surplus effect; given \( \rho \psi \), higher \( \alpha \) implies stronger expected seigniorage effect. Consequently, the correlation between \( \tilde{b}_{t-1} \) and \( \tilde{\psi}_t \) can be either positive or negative, which can be seen in Figure 2.

The economics underlying case IV is straightforward. From the solution of \( \tilde{b}_{t-1} \) that is based on the stable difference equation in (62), it can be shown that a tax cut at \( t-1 \) raises \( \tilde{b}_{t-1} \) immediately. According to the tax rule, higher \( \tilde{b}_{t-1} \) raises lump-sum tax at \( t \), which neutralizes the effect of a tax cut to the point that inflation at \( t-1 \) keeps unchanged. This is the standard Ricardian equivalence. On the other hand, a tax cut at \( t-1 \) reduces \( \tilde{\psi}_{t-1} \) and tends to reduce \( \tilde{\psi}_t \), provided \( \rho \psi > 0 \). So \( \tilde{b}_{t-1} \) and \( \tilde{\psi}_t \) tend to be negatively correlated, which is consistent with the negative covariance (29).

From these cases, it is clear that simultaneity bias associated with OLS estimator of the tax rule is ubiquitous no matter what the underlying DGP is. This could make correct identification of tax policy behavior very difficult. With non-Ricardian DGP, on one hand, taxes respond to lagged government debt weakly. As \( \rho \psi \) gets smaller, taxes becomes more close to exogenous and fiscal policy becomes more active. On the other hand, equilibrium condition (30) always holds, linking \( b_{t-1} \) to future taxes. When the econometrician runs OLS regression on the tax rule using equilibrium data, it will be hard to recover the true tax policy behavior because \( \hat{\gamma}_{OLS} \) captures the correlation between \( b_{t-1} \) and \( \tau_t \) that exists because of (30). It is possible that \( \gamma \) is very small while \( \hat{\gamma}_{OLS} \) is significantly large. With Ricardian DGP, both tax policy rule and (30) imply relatively strong correlation between \( b_{t-1} \) and \( \tau_t \). So \( \hat{\gamma}_{OLS} \) actually mixes the correlation information in both the tax policy rule and (30), which obscures the true fiscal policy behavior.
3.6 Monte Carlo Experiments

In order to illustrate the negative effect of simultaneity bias associated with $\hat{\gamma}_{OLS}$ in small samples, we implement the following Monte Carlo experiments.

Firstly, we assume the DGP to be non-Ricardian. In the space of $(\alpha, \gamma)$, the parameters governing monetary and fiscal policies, we set up a grid where both parameters range from 0.01 to 0.99 with increment 0.01. Then we assume the DGP to be Ricardian and the corresponding grid ranges from 1.01 to 1.99 with increment 0.01 for both parameters. For both cases, we calibrate $(y, \beta, \delta, \gamma_0)$ in the same way as case III in the previous section. We assume $\rho_\theta$ and $\rho_\psi$ to be 0.75 and 0.9, respectively. At each grid point, we calibrate $\sigma_\theta$ and $\sigma_\psi$ to match standard deviations of $\hat{\pi}_t$ and $\hat{b}_t$ in the U.S. data, which are 0.007 and 0.3, respectively. For each combination of $\alpha$ and $\gamma$, we simulate 1000 independent data sets based on (17) and (18), each with 300 quarters. Based on each data set, the econometrician runs OLS regression on (19) and does the following hypothesis testings:

Test 1 (Non-Ricardian): $H_0 : \gamma \leq 1$ $H_1 : \gamma > 1$

Test 2 (Non-Ricardian): $H_0 : \gamma \geq 0$ $H_1 : \gamma < 0$

Test 3 (Ricardian): $H_0 : \gamma \geq 1$ $H_1 : \gamma < 1$

When the DGP is non-Ricardian, $\gamma \in (0, 1)$ and the econometrician does Test 1 and 2, which contain true null hypotheses. For Ricardian DGP, $\gamma \geq 1$ and the econometrician does Test 3, which contains a true null hypothesis. We set nominal size of all the tests to be 5%. If the econometrician rejects the corresponding null hypothesis, we record that as a wrong decision. Finally, we calculate the empirical probabilities of making wrong decisions among the 1000 trials at each grid point.

Figure 3 illustrates the simultaneity bias for the non-Ricardian case. From the left panel, we find that the bias is positive when $\alpha$ is smaller than 0.3 and turns to negative as $\alpha$ gets larger. Besides, the negative bias is nonlinear in $\alpha$, which is indicated by a U-shaped plane. To get better understanding of this pattern, we plot the simultaneity bias against $\alpha$ in the right panel, where the bias has been taken average over $\gamma$. When $\alpha$ is small, the expected future surplus effect dominates the expected future seigniorage.
effect, which implies positive correlation between $\tilde{b}_{t-1}$ and $\tilde{\psi}_t$ and positive bias. And because monetary policy is very passive, nominal interest rate is nearly pegged. So the volatility of expected inflation and seigniorage is relatively small compared to the case when $\alpha$ is large. From (30), we know the volatility of $\tilde{b}_{t-1}$ must be relatively small, which makes the simultaneity bias relatively large. The simulation result attains a maximum bias 0.68 when $\alpha = 0.01$ and $\gamma = 0.7$. As $\alpha$ gets larger, the seigniorage effect starts to dominate, which turns the correlation between $\tilde{b}_{t-1}$ and $\tilde{\psi}_t$ to negative. However, the seigniorage effect is not linear in $\alpha$. There is actually a Laffer curve underlying the nonlinearity. As the economy gets to the downward-sloping side of the Laffer curve, higher $\alpha$ induces weaker seigniorage effect, which means that the surplus effect dominates again at some point. So the bias has an overturn and approaches back to zero.

In Figure 4, the two panels on the left show us the empirical probabilities of making wrong decisions when the econometrician does Test 1 and 2. The two panels on the right are averaged probabilities according to $\alpha$ after taking average over $\gamma$. From the upper-right panel, it is clear that when $\alpha$ is smaller than 0.3, the empirical probabilities of making wrong decisions are generally larger than 0.05, the nominal size of the test. This means that the econometrician is very likely to make mistakes in these cases by identifying the non-Ricardian DGP as Ricardian. As shown in the lower-right panel, under the highly negative bias associated with high levels of $\alpha$, empirical probabilities of making wrong decisions are larger than 0.05 over a very wide range of $\alpha$, which indicates high propensity of making mistakes as the econometrician does Test 2.

Figure 5 illustrates the simultaneity bias for the Ricardian case. Being consistent with the theory, the bias is always negative and turns out to be a downward-sloping plane in $\gamma$. This is because as $\gamma$ gets larger, taxes respond to government debt more strongly. This leaves less room for expected seigniorage to vary because the wealth effect is weaker. Through (30), $\text{var}(\tilde{b}_{t-1})$ gets smaller, which makes the bias become larger. The right panel of Figure 5 is the bias in $\gamma$ after taking average over $\alpha$.

The left panel of Figure 6 shows us the empirical probabilities of making wrong decisions when the econometrician does Test 3. The right panel is the probabilities in $\gamma$ after taking average over $\alpha$. Even though the bias is decreasing in $\gamma$, in small
samples the probability of making wrong decisions is relatively high when $\gamma$ is relatively small, because $\gamma$ is close to the boundary value. However, the probability of making mistakes is not so high when both $\gamma$ and the bias become larger. This is because in small samples, the bias is much smaller than its limit in magnitude and is not negative enough to push $\hat{\gamma}_{OLS}$ down below 1 significantly.

As a summary, OLS estimation on the tax policy rule is very distorting in small samples. We need a more robust estimation scheme which is free of simultaneity bias and is able to draw reliable inference on fiscal policy behavior.

4 Statistical Inference II: Vector Autoregression

As mentioned in the previous section, $\hat{\gamma}_{OLS}$ based on equilibrium data captures correlation information of both the tax policy rule (19) and the ubiquitous equilibrium condition (30), which causes simultaneity bias in $\hat{\gamma}_{OLS}$. To solve the problem, we propose to estimate tax policy rule in the VAR framework together with information implied by the equilibrium condition (30). The restricted VAR may avoid simultaneity bias and help to recover the true fiscal policy behavior.

In this section, we firstly log-linearize condition (30) and derive the restrictions to be imposed on the unrestricted VAR. Then we briefly sketch the estimation procedure of the restricted VAR. The technique closely follows Chung and Leeper (2007) and Traum (2007). Finally, we report Monte Carlo experiment results.

4.1 Restrictions on the VAR

Firstly, we express (30) in real terms as

$$b_{t-1} = E_{t-1} \sum_{i=0}^{\infty} \left( \prod_{j=0}^{i} \pi_{t+j} R_{t+j-1}^{-1} \right) \left[ \tau_{t+i} + m_{t+i} - \frac{1}{\pi_{t+i}} m_{t+i-1} \right]$$  (31)
Log linearizing (31) around the deterministic steady state and expressing variables in logarithm, we get\(^6\)

\[
\tilde{b}_{t-1} = E_{t-1} \sum_{i=0}^{\infty} \delta^i \left[ \left( \tau + m - \frac{m}{\pi} \right) \left( \frac{\delta}{1 - \delta} \right) \left( \tilde{\pi}_{t+i} - \tilde{R}_{t+i-1} \right) \right] + E_{t-1} \sum_{i=0}^{\infty} \delta^{i+1} \left[ \tau \tilde{\pi}_{t+i} + R \left( \frac{\delta c - m}{R - 1} \right) \tilde{R}_{t+i} - \frac{R}{\pi} \left( \frac{\delta c - m}{R - 1} \right) \tilde{R}_{t+i-1} + \frac{m}{\pi} \tilde{\pi}_{t+i} \right] + k
\]

where \( \delta \equiv \pi / R \) is the steady state discount factor and

\[
k = b \ln(b) - \frac{\delta}{(1 - \delta)^2} \left( \tau + m - \frac{m}{\pi} \right) \ln(\delta) + \frac{\delta}{1 - \delta} \left[ \frac{R}{\pi} \left( \frac{\delta c - m}{R - 1} \right) (1 - \pi) \ln(R) - \tau \ln(\tau) - \frac{m}{\pi} \ln(\pi) \right]
\]

Following Chung and Leeper (2007) and Traum (2007), we suppose the state of the economy is characterized by the \( n \)-dimensional factors \( f_t \) which evolve according to the unrestricted VAR(p) process

\[
f_t = B_0 + f_{t-1}B_1 + \cdots + f_{t-p}B_p + u_t
\]

In our application, \( n = 4 \) and \( f_t = [\tilde{\pi}_t, \tilde{b}_t, \tilde{\tau}_t, \tilde{R}_t] \) is a \((1 \times 4)\) vector, \( B_0 \) and \( u_t \) are \((1 \times n)\) vectors, \( B_i \) is an \((n \times n)\) matrix for \( i = 1, \ldots, p \). We also suppose all model variables \( \tilde{x}_t \) are related to the factors \( f_t \) through the mapping

\[
\tilde{x}_t = f_tC_x
\]

where the selection vector \( C_x \) is simply defined as \( C_\pi = [1, 0, 0, 0]' \), \( C_b = [0, 1, 0, 0]' \), \( C_\tau = [0, 0, 1, 0]' \) and \( C_R = [0, 0, 0, 1]' \). Both \( f_t \) and \( \tilde{x}_t \) are measured in logarithm.

To derive the restrictions on the VAR, it is convenient to fit (34) in companion form and express it as a VAR(1) process

\[
\tilde{f}_t = \tilde{B}_0 + \tilde{f}_{t-1}B + \tilde{u}_t
\]

\(^6\)See Appendix D
where \( \bar{f}_t = [f_t, f_{t-1}, \ldots, f_{t-p+1}] \), \( B_0 = [B_0, 0_{1 \times n(p-1)}] \) and \( \bar{u}_t = [u_t, 0_{1 \times n(p-1)}] \) are all \((1 \times np)\) vectors and \( B \) is an \((np \times np)\) matrix which is defined as

\[
B = \begin{bmatrix}
B_1 & I_n & 0_n & \cdots & 0_n \\
B_2 & 0_n & I_n & \cdots & 0_n \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
B_{p-1} & 0_n & 0_n & \cdots & I_n \\
B_p & 0_n & 0_n & \cdots & 0_n
\end{bmatrix}
\]

where \( I_n \) is an \( n \)-dimensional identity matrix and \( 0_n \) is an \((n \times n)\) zero square matrix.

Accordingly, the mapping (35) becomes

\[
\tilde{x}_t = \bar{f}_t \tilde{C}_x
\]

where all \( \tilde{C}_x \) are now \((np \times 1)\) vectors, such as \( \tilde{C}_\pi = [C'_\pi, 0'_{n(p-1) \times 1}]' \).

As shown in Appendix E, condition (32) implies the following restrictions on the VAR coefficients \( B_0 \) and \( B \), which will be imposed in the estimation.

\[
\delta B \left\{ b \bar{C}_b + \left[ (\tau - m) \left( \frac{1}{1 - \delta} \right) + \frac{m}{\pi} \right] \tilde{C}_\pi + \tau \bar{C}_\tau + R \left( \frac{\delta c - m}{R - 1} \right) \tilde{C}_R \right\} = b \bar{C}_b + \left[ (\tau - m) \left( \frac{\delta}{1 - \delta} \right) + \frac{\delta m}{\pi} \left( \frac{\delta c - m}{R - 1} \right) \right] \tilde{C}_R
\]

\[
\delta B \left\{ \left[ (\tau - m) \left( \frac{1}{1 - \delta} \right) + \frac{m}{\pi} \right] \left( \frac{1}{1 - \delta} \right) B_0 (I - \delta B)^{-1} \tilde{C}_\pi + \left( \frac{\delta \tau}{1 - \delta} \right) B_0 (I - \delta B)^{-1} \tilde{C}_\tau + \left[ \frac{\delta R}{1 - \delta} \left( \frac{\delta c - m}{R - 1} \right) \left( 1 - \frac{1}{R} \right) - (\tau - m) \left( \frac{\delta}{1 - \delta} \right)^2 \right] B_0 (I - \delta B)^{-1} \tilde{C}_R \right\} = -k
\]

Since the elements in \( B_0 \) are irrelevant to our analysis, we will only impose restriction (38) in the estimation, which is consistent with Chung and Leeper (2007) and Traum (2007).

### 4.2 Estimation Procedure

Suppose we want to estimate the reduced-form VAR(p) system (34)

\[
f_t = B_0 + f_{t-1}B_1 + \cdots + f_{t-p}B_p + u_t
\]
or more compactly
\[ f_t = B_0 + f_- B_- + u_t \] (40)
where \( f_- = [f_{t-1}, \ldots, f_{t-p}] \) is a \((1 \times np)\) lagged data vector and \( B_- = [B'_1, \ldots, B'_p]' \) is an \((np \times n)\) matrix. Then we stack the \( T + p \) observations in the form of (40) and get
\[ F = 1B_0 + F_- B_- + u \] (41)
where \( F \) is a \((T \times n)\) data matrix, \( 1 \) is a \((T \times 1)\) vector consisting of all 1’s, \( F_- \) is a \((T \times np)\) lagged data matrix and \( u \) is a \((T \times n)\) matrix. Define \( X \equiv [1, F_-] \), a \((T \times (np + 1))\) matrix, \( \bar{C}_0 \), an \((np \times 1)\) vector, and \( b \equiv vec([B'_0, B'_-]) \), an \((n(np + 1) \times 1)\) vector. Applying vec operator, we can express (41) as
\[ vec(F) = (I_n \otimes X)b + vec(u) \] (42)
The log likelihood function corresponding to (42) is
\[ L(b, \Sigma) = -\frac{nT}{2} \ln(2\pi) + \frac{1}{2} \ln |(\Sigma \otimes I_T)^{-1}| - \frac{1}{2} \left[ vec(u)'(\Sigma \otimes I_T)^{-1}vec(u) \right] \] (43)
where \( \Sigma \) is the across-equation variance and covariance matrix.

Restriction (38) to be imposed on the unrestricted VAR can be simplified as
\[ B_- C_1 = \bar{C}_0 \] (44)
where \( C_1 \), an \((n \times 1)\) vector, and \( \bar{C}_0 \), an \((np \times 1)\) vector, are defined as
\[ C_1 \equiv \delta \left\{ bC_b + \left( \tau + m - \frac{m}{\pi} \right) \left( \frac{1}{1-\delta} \right) + \frac{m}{\pi} \right\} C_\pi + \tau C_\tau + R \left( \frac{\delta c - m}{R - 1} \right) C_R \] (45)
\[ \bar{C}_0 \equiv b\bar{C}_b + \left[ \left( \tau + m - \frac{m}{\pi} \right) \left( \frac{\delta}{1-\delta} \right) + \frac{\delta R}{\pi} \left( \frac{\delta c - m}{R - 1} \right) \right] \bar{C}_R \] (46)
(44) can be rewritten as
\[ [0_{np \times 1}, I_{np}][B'_0, B'_-]'C_1 = \bar{C}_0 \] (47)
Applying vec operator on (47), we get
\[ Vb = vec(\bar{C}_0) \] (48)
where \( V \equiv (C_1' \otimes [0_{np \times 1}, I_{np}]) \).

The goal of the estimation is to choose \( b \) to maximize (43) subject to constraint (48). Appendix F shows that \( \tilde{b} \), the consistent estimator of \( b \) of the restricted VAR, is given by

\[
\tilde{b} = \hat{b} + S^{-1}V'(VS^{-1}V')^{-1}(\text{vec}(\tilde{C}_0) - V\hat{b})
\]

(49)

where \( \hat{b} = \text{vec}((X'X)^{-1}(X'F)) \) is the unrestricted VAR estimator of \( b \) and \( S \equiv (\Sigma^{-1} \otimes X'X) \). Following Chung and Leeper (2007) and Traum (2007), we use bootstrap method to estimate standard errors of \( \tilde{b} \) and a feasible GLS procedure based on (49), iterated until convergence, to get a consistent estimator of \( \Sigma \).

### 4.3 Monte Carlo Experiments

In the Monte Carlo experiments, we suppose the econometrician wants to estimate VAR to identify fiscal policy behavior. Again, we treat the DSGE model as DGP and the simulation is based on (17) and a modified version of (18). Since the VAR contains four variables and there are only two exogenous shocks in the DSGE model, we introduce two measurement errors to avoid stochastic singularity. The modified observation equation is

\[
\begin{bmatrix}
\hat{\pi}_{t+1} \\
\hat{b}_{t+1} \\
\end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
\end{bmatrix} \begin{bmatrix}
\hat{\pi}_{t+1} \\
\hat{b}_{t+1} \\
\hat{\theta}_{t+1} \\
\hat{\psi}_{t+1} \\
\end{bmatrix} + \begin{bmatrix}
\zeta_{1t+1} \\
\zeta_{2t+1} \\
\end{bmatrix}
\]

(50)

where \( \zeta_{1t+1} \) and \( \zeta_{2t+1} \) are measurement errors that are normally distributed with means 0 and standard deviations \( \sigma_{\zeta_1} \) and \( \sigma_{\zeta_2} \), respectively.

For the non-Ricardian DGP, we set up a grid where \( \alpha \) and \( \gamma \) range from 0.05 to 0.95 with increment 0.05. For the Ricardian DGP, both parameters range from 1.05 to 1.95 with increment 0.05. The other parameters are calibrated in the same way as the previous section. At each grid point, we simulate 1000 independent data sets based on (17) and (50), each with 300 quarters.

In the benchmark experiment, the econometrician firstly runs OLS regression on (19) and then estimates an unrestricted VAR (34) based on each data set. The optimal VAR lag length \( p \) is determined by Akaike Information Criteria (AIC). The
econometrician then does hypothesis testings, which are specified in the previous section, corresponding to the underlying DGP. The nominal size of all the tests are kept at 5%. Finally, we calculate the empirical probabilities of making wrong decisions among the 1000 trials at each grid point for each estimation scheme.

In order to evaluate the relative performance of the restricted VAR, we implement a second experiment, where we still simulate 1000 data sets per grid point. This time, the econometrician firstly runs OLS regression on (19) and then estimates a restricted VAR based on (49) for each data set. The determination of optimal VAR lag length and hypothesis testings are carried out in the same way as in the benchmark experiment. The econometrician applies bootstrap method to estimate standard errors of $\tilde{b}$, which are used to calculate relevant test statistics. Finally, we calculate the empirical probabilities of making wrong decisions among the 1000 trials at each grid point for each estimation scheme.

Figure 7 and 8 show the empirical probabilities of making wrong decisions for the benchmark experiment, where Figure 7 is for non-Ricardian DGP and Figure 8 is for Ricardian DGP. In each figure, panels on the left are based on OLS regression and those on the right are based on unrestricted VAR. For the non-Ricardian DGP, as shown in Figure 7, unrestricted VAR is more robust than OLS regression in identifying fiscal policy behavior. Even though there are regions where OLS has empirical probabilities close to zero, there are also large regions where OLS is unreliable at all. On the other hand, the unrestricted VAR has more stable performance, with average empirical probability around 0.05, the nominal size of the test. For the Ricardian DGP, as shown in Figure 8, unrestricted VAR has no advantage over OLS, where in some regions unrestricted VAR is even inferior to OLS.

To compare OLS with unrestricted VAR more clearly, Table 1 to 3 report the empirical probabilities of making wrong decisions for both DGPs and both estimation schemes. In each table, rows are for $\gamma$ and columns are for $\alpha$. The empirical probability before “/” is for OLS and that after “/” is for unrestricted VAR. The last row and last column are averages taken over $\gamma$ and $\alpha$, respectively. The entry in the lower-right corner is total average of the whole table. Specifically, Table 1 to 3 are corresponding to the upper two panels of Figure 7, the lower two panels of Figure 7 and Figure 8,
respectively.

From the benchmark experiment, it seems that unrestricted VAR is able to correct simultaneity bias associated with OLS estimator if the DGP is non-Ricardian. For Ricardian DGP, unrestricted VAR is not superior to OLS.

5 Conclusion

How can we identify U.S. fiscal policy behavior in a reliable way? We start by questioning on the validity of OLS regression on the tax rule, one of the common approaches to identify U.S. fiscal policy behavior in the current literature. We cast doubt on OLS method because in general equilibrium framework, OLS regression isolates the tax rule from a system of equations implied by the whole model, which causes simultaneity bias in the OLS estimator. In a very simple DSGE model with monetary and fiscal policy interactions, we analytically show that the simultaneity bias is ubiquitous for both non-Ricardian and Ricardian equilibria. In small samples, OLS-based identification of fiscal policy behavior is unreliable due to the simultaneity bias.

As a solution, unrestricted VAR corrects simultaneity bias and provides reliable inference on fiscal policy behavior for non-Ricardian DGP, but not for Ricardian DGP. VAR estimation subject to restrictions implied by the intertemporal government budget constraint may be a solution to this puzzle. Another direction for future research is to apply the technique of this paper to real U.S. data and compare the results with Bohn (1998).

Even though the analysis of this paper is based on the particular model specification and calibrated parameter values, the simultaneity bias problem associated with OLS estimation always exists in general equilibrium framework, which calls for extra caution of the empirical macroeconomists.
References


Appendix A: Solution of the Model

To solve the model, we firstly log-linearize (7) and (9) as

\[
\hat{R}_t = E_t \hat{\pi}_{t+1} \tag{51}
\]
\[
\hat{R}_t = \alpha \hat{\pi}_t + \hat{\theta}_t \tag{52}
\]

To get (51), we have imposed steady state condition \( R = \beta^{-1} \). For simplicity, we assume \( \pi = 1 \). Combining (51) and (52), we get

\[
E_t \hat{\pi}_{t+1} = \alpha \hat{\pi}_t + \hat{\theta}_t \tag{53}
\]

We then define the one-period-ahead endogenous forecast error \( \eta_{t+1} \equiv \hat{\pi}_{t+1} - E_t \hat{\pi}_{t+1} \) and express (53) as

\[
\hat{\pi}_{t+1} = \alpha \hat{\pi}_t + \hat{\theta}_t + \eta_{t+1} \tag{54}
\]

Then we log-linearize (5), (8) and (10) as

\[
m \hat{m}_t + b \hat{b}_t + \tau \hat{\pi}_t = \hat{m}_{t-1} - \hat{m}_{t+1} + Rb \hat{R}_{t-1} + Rb \hat{b}_{t-1} - Rb \hat{\pi}_t \tag{55}
\]
\[
m(R - 1) \hat{m}_t = R(\delta c - m) \hat{R}_t \tag{56}
\]
\[
\hat{\pi}_t = \gamma \hat{b}_{t-1} + \hat{\psi}_t \tag{57}
\]

where relevant steady state conditions have been imposed. Combining (52), (55)-(57) and rearranging terms, we get

\[
\varphi_1 \hat{\pi}_t + \hat{b}_t + \varphi_2 \hat{\pi}_{t-1} - \left[ \beta^{-1} - \gamma (\beta^{-1} - 1) \right] \hat{b}_{t-1} + \varphi_3 \hat{\theta}_t + (\beta^{-1} - 1) \hat{\psi}_t + \varphi_4 \hat{\theta}_{t-1} = 0 \tag{58}
\]

where

\[
\varphi_1 \equiv \frac{m}{b}(\alpha \chi + 1) + \beta^{-1}
\]
\[
\varphi_2 \equiv -\alpha \left[ \frac{m}{b} \chi + \beta^{-1} \right]
\]
\[
\varphi_3 \equiv \frac{m}{b} \chi
\]
\[
\varphi_4 \equiv - \left[ \frac{m}{b} \chi + \beta^{-1} \right]
\]
and \( \chi \equiv 1/(1 - R) \) is the interest elasticity of money demand.

It is straightforward to log-linearize (11) and (12) as

\[
\hat{\theta}_t = \rho_\theta \hat{\theta}_{t-1} + \varepsilon_{\theta_t} \quad (59) \\
\hat{\psi}_t = \rho_\psi \hat{\psi}_{t-1} + \varepsilon_{\psi_t} \quad (60)
\]

Then substituting (59) and (60) into (58), we get

\[
\varphi_1 \hat{\pi}_{t+1} + \hat{b}_{t+1} = -\varphi_2 \hat{\pi}_t + [\beta^{-1} - \gamma(\beta^{-1} - 1)] \hat{b}_t - (\varphi_3 \rho_\theta + \varphi_4) \hat{\theta}_t - (\beta^{-1} - 1) \rho_\psi \hat{\psi}_t - \varphi_3 \varepsilon_{\theta_{t+1}} - (\beta^{-1} - 1) \varepsilon_{\psi_{t+1}}
\]

(61)

So far, (54), (61), (59) and (60) form a self-contained system governing the dynamics of \( \hat{\pi}_t, \hat{b}_t, \hat{\theta}_t \) and \( \hat{\psi}_t \). We organize the linearized system in the following matrix form

\[
\Gamma_0 Y_{t+1} = \Gamma_1 Y_t + \Pi \eta_{t+1} + \Psi \varepsilon_{t+1}
\]

where \( Y_{t+1} = [\hat{\pi}_{t+1}, \hat{b}_{t+1}, \hat{\theta}_{t+1}, \hat{\psi}_{t+1}]' \), \( \varepsilon_{t+1} = [\varepsilon_{\theta_{t+1}}, \varepsilon_{\psi_{t+1}}]' \), \( \Pi = [\begin{array}{cc} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{array}] \), \( \Psi = [\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array}] \)

\[
\Gamma_0 = \begin{bmatrix}
1 & 0 & 0 & 0 \\
\varphi_1 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}, \quad \Pi = \begin{bmatrix}
1 \\
0 \\
0 \\
0
\end{bmatrix}, \quad \Psi = \begin{bmatrix}
-\varphi_3 & -1 \\
1 & 0 \\
0 & 1
\end{bmatrix}
\]

Since \( \Gamma_0 \) is invertible, (13) can be expressed as

\[
Y_{t+1} = \Gamma_1^* Y_t + \Pi^* \eta_{t+1} + \Psi^* \varepsilon_{t+1}
\]

where \( \Gamma_1^* = \Gamma_0^{-1} \Gamma_1 \), \( \Pi^* = \Gamma_0^{-1} \Pi \) and \( \Psi^* = \Gamma_0^{-1} \Psi \). Applying Jordan decomposition on \( \Gamma_1^* \), the above equation becomes:

\[
Y_{t+1} = P \Lambda P^{-1} Y_t + \Pi^* \eta_{t+1} + \Psi^* \varepsilon_{t+1} \\
\Rightarrow \quad P^{-1} Y_{t+1} = \Lambda P^{-1} Y_t + P^{-1} \Pi^* \eta_{t+1} + P^{-1} \Psi^* \varepsilon_{t+1}
\]

(62)
where \( \Lambda \) is diagonal matrix with eigenvalues of \( \Gamma_1^* \) on the main diagonal, which is in the following form

\[
\Lambda = \begin{bmatrix}
\alpha & 0 & 0 & 0 \\
0 & \beta^{-1} - \gamma(\beta^{-1} - 1) & 0 & 0 \\
0 & 0 & \rho_\theta & 0 \\
0 & 0 & 0 & \rho_{\psi}
\end{bmatrix}
\]

\( P \) is a matrix, each column of which is the eigenvector of \( \Gamma_1^* \) and is corresponding to the eigenvalue in \( \Lambda \). Since there are no repeated eigenvalues, \( P \) has full column rank and \( P^{-1} \) can be shown as

\[
P^{-1} = \begin{bmatrix}
\frac{\alpha \varphi_1 + \varphi_2}{\beta^{-1} - \gamma(\beta^{-1} - 1) - \alpha} & 0 & \frac{\alpha \varphi_1 + \varphi_2}{|\beta^{-1} - \gamma(\beta^{-1} - 1) - \alpha|} & 0 \\
-\frac{\alpha \varphi_1 + \varphi_2}{\beta^{-1} - \gamma(\beta^{-1} - 1) - \alpha} & 1 & \frac{P_{2,3}}{\beta^{-1} - \gamma(\beta^{-1} - 1) - \rho_{\psi}} & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}
\]

where \( P_{2,3} = -\frac{\varphi_1[\beta^{-1} - \gamma(\beta^{-1} - 1)] + \varphi_2}{|\beta^{-1} - \gamma(\beta^{-1} - 1) - \alpha|} - \frac{\varphi_3 \rho_\theta + \varphi_4}{\beta^{-1} - \gamma(\beta^{-1} - 1) - \rho_{\psi}} \). If the bounded equilibrium is determinate, the following conditions suppress the unstable root of the system and must hold for all \( t \):

\[
P^i Y_t = 0 \\
P^i \Pi^* \eta_{t+1} + P^i \Psi^* \epsilon_{t+1} = 0
\]

where \( P^i \) is the \( i \)th row of \( P^{-1} \) and the \( i \)th eigenvalue in \( \Lambda \) is unstable.

**Appendix B: Derivation of (20)**

For convenience, let us rewrite (19) as below.

\[
\tilde{\tau}_t = \gamma_0 + \gamma \tilde{b}_{t-1} + \tilde{\psi}_t
\]

Stacking all observations in vectors, we get

\[
\tilde{\tau} = \begin{bmatrix} 1 \gamma_0 + \gamma \tilde{b}_{t-1} + \tilde{\psi} \end{bmatrix} = \tilde{X} \Gamma + \tilde{\psi}
\]

(63)
where \( \hat{\tau} = [\hat{\tau}_2, \ldots, \hat{\tau}_T]' \), \( \hat{b}_{-1} = [\hat{b}_1, \ldots, \hat{b}_{T-1}]' \), \( \hat{\psi} = [\hat{\psi}_2, \ldots, \hat{\psi}_T]' \), \( \tilde{X} = [I, \hat{b}_{-1}] \), \( \Gamma = [\gamma_0, \gamma]' \) and \( I \) is a \(((T-1) \times 1)\) vector consisting of all 1's. The OLS estimator of \( \Gamma \) based on (63) is

\[
\hat{\gamma}_{OLS} = \left[ \begin{array}{c} \hat{\gamma}_{0,OLS} \\ \hat{\gamma}_{OLS} \end{array} \right] = (\tilde{X}'\tilde{X})^{-1}(\tilde{X}'\hat{\tau}) = \Gamma + (\tilde{X}'\tilde{X})^{-1}(\tilde{X}'\hat{\psi})
\]

\[
= \left[ \begin{array}{c} \gamma_0 \\ \gamma \end{array} \right] + \left[ \begin{array}{c} 1'1 \\ \tilde{b}_{-1}'1 \\ \tilde{b}_{-1}'\hat{b}_{-1} \end{array} \right]^{-1} \left[ \begin{array}{c} 1'\hat{\psi} \\ \tilde{\psi}'1 \end{array} \right]
\]

\[
= \left[ \begin{array}{c} \gamma_0 \\ \gamma \end{array} \right] + \frac{1}{(T-1)\sum_{t=2}^T \tilde{b}_{t-1}^2 - (\sum_{t=2}^T \tilde{b}_{t-1})^2} \left[ \sum_{t=2}^T \tilde{b}_{t-1}^2 \sum_{t=2}^T \tilde{\psi}_t - \sum_{t=2}^T \tilde{b}_{t-1} \sum_{t=2}^T (\tilde{b}_{t-1} \tilde{\psi}_t) \right]
\]

\[
\left( T - 1 \right) \sum_{t=2}^T (\tilde{b}_{t-1} \tilde{\psi}_t) - \sum_{t=2}^T \tilde{b}_{t-1} \sum_{t=2}^T \tilde{\psi}_t \]

Therefore, the probability limit of \( \hat{\gamma}_{OLS} \) is

\[
\text{plim} \hat{\gamma}_{OLS} = \gamma + \text{plim} \frac{(T-1)\sum_{t=2}^T (\tilde{b}_{t-1} \tilde{\psi}_t) - \sum_{t=2}^T \tilde{b}_{t-1} \sum_{t=2}^T \tilde{\psi}_t}{(T-1)\sum_{t=2}^T \tilde{b}_{t-1}^2 - (\sum_{t=2}^T \tilde{b}_{t-1})^2}
\]

\[
= \gamma + \frac{E(\tilde{b}_{t-1} \tilde{\psi}_t) - E(\tilde{b}_{t-1})E(\tilde{\psi}_t)}{E(\tilde{b}_{t-1}^2) - (E(\tilde{b}_{t-1}))^2} = \gamma + \frac{\text{cov}(\tilde{b}_{t-1}, \tilde{\psi}_t)}{\text{var}(\tilde{b}_{t-1})}
\]

where \( E(\cdot) \) is expectation operator.

**Appendix C: Derivation of Asymptotic Bias**

- **Case III (Non-Ricardian):** \( \gamma = \rho_\theta = 0, \alpha \neq 0, \rho_\psi \neq 0 \)

Since \( a_{21} \neq 0 \) in (26), we substitute (26) into (58) and get

\[
\left( \frac{\varphi_1}{a_{21}} - 1 \right) \hat{b}_t + \left( \frac{\varphi_2}{a_{21}} + \beta^{-1} - \gamma(\beta^{-1} - 1) \right) \hat{b}_{t-1} = \left( \frac{\varphi_3 - \varphi_1 a_{23}}{a_{21}} \right) \hat{\theta}_t + \left( \beta^{-1} - 1 - \frac{\varphi_1 a_{24}}{a_{21}} \right) \hat{\psi}_t + \left( \varphi_4 - \frac{\varphi_2 a_{23}}{a_{21}} \right) \hat{\theta}_{t-1} + \left( -\frac{\varphi_2 a_{24}}{a_{21}} \right) \hat{\psi}_{t-1}
\]

28
Therefore,
\[
\hat{b}_t = \ln(b) + \frac{A_3 + A_5}{(A_1 + A_2 L)(1 - \rho_\theta L)}\varepsilon_{\theta_t} + \frac{A_1 + A_6 L}{(A_1 + A_2 L)(1 - \rho_\psi L)}\varepsilon_{\psi_t} \\
= \ln(b) + \frac{\frac{A_3}{A_1} + \frac{A_5}{A_1} L}{(1 + \frac{A_3}{A_1} L)(1 - \rho_\theta L)}\varepsilon_{\theta_t} + \frac{\frac{A_1}{A_1} + \frac{A_6}{A_1} L}{(1 + \frac{A_3}{A_1} L)(1 - \rho_\psi L)}\varepsilon_{\psi_t} \\
= \ln(b) + \frac{C_3 + C_5 L}{B_1(1 + C_2 L)(1 - \rho_\theta L)}\varepsilon_{\theta_t} + \frac{C_4 + C_6 L}{B_2(1 + C_2 L)(1 - \rho_\psi L)}\varepsilon_{\psi_t},
\]

where \(C_i = A_i/A_1\), for \(i = 2, 3, 4, 5, 6\).

It can be shown that
\[
B_1 = \frac{m_1}{1 + C_2 L} - \frac{n_1}{1 - \rho_\theta L} \quad \text{and} \quad B_2 = \frac{m_2}{1 + C_2 L} - \frac{n_2}{1 - \rho_\psi L}
\]

where
\[
m_1 = \frac{C_2 C_3 - C_5}{\rho_\theta + C_2}, \quad m_2 = \frac{C_2 C_4 - C_6}{\rho_\theta + C_2}, \quad n_1 = -\frac{C_5 \rho_\theta + C_5}{\rho_\theta + C_2}, \quad n_2 = -\frac{C_4 \rho_\psi + C_6}{\rho_\psi + C_2}
\]

So we have
\[
\hat{b}_{t-1} = \ln(b) + B_1 \varepsilon_{\theta_{t-1}} + B_2 \varepsilon_{\psi_{t-1}} \\
= \ln(b) + (m_1 - n_1) \varepsilon_{\theta_{t-1}} - (m_1 C_2 + n_1 \rho_\theta) \varepsilon_{\theta_{t-2}} + (m_1 C_2^2 - n_1 \rho_\theta^2) \varepsilon_{\theta_{t-3}} - \cdots \\
+ (m_2 - n_2) \varepsilon_{\psi_{t-1}} - (m_2 C_2 + n_2 \rho_\psi) \varepsilon_{\psi_{t-2}} + (m_2 C_2^2 - n_2 \rho_\psi^2) \varepsilon_{\psi_{t-3}} - \cdots
\]

Since
\[
\tilde{\psi}_t = \frac{\varepsilon_{\psi_t}}{1 - \rho_\psi L} = \varepsilon_{\psi_t} + \rho_\psi \varepsilon_{\psi_{t-1}} + \rho_\psi^2 \varepsilon_{\psi_{t-2}} + \cdots
\]

We can show that
\[
cov(\hat{b}_{t-1}, \tilde{\psi}_t) = (m_2 - n_2) \rho_\psi \sigma_{\psi}^2 - (m_2 C_2 + n_2 \rho_\psi) \rho_\psi^2 \sigma_{\psi}^2 + (m_2 C_2^2 - n_2 \rho_\psi^2) \rho_\psi^3 \sigma_{\psi}^2 - \cdots \\
= [(m_2 \rho_\psi - m_2 C_2 \rho_\psi^2 + m_2 C_2^2 \rho_\psi^3 - \cdots) - (n_2 \rho_\psi + n_2 \rho_\psi^3 + n_2 \rho_\psi^5 + \cdots)] \sigma_{\psi}^2 \\
= \left(\frac{m_2 \rho_\psi}{1 + C_2 \rho_\psi} - \frac{n_2 \rho_\psi}{1 - \rho_\psi^2}\right) \sigma_{\psi}^2
\]
where the last equality holds because in non-Ricardian equilibrium, $C_2 = -\alpha \beta > -1$ and $|C_2 \rho_\psi| < 1$.

Finally, we can show that

\[
\text{var}(\tilde{b}_{t-1}) = \text{cov}(\tilde{b}_{t-1}, \tilde{b}_{t-1}) \\
= [(m_1 - n_1)^2 + (m_1 C_2 + n_1 \rho_\theta)^2 + (m_1 C^2_2 - n_1 \rho_\theta^2)^2 + \cdots] \sigma_\theta^2 \\
+ [(m_2 - n_2)^2 + (m_2 C_2 + n_2 \rho_\psi)^2 + (m_2 C^2_2 - n_2 \rho_\psi^2)^2 + \cdots] \sigma_\psi^2 \\
= \frac{m_1^2 \sigma_\theta^2 + m_2^2 \sigma_\psi^2}{1 - C_2^2} + \frac{n_1^2 \sigma_\theta^2 + n_2^2 \sigma_\psi^2}{1 - \rho_\theta^2 + 1 - \rho_\psi^2} - \frac{2m_1 n_1 \sigma_\theta^2}{1 + C_2 \rho_\theta} - \frac{2m_2 n_2 \sigma_\psi^2}{1 + C_2 \rho_\psi}
\]

(65)

- Case IV (Ricardian):

Recall (27) and (28),

\[
\tilde{\pi}_t = -\frac{1}{\alpha - \rho_\theta} \tilde{\theta}_t \\
\eta_t = -\frac{1}{\alpha - \rho_\theta} \varepsilon_{\theta_t}
\]

Let us define $g \equiv -1/(\alpha - \rho_\theta)$ and $\lambda = \beta^{-1} - \gamma(\beta^{-1} - 1)$. Since the equilibrium is Ricardian, the second row of system (62) is a stable first-order difference equation, from which we can solve for $\tilde{b}_t$ as a function of $\varepsilon_\theta$ and $\varepsilon_\psi$:

\[
\tilde{b}_t = \ln(b) + \frac{D_3 + D_5 L}{(D_1 + D_2 L)(1 - \rho_\theta L)} \varepsilon_{\theta_t} + \frac{D_4 + D_6 L}{(D_1 + D_2 L)(1 - \rho_\psi L)} \varepsilon_{\psi_t} \\
+ \frac{D_7}{D_1 + D_2 L} \varepsilon_{\theta_t} + \frac{D_8}{D_1 + D_2 L} \varepsilon_{\psi_t}
\]
where

\[ D_1 = 1 \]
\[ D_2 = -\lambda \]
\[ D_3 = -(a_{21}g + a_{23}) \]
\[ D_4 = -a_{24} \]
\[ D_5 = \lambda(a_{21}g + a_{23}) \]
\[ D_6 = \lambda a_{24} \]
\[ D_7 = \left( -\frac{\varphi_1 \alpha + \varphi_2}{\lambda - \alpha} - \varphi_1 \right) g - \varphi_3 + a_{23} \]
\[ D_8 = -(\beta^{-1} - 1) - \frac{(\beta^{-1} - 1)\rho_\psi}{\lambda - \rho_\psi} \]

It can be shown that

\[
\frac{D_3 + D_5 L}{(D_1 + D_2 L)(1 - \rho_\theta L)} = \frac{m_1}{1 - \lambda L} - \frac{n_1}{1 - \rho_\theta L} = m_1 + \lambda m_1 L + \lambda^2 m_1 L^2 + \cdots
\]
\[-n_1 - \rho_\theta n_1 L - \rho_\theta^2 n_1 L^2 - \cdots \]

\[
\frac{D_4 + D_6 L}{(D_1 + D_2 L)(1 - \rho_\psi L)} = \frac{m_2}{1 - \lambda L} - \frac{n_2}{1 - \rho_\psi L} = m_2 + \lambda m_2 L + \lambda^2 m_2 L^2 + \cdots
\]
\[-n_2 - \rho_\psi n_2 L - \rho_\psi^2 n_2 L^2 - \cdots \]

where

\[
m_1 = \frac{D_5 + \lambda D_3}{\lambda - \rho_\theta}, \quad m_2 = \frac{D_6 + \lambda D_4}{\lambda - \rho_\psi}, \quad n_1 = \frac{D_5 + \rho_\theta D_3}{\lambda - \rho_\theta}, \quad n_2 = \frac{D_6 + \rho_\psi D_4}{\lambda - \rho_\psi} \]

Besides,

\[
\frac{D_7}{D_1 + D_2 L} = D_7 + \lambda D_7 L + \lambda^2 D_7 L^2 + \cdots
\]
\[
\frac{D_8}{D_1 + D_2 L} = D_8 + \lambda D_8 L + \lambda^2 D_8 L^2 + \cdots
\]
With some algebra, we can show
\[
\text{cov}(\tilde{b}_{t-1}, \tilde{\psi}_t) = -\rho_{\psi} \sigma_{\psi}^2 \left[ \frac{1}{1 - \lambda \rho_{\psi}} - \frac{1}{1 - \rho_{\psi}^2} \right] < 0
\]
and
\[
\text{var}(\tilde{b}_{t-1}) = \left[ \frac{(m_1 + D_7)^2}{1 - \rho_{\theta}^2} + \frac{n_2^2}{1 - \rho_{\psi}^2} - \frac{2n_1(m_1 + D_7)}{1 - \lambda \rho_{\theta}} \right] \sigma_{\theta}^2
\]

Obviously, the simultaneity bias in the case of Ricardian equilibrium is always negative unless \( \rho_{\psi} = 0 \).

**Appendix D: Derivation of (32) and (33)**

Firstly, we rewrite (31) as
\[
b_{t-1} = E_{t-1} \sum_{i=0}^{\infty} \left( \prod_{j=0}^{i} \pi_{t+j} R_{t+j-1}^{-1} \right) \left[ \tau_{t+i} + m_{t+i} - \frac{1}{\pi_{t+i}} m_{t+i-1} \right]
\]
\[
= E_{t-1} \sum_{i=0}^{\infty} \left( \prod_{j=0}^{i} \pi_{t+j} R_{t+j-1}^{-1} \right) \tau_{t+i} + E_{t-1} \sum_{i=0}^{\infty} \left( \prod_{j=0}^{i} \pi_{t+j} R_{t+j-1}^{-1} \right) m_{t+i}
\]
\[
- E_{t-1} \sum_{i=0}^{\infty} \left( \prod_{j=0}^{i} \pi_{t+j} R_{t+j-1}^{-1} \right) \frac{m_{t+i-1}}{\pi_{t+i}}
\]
(66)

Let us define \( \Delta \equiv \prod_{j=0}^{i} \pi_{t+j} R_{t+j-1}^{-1} \). Log linearizing (66) around the deterministic steady state and rearranging terms, we get
\[
\dot{b}_{t-1} = E_{t-1} \sum_{i=0}^{\infty} \delta^{i+1} \left( \tau + m - \frac{m}{\pi} \right) \hat{\Delta} + E_{t-1} \sum_{i=0}^{\infty} \delta^{i+1} \left( \tau \hat{\tau}_{t+i} + m \hat{m}_{t+i} - \frac{m}{\pi} \hat{m}_{t+i-1} + \frac{m}{\pi} \hat{\pi}_{t+i} \right)
\]
(67)
where $\delta \equiv \pi / R$ is the steady state discount factor and $\Delta = \sum_{j=0}^{i} (\hat{\pi}_{t+j} - \hat{R}_{t+j-1})$. Further simplifying (67) and applying (56), we get

$$b\hat{b}_{t-1} = E_{t-1} \sum_{i=0}^{\infty} \delta^i \left[ (\tau + m - \frac{m}{\pi}) \left( \frac{\delta}{1 - \delta} \right) (\hat{\pi}_{t+i} - \hat{R}_{t+i-1}) \right]$$

$$+ E_{t-1} \sum_{i=0}^{\infty} \delta^{i+1} \left[ \tau \hat{\pi}_{t+i} + R \left( \frac{\delta c - m}{R - 1} \right) \hat{R}_{t+i-1} - \frac{R}{\pi} \left( \frac{\delta c - m}{R - 1} \right) \hat{R}_{t+i-1} + \frac{m}{\pi} \hat{\pi}_{t+i} \right]$$

(68)

Expressing variables in log, (68) can be written as

$$\tilde{b}_{t-1} = E_{t-1} \sum_{i=0}^{\infty} \delta^i \left[ (\tau + m - \frac{m}{\pi}) \left( \frac{\delta}{1 - \delta} \right) (\tilde{\pi}_{t+i} - \tilde{R}_{t+i-1}) \right]$$

$$+ E_{t-1} \sum_{i=0}^{\infty} \delta^{i+1} \left[ \tau \tilde{\pi}_{t+i} + R \left( \frac{\delta c - m}{R - 1} \right) \tilde{R}_{t+i-1} - \frac{R}{\pi} \left( \frac{\delta c - m}{R - 1} \right) \tilde{R}_{t+i-1} + \frac{m}{\pi} \tilde{\pi}_{t+i} \right] + k$$

where

$$k = b \ln(b) - \frac{\delta}{(1 - \delta)^2} \left( \tau + m - \frac{m}{\pi} \right) \ln(\delta)$$

$$+ \frac{\delta}{1 - \delta} \left[ \frac{R}{\pi} \left( \frac{\delta c - m}{R - 1} \right) (1 - \pi) \ln(R) - \tau \ln(\tau) - \frac{m}{\pi} \ln(\pi) \right]$$

Appendix E: Derivation of (38) and (39)

Firstly, we apply (37) in (32) and get

$$b\tilde{f}_{t-1} \tilde{C}_b = k + E_{t-1} \sum_{i=0}^{\infty} \delta^i \left[ (\tau + m - \frac{m}{\pi}) \left( \frac{\delta}{1 - \delta} \right) (\tilde{f}_{t+i} \tilde{C}_\pi - \tilde{f}_{t+i-1} \tilde{C}_R) \right]$$

$$+ E_{t-1} \sum_{i=0}^{\infty} \delta^{i+1} \left[ \tau \tilde{f}_{t+i} \tilde{C}_\tau + R \left( \frac{\delta c - m}{R - 1} \right) \tilde{f}_{t+i} \tilde{C}_R - \frac{R}{\pi} \left( \frac{\delta c - m}{R - 1} \right) \tilde{f}_{t+i-1} \tilde{C}_R + \frac{m}{\pi} \tilde{f}_{t+i} \tilde{C}_\pi \right]$$

(69)
According to (36), we have
\[ E_{t-1} \bar{f}_{t+i} = \bar{f}_{t-1} B^{i+1} + \bar{B}_0 \sum_{k=0}^{i} B^k \]  
(70)

After applying (70) in (69) and rearranging terms, we get
\[ b \bar{f}_{t-1} \bar{C}_b = k + \sum_{i=0}^{\infty} \delta^{i+1} \left[ \left( \tau + m - \frac{m}{\pi} \right) \left( 1 - \frac{1}{1 - \delta} \right) + \frac{m}{\pi} \right] \left( \bar{f}_{t-1} B^{i+1} + \bar{B}_0 \sum_{k=0}^{i} B^k \right) \bar{C}_\pi \]
\[ + \sum_{i=0}^{\infty} \delta^{i+1} \left[ \left( \bar{f}_{t-1} B^{i+1} + \bar{B}_0 \sum_{k=0}^{i} B^k \right) \tau \bar{C}_\tau + \left( \frac{\delta c - m}{R - 1} \right) \left( \bar{f}_{t-1} B^{i+1} + \bar{B}_0 \sum_{k=0}^{i} B^k \right) R \bar{C}_R \right] \]
\[ - \sum_{i=0}^{\infty} \delta^{i+1} \left[ \left( \tau + m - \frac{m}{\pi} \right) \left( \frac{\delta}{1 - \delta} + \frac{\delta R}{\pi} \left( \frac{\delta c - m}{R - 1} \right) \right) \right] \left( \bar{f}_{t-1} B^{i+1} + \bar{B}_0 \sum_{k=0}^{i} B^k \right) \bar{C}_R \]
\[ - \left[ \left( \tau + m - \frac{m}{\pi} \right) \left( \frac{\delta}{1 - \delta} + \frac{\delta R}{\pi} \left( \frac{\delta c - m}{R - 1} \right) \right) \right] \bar{f}_{t-1} \bar{C}_R \]  
(71)

Assuming the VAR(1) process (36) is stationary, it can be shown that
\[ \sum_{i=0}^{\infty} \delta^i \sum_{k=0}^{\infty} B^k = \frac{1}{1 - \delta} \sum_{i=0}^{\infty} (\delta B)^i = \frac{1}{1 - \delta} (I - \delta B)^{-1} \]  
(72)

Applying (72) in (71), we get
\[ b \bar{f}_{t-1} \bar{C}_b = k + \left[ \left( \tau + m - \frac{m}{\pi} \right) \left( \frac{\delta}{1 - \delta} + \frac{\delta m}{\pi} \right) \bar{f}_{t-1} (I - \delta B)^{-1} B \bar{C}_\pi \right] \]
\[ + \left[ \left( \tau + m - \frac{m}{\pi} \right) \left( \frac{\delta}{1 - \delta} + \frac{\delta m}{\pi} \right) \bar{C}_\pi \right] \bar{f}_{t-1} B \bar{C}_\tau \]
\[ + \delta \tau \bar{f}_{t-1} (I - \delta B)^{-1} B \bar{C}_\tau \]
\[ + \delta R \left( \frac{\delta c - m}{R - 1} \right) \bar{f}_{t-1} (I - \delta B)^{-1} B \bar{C}_R \]
\[ + \delta R \left( \frac{\delta c - m}{R - 1} \right) \bar{C}_\pi \]
\[ - \left[ \left( \tau + m - \frac{m}{\pi} \right) \left( \frac{\delta}{1 - \delta} + \frac{\delta R}{\pi} \left( \frac{\delta c - m}{R - 1} \right) \right) \right] \bar{f}_{t-1} (I - \delta B)^{-1} B \bar{C}_R \]
\[ - \left[ \left( \tau + m - \frac{m}{\pi} \right) \left( \frac{\delta}{1 - \delta} + \frac{\delta R}{\pi} \left( \frac{\delta c - m}{R - 1} \right) \right) \right] \bar{f}_{t-1} \bar{C}_R \]  
(73)
After collecting terms of (73) that are with $\bar{f}_{t-1}$, we obtain the following two restrictions on $\bar{B}_0$ and $B$.

$$
\delta B \left\{ b \bar{C}_b + \left[ \left( \tau + m - \frac{m}{\pi} \right) \left( \frac{1}{1-\delta} \right) + \frac{m}{\pi} \right] \bar{C}_x + \tau \bar{C}_x + R \left( \frac{\delta c - m}{R-1} \right) \bar{C}_R \right\}
$$

$$
= b \bar{C}_b + \left[ \left( \tau + m - \frac{m}{\pi} \right) \left( \frac{\delta}{1-\delta} \right) + \frac{\delta R}{\pi} \left( \frac{\delta c - m}{R-1} \right) \right] \bar{C}_R
$$

$$
+ \left[ \left( \tau + m - \frac{m}{\pi} \right) \left( \frac{\delta}{1-\delta} \right) + \frac{\delta m}{\pi} \left( \frac{1}{1-\delta} \right) \right] \bar{B}_0 (I - \delta B)^{-1} \bar{C}_x + \left( \frac{\delta \tau}{1-\delta} \right) \bar{B}_0 (I - \delta B)^{-1} \bar{C}_x
$$

$$
+ \left[ \frac{\delta R}{1-\delta} \left( \frac{\delta c - m}{R-1} \right) \left( 1 - \frac{1}{R} \right) - \frac{\delta}{1-\delta} \left( \frac{\delta}{1-\delta} \right)^2 \right] \bar{B}_0 (I - \delta B)^{-1} \bar{C}_R = -k
$$

(74)

(75)

**Appendix F: Derivation of (49)**

As shown in Chung and Leeper (2007) and Traum (2007), the part of the objective function (43) that is related to $b$ can be rewritten as $-(b - \hat{b})'S(b - \hat{b})$, where $\hat{b} = vec((X'X)^{-1}(X'F))$ and $S \equiv (\Sigma^{-1} \otimes X'X)$. Maximizing (43) subject to (48), the first-order condition can be shown as $\tilde{b} = \hat{b} - S^{-1}V'\xi$, where $\xi$ is the Lagrangian multiplier of the constrained optimization problem. From the first-order condition, we can derive $\xi = (VS^{-1}V')^{-1}(V\tilde{b} - V\hat{b})$. Substituting this expression for $\xi$ back to the first-order condition, we get

$$
\tilde{b} = \hat{b} + S^{-1}V'(VS^{-1}V')^{-1}(vec(\bar{C}_0) - V\tilde{b})
$$

where we have imposed $V\tilde{b} = vec(\bar{C}_0)$.  

35
Figure 1: First Quadrant of (α, γ) Space

Figure 2: Simultaneity Bias in the Numerical Example (Case III)
Figure 3: Simultaneity Bias (Non-Ricardian)

Figure 4: Empirical Probabilities of Making Wrong Decisions (Non-Ricardian)
Figure 5: Simultaneity Bias (Ricardian)

Figure 6: Empirical Probabilities of Making Wrong Decisions (Ricardian)
Figure 7: Empirical Probabilities of Making Wrong Decisions (Non-Ricardian): OLS v.s. Unrestricted VAR, Test 1 and 2.

Figure 8: Empirical Probabilities of Making Wrong Decisions (Ricardian): OLS v.s. Unrestricted VAR, Test 3.
Table 1: Empirical Probabilities of Making Wrong Decisions (Non-Ricardian & Test 1): OLS v.s. Unrestricted VAR. Rows are for $\gamma$, columns are for $\alpha$. The empirical probability before / is for OLS and that after / is for unrestricted VAR. The last row and last column are averages taken over $\gamma$ and $\alpha$. The lower-right corner is the total average.

<table>
<thead>
<tr>
<th>$\gamma \setminus \alpha$</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>0.01/0.02</td>
<td>0/0.02</td>
<td>0/0.02</td>
<td>0/0.02</td>
<td>0/0.02</td>
<td>0/0.02</td>
<td>0/0.02</td>
<td>0/0.02</td>
<td>0/0.01</td>
<td>0/0.02</td>
</tr>
<tr>
<td>0.2</td>
<td>0.03/0.04</td>
<td>0.01/0.02</td>
<td>0/0.01</td>
<td>0/0.01</td>
<td>0/0.02</td>
<td>0/0.02</td>
<td>0/0.02</td>
<td>0/0.03</td>
<td>0/0.01</td>
<td>0/0.02</td>
</tr>
<tr>
<td>0.3</td>
<td>0.16/0.02</td>
<td>0.04/0.02</td>
<td>0.01/0.01</td>
<td>0/0.02</td>
<td>0/0.02</td>
<td>0/0.02</td>
<td>0/0.01</td>
<td>0/0.02</td>
<td>0/0.02</td>
<td>0.02/0.02</td>
</tr>
<tr>
<td>0.4</td>
<td>0.36/0.02</td>
<td>0.12/0.02</td>
<td>0.02/0.02</td>
<td>0/0.02</td>
<td>0/0.01</td>
<td>0/0.02</td>
<td>0/0.02</td>
<td>0/0.02</td>
<td>0/0.02</td>
<td>0.06/0.02</td>
</tr>
<tr>
<td>0.5</td>
<td>0.49/0.02</td>
<td>0.2/0.01</td>
<td>0.03/0.02</td>
<td>0/0.02</td>
<td>0/0.02</td>
<td>0/0.02</td>
<td>0/0.01</td>
<td>0/0.01</td>
<td>0/0.01</td>
<td>0.08/0.02</td>
</tr>
<tr>
<td>0.6</td>
<td>0.69/0.01</td>
<td>0.31/0.02</td>
<td>0.06/0.01</td>
<td>0.01/0.01</td>
<td>0/0.02</td>
<td>0/0.01</td>
<td>0/0.01</td>
<td>0/0.01</td>
<td>0/0.01</td>
<td>0.12/0.01</td>
</tr>
<tr>
<td>0.7</td>
<td>1/0.02</td>
<td>0.86/0.02</td>
<td>0.12/0.01</td>
<td>0/0.02</td>
<td>0/0.02</td>
<td>0/0.02</td>
<td>0/0.02</td>
<td>0/0.02</td>
<td>0/0.01</td>
<td>0.22/0.02</td>
</tr>
<tr>
<td>0.8</td>
<td>1/0.01</td>
<td>0.99/0.01</td>
<td>0.14/0.01</td>
<td>0/0.01</td>
<td>0/0.01</td>
<td>0/0.02</td>
<td>0/0.01</td>
<td>0/0.01</td>
<td>0/0.02</td>
<td>0.24/0.01</td>
</tr>
<tr>
<td>0.9</td>
<td>1/0</td>
<td>0.99/0</td>
<td>0.17/0.01</td>
<td>0/0.01</td>
<td>0/0.01</td>
<td>0/0.01</td>
<td>0/0</td>
<td>0/0.01</td>
<td>0/0</td>
<td>0.24/0.01</td>
</tr>
<tr>
<td>Mean</td>
<td>0.53/0.02</td>
<td>0.39/0.02</td>
<td>0.06/0.01</td>
<td>0/0.01</td>
<td>0/0.01</td>
<td>0/0.02</td>
<td>0/0.02</td>
<td>0/0.02</td>
<td>0/0.01</td>
<td>0.11/0.02</td>
</tr>
<tr>
<td>( \gamma ) ( \alpha )</td>
<td>0.1</td>
<td>0.2</td>
<td>0.3</td>
<td>0.4</td>
<td>0.5</td>
<td>0.6</td>
<td>0.7</td>
<td>0.8</td>
<td>0.9</td>
<td>Mean</td>
</tr>
<tr>
<td>--------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>----------</td>
</tr>
<tr>
<td>0.1</td>
<td>0.02/0.06</td>
<td>0.08/0.05</td>
<td>0.15/0.04</td>
<td>0.26/0.04</td>
<td>0.41/0.04</td>
<td>0.57/0.05</td>
<td>0.65/0.06</td>
<td>0.71/0.05</td>
<td>0.65/0.07</td>
<td>0.39/0.05</td>
</tr>
<tr>
<td>0.2</td>
<td>0.01/0.04</td>
<td>0.03/0.05</td>
<td>0.1/0.06</td>
<td>0.17/0.05</td>
<td>0.31/0.05</td>
<td>0.44/0.04</td>
<td>0.56/0.04</td>
<td>0.59/0.05</td>
<td>0.44/0.06</td>
<td>0.29/0.05</td>
</tr>
<tr>
<td>0.3</td>
<td>0/0.04</td>
<td>0.01/0.04</td>
<td>0.05/0.05</td>
<td>0.15/0.04</td>
<td>0.31/0.05</td>
<td>0.45/0.04</td>
<td>0.58/0.06</td>
<td>0.59/0.04</td>
<td>0.42/0.04</td>
<td>0.28/0.04</td>
</tr>
<tr>
<td>0.4</td>
<td>0/0.05</td>
<td>0.01/0.05</td>
<td>0.03/0.05</td>
<td>0.1/0.04</td>
<td>0.27/0.04</td>
<td>0.45/0.05</td>
<td>0.57/0.05</td>
<td>0.58/0.05</td>
<td>0.45/0.05</td>
<td>0.27/0.05</td>
</tr>
<tr>
<td>0.5</td>
<td>0/0.06</td>
<td>0.01/0.04</td>
<td>0.03/0.04</td>
<td>0.09/0.05</td>
<td>0.22/0.04</td>
<td>0.32/0.05</td>
<td>0.44/0.05</td>
<td>0.48/0.04</td>
<td>0.32/0.03</td>
<td>0.21/0.04</td>
</tr>
<tr>
<td>0.6</td>
<td>0/0.05</td>
<td>0/0.03</td>
<td>0.03/0.05</td>
<td>0.08/0.05</td>
<td>0.22/0.05</td>
<td>0.37/0.06</td>
<td>0.51/0.06</td>
<td>0.54/0.04</td>
<td>0.39/0.03</td>
<td>0.24/0.05</td>
</tr>
<tr>
<td>0.7</td>
<td>0/0.05</td>
<td>0/0.04</td>
<td>0.01/0.05</td>
<td>0.25/0.06</td>
<td>0.85/0.05</td>
<td>0.92/0.06</td>
<td>1/0.04</td>
<td>0.75/0.05</td>
<td>1/0.03</td>
<td>0.53/0.05</td>
</tr>
<tr>
<td>0.8</td>
<td>0/0.05</td>
<td>0/0.05</td>
<td>0/0.04</td>
<td>0.35/0.05</td>
<td>0.87/0.06</td>
<td>1/0.07</td>
<td>0.91/0.04</td>
<td>0.86/0.04</td>
<td>1/0.03</td>
<td>0.55/0.05</td>
</tr>
<tr>
<td>0.9</td>
<td>0/0.05</td>
<td>0/0.05</td>
<td>0/0.05</td>
<td>0.09/0.05</td>
<td>0.06/0.06</td>
<td>1/0.05</td>
<td>0.4/0.05</td>
<td>1/0.05</td>
<td>0.89/0.03</td>
<td>0.38/0.05</td>
</tr>
<tr>
<td>Mean</td>
<td>0/0.05</td>
<td>0.02/0.04</td>
<td>0.04/0.05</td>
<td>0.17/0.05</td>
<td>0.39/0.05</td>
<td>0.61/0.05</td>
<td>0.62/0.05</td>
<td>0.68/0.05</td>
<td>0.62/0.04</td>
<td>0.35/0.05</td>
</tr>
</tbody>
</table>

Table 2: Empirical Probabilities of Making Wrong Decisions (Non-Ricardian & Test 2): OLS v.s. Unrestricted VAR. Rows are for \( \gamma \), columns are for \( \alpha \). The empirical probability before / is for OLS and that after / is for unrestricted VAR. The last row and last column are averages taken over \( \gamma \) and \( \alpha \). The lower-right corner is the total average.
<table>
<thead>
<tr>
<th>$\gamma \setminus \alpha$</th>
<th>1.1</th>
<th>1.2</th>
<th>1.3</th>
<th>1.4</th>
<th>1.5</th>
<th>1.6</th>
<th>1.7</th>
<th>1.8</th>
<th>1.9</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>0.22/0.14</td>
<td>0.17/0.13</td>
<td>0.19/0.12</td>
<td>0.16/0.14</td>
<td>0.16/0.14</td>
<td>0.16/0.14</td>
<td>0.17/0.14</td>
<td>0.16/0.12</td>
<td>0.16/0.13</td>
<td>0.17/0.13</td>
</tr>
<tr>
<td>1.2</td>
<td>0.15/0.14</td>
<td>0.16/0.12</td>
<td>0.17/0.13</td>
<td>0.14/0.12</td>
<td>0.15/0.14</td>
<td>0.17/0.11</td>
<td>0.15/0.1</td>
<td>0.18/0.15</td>
<td>0.14/0.13</td>
<td>0.16/0.13</td>
</tr>
<tr>
<td>1.3</td>
<td>0.11/0.11</td>
<td>0.12/0.11</td>
<td>0.11/0.11</td>
<td>0.14/0.12</td>
<td>0.15/0.12</td>
<td>0.13/0.13</td>
<td>0.14/0.11</td>
<td>0.15/0.13</td>
<td>0.14/0.13</td>
<td>0.13/0.12</td>
</tr>
<tr>
<td>1.4</td>
<td>0.07/0.12</td>
<td>0.07/0.15</td>
<td>0.1/0.1</td>
<td>0.15/0.12</td>
<td>0.09/0.11</td>
<td>0.11/0.12</td>
<td>0.1/0.1</td>
<td>0.13/0.11</td>
<td>0.13/0.12</td>
<td>0.11/0.11</td>
</tr>
<tr>
<td>1.5</td>
<td>0.06/0.13</td>
<td>0.05/0.11</td>
<td>0.06/0.13</td>
<td>0.1/0.12</td>
<td>0.07/0.13</td>
<td>0.1/0.13</td>
<td>0.09/0.13</td>
<td>0.1/0.11</td>
<td>0.11/0.11</td>
<td>0.08/0.12</td>
</tr>
<tr>
<td>1.6</td>
<td>0.05/0.12</td>
<td>0.07/0.1</td>
<td>0.08/0.1</td>
<td>0.07/0.12</td>
<td>0.06/0.12</td>
<td>0.08/0.11</td>
<td>0.09/0.12</td>
<td>0.08/0.12</td>
<td>0.07/0.13</td>
<td>0.07/0.12</td>
</tr>
<tr>
<td>1.7</td>
<td>0.03/0.11</td>
<td>0.06/0.12</td>
<td>0.02/0.09</td>
<td>0.09/0.1</td>
<td>0.06/0.12</td>
<td>0.09/0.11</td>
<td>0.05/0.11</td>
<td>0.07/0.13</td>
<td>0.1/0.09</td>
<td>0.06/0.11</td>
</tr>
<tr>
<td>1.8</td>
<td>0.01/0.13</td>
<td>0.01/0.1</td>
<td>0.04/0.11</td>
<td>0.01/0.11</td>
<td>0.01/0.12</td>
<td>0.05/0.13</td>
<td>0.03/0.11</td>
<td>0.04/0.11</td>
<td>0.04/0.13</td>
<td>0.03/0.12</td>
</tr>
<tr>
<td>1.9</td>
<td>0/0.11</td>
<td>0.01/0.12</td>
<td>0.01/0.12</td>
<td>0.03/0.11</td>
<td>0.04/0.12</td>
<td>0.02/0.11</td>
<td>0.03/0.11</td>
<td>0.03/0.11</td>
<td>0.02/0.11</td>
<td>0.02/0.11</td>
</tr>
<tr>
<td>Mean</td>
<td>0.08/0.12</td>
<td>0.08/0.12</td>
<td>0.09/0.11</td>
<td>0.1/0.12</td>
<td>0.09/0.12</td>
<td>0.1/0.12</td>
<td>0.1/0.11</td>
<td>0.1/0.12</td>
<td>0.1/0.12</td>
<td>0.09/0.12</td>
</tr>
</tbody>
</table>

Table 3: Empirical Probabilities of Making Wrong Decisions (Ricardian & Test 3): OLS v.s. Unrestricted VAR. Rows are for $\gamma$, columns are for $\alpha$. The empirical probability before / is for OLS and that after / is for unrestricted VAR. The last row and last column are averages taken over $\gamma$ and $\alpha$. The lower-right corner is the total average.