Discussion Paper 37

Institute for Empirical Macroeconomics
Federal Reserve Bank of Minneapolis
250 Marquette Avenue
Minneapolis, Minnesota 55480

January 1991

THE MACROECONOMIC EFFECTS OF DISTORTIONARY TAXATION

Ellen R. McGrattan
Institute for Empirical Macroeconomics
and Department of Economics
Duke University

I would like to thank Tom Sargent, Richard Rogerson, Ken Judd, Sumuru Altug, Randy Wright, and Toni Braun for helpful comments and discussions.

This material is based on work supported by the National Science Foundation under Grant No. SES-8722451. The Government has certain rights to this material.

ABSTRACT

This paper develops a model of competitive economy which is used to study the effect that distortionary taxes have on the business cycle and on agents' welfare. In the presence of distortions, the equilibria are not Pareto optimal and standard computational techniques cannot be used. Instead, methods that take into account the presence of distorting taxes are applied. Maximum likelihood estimates of taste, technology and policy parameters from U.S. post-war time series are used to obtain several results. I find that a significant portion of the variance of the aggregate consumption, output, hours worked, capital stock, and investment can be attributed to the factor tax and government spending processes. Also, I compute the deadweight loss due to alternative tax changes and compare these estimates to others in the literature. Specification of taxes as constant versus state-contingent can have a significant effect on the results.

The Macroeconomic Effects of Distortionary Taxation

Ellen R. McGrattan †

Duke University

December, 1990

Abstract: This paper develops a model of a competitive economy which is used to study the effect that distortionary taxes have on the business cycle and on agents' welfare. In the presence of distortions, the equilibria are not Pareto optimal and standard computational techniques cannot be used. Instead, methods that take into account the presence of distorting taxes are applied. Maximum likelihood estimates of taste, technology and policy parameters from U.S. post-war time series are used to obtain several results. I find that a significant portion of the variance of the aggregate consumption, output, hours worked, capital stock, and investment can be attributed to the factor tax and government spending processes. Also, I compute the deadweight loss due to alternative tax changes and compare these estimates to others in the literature. Specification of taxes as constant versus state-contingent can have a significant effect on the results.

† Department of Economics, Duke University, Durham, NC 27706. I would like to thank Tom Sargent, Richard Rogerson, Ken Judd, Sumru Altug, Randy Wright, and Toni Braun for helpful comments and discussions.

1. Introduction

Real business cycle models in the tradition of Kydland and Prescott (1982)[1] assume that technological change is the driving force behind growth and fluctuations observed in post-war U.S. data. While these models have been successful in accounting for a large fraction of the variability and comovements of the aggregate time series, they do not do well along some dimensions. Relative to the data, the variability of consumption, hours of work, and output is too low, and the variability of investment and the correlation of real wages and hours are too high. One reason for this limited success may be the exclusion of monetary or fiscal shocks.

In this paper, the basic framework of Kydland and Prescott (1982) is extended to include a public sector. It is assumed that the government levies distortionary taxes on factors of production to finance its expenditures. As Christiano and Eichenbaum (1988) note, the inclusion of a public sector has the potential to improve some of the predictions of Kydland and Prescott's (1982) model. For example, with only productivity shocks driving fluctuations, the model of Kydland and Prescott cannot capture the fact that the correlation between hours and real wages is approximately zero. A shock to technology shifts the labor demand curve but not the labor supply curve. This explains the high correlation between hours and wages that Kydland and Prescott predict. To improve this prediction, Christiano and Eichenbaum (1988) study a real business cycle model in which government purchases affect agents' utility. The expenditures are financed through lump-sum taxes. In this case, shocks to expenditures shift the labor supply curve. However, they predict that while the hours and wage correlation comes closer to that observed, it is significantly positive.

Christiano and Eichenbaum (1988) do not allow for distortionary taxation. Like government expenditures, changes in the tax rates affect labor supply. Thus, tax rates provide another mechanism for explaining the observed correlation between hours and wages. Furthermore, if the tax rates are state contingent, then the model has the potential to improve the predictions of variability of output, consumption, investment, and hours of work. How the dynamics are altered will depend, however, on the particular form of the tax rules.

This paper shows that fiscal variables can be important determinants of cyclical behavior. This claim is quantified by "innovation accounting." Following Sims (1980), variances

[1] See, for example, Prescott (1986), Hansen (1985), Benhabib, Rogerson and Wright (1989), and Long and Plosser (1983).

are decomposed with fractions attributed to innovations in technology, government expenditures, labor tax rate, and capital tax rate processes. It is shown that a significant portion of the variance of the aggregate consumption, output, hours worked, capital stock, and investment can be attributed to the government expenditures and factor tax processes.

An advantage of the model developed in this paper is that it provides a framework for calculating the welfare costs due to capital and labor taxation. There exists a wide range of estimates of the costs of taxes on factors of production. This is true across models and across parameterizations. For example, Judd (1987) measures deadweight loss due to capital and labor taxation for different choices of a single agent's utility function and for various tax changes. In the case of a permanent increase in the tax on labor income the cost figures range from 2 cents per dollar revenue to over 1 dollar. For the capital income, the range is 15 cents to over 20 dollars and in one case the revenues fell with an increase in the capital tax rate. McGrattan (1989) also finds that the estimates are sensitive to parameterizations of preferences and tax policies. Results depend on whether the tax rates are assumed to be constant or state-contingent.

Because the welfare cost calculations are sensitive to parameterizations, the parameters of the model are estimated using U.S. aggregate time series. As in Altug (1989), the parameter estimates are those that maximize the likelihood function. One reason for using maximum likelihood over methods such as generalized method of moments (GMM) is that tax rates are assumed unobserved. The GMM procedure assumes that the econometrician observes all variables that the agents observe. For maximum likelihood, the identification of parameters is possible with latent variables because of the cross-equation restrictions imposed by the assumption of rational expectations.

To construct the likelihood function, it is necessary to first compute an equilibrium. The methods used to compute equilibria in this paper differ from those of Kydland and Prescott (1982). With distortionary taxes, the equivalence of the equilibria of the competitive system and the "planner's problem" does not obtain. Assuming a planner that maximizes utility for the representative agent subject to resource constraints does not work. Instead, a competitive equilibrium is computed directly. The price functions and laws of motion for aggregate quantities are determined endogenously and must be computed along with the decision rules of the households and firms.

With parameters of preferences, technologies, and policies given by the maximum likelihood estimates, welfare costs for a permanent changes in tax rates are computed.

Estimates fall in the range of those computed by Judd (1987) and McGrattan (1989). As in Judd (1987), the cost of the capital tax exceeds that of the labor tax. However, the estimates of the cost of the labor tax are more robust across specifications.

The model is described in Section 2. In Section 3, a method for computing equilibria is described. The technique involves approximating the true preferences of the household with a function that is quadratic. This approximation technique is used by Kydland and Prescott (1982). The estimation strategy is outlined in Section 4. The parameter estimates are discussed in Section 5.1 and comparisons are made to related empirical studies. With estimates imposed, the fractions of variance in aggregate time series due to productivity shocks and to government purchases and tax rate processes are calculated. These results are given in Section 5.2. The welfare costs for the estimated parameters and several others are reported in Section 5.3. A comparison of methods for solving equilibria, estimation strategies, and results is made to several recent related papers.[2] Concluding remarks are given in Section 6.

2. The Model

The economy in this paper is comprised of a representative household, a representative firm, and a government, all of whom are infinitely-lived. The household chooses sequences of consumption, $\{c_t\}$, investment, $\{i_t\}$, and hours worked, $\{n_t\}$, to maximize expected utility

$$E_0 \sum_{t=0}^{\infty} \beta^t U(c_t + \psi G_t, \alpha(L)\,\ell_t), \qquad 0 < \beta < 1 \tag{1}$$

where expectations are conditioned on the information set of the household at time 0. Current period utility depends on the number of goods consumed, government expenditures, G_t, and current and past hours of leisure, $\ell_t, \ell_{t-1}, \ldots$. The specification of (1) assumes that past leisure decisions affect current leisure services. If $\alpha(L) = \sum_{j=0}^{\infty} \alpha_j L^j$, where L is the lag operator, then one hour of leisure at t gives α_j hours of leisure services at $t+j$. If $\alpha_j = (1-\eta)^{j-1}\alpha_1$, $0 \leq \eta \leq 1$, $\alpha(1) = 1$ as in Kydland and Prescott (1982), the future services from one unit today are declining over time. Let $h_t = \sum_{j=1}^{\infty}(1-\eta)^{j-1} n_{t-j}$ assuming $\alpha(L)$ has declining weights. Let $\bar{H}$ be the time endowment each period so that

$$n_t + \ell_t = \bar{H}. \tag{2}$$

[2] See for example Braun (1990), Cassou (1990), Greenwood and Huffman (1989).

Then

$$\alpha(L)\ell_t = \bar{H} - \alpha_0 n_t - \eta(1-\alpha_0)h_t \tag{3}$$

$$h_{t+1} = (1-\eta)h_t + n_t \tag{4}$$

which is a recursive representation for past hours and a way to avoid The specification of the first argument of (1) assumes that government expenditures can influence the household's utility if $\psi \neq 0$. If $\psi > 0$, the marginal utility of consumption decreases with an increase in G_t. If $\psi < 0$, the opposite is true.

The budget constraints of the household are given by

$$c_t + i_t \le (1-\tau_{kt})r_t k_t + (1-\tau_{nt})w_t n_t + \delta\tau_{kt}k_t + T_t, \quad t = 1, 2, \ldots. \tag{5}$$

Each period, the household purchases c_t and investment goods i_t with after-tax income from renting the factors of production that they own to the firm. The capital income is $r_t k_t$, where r is the (real) price of renting capital and k is the capital stock of the household. Labor income is $w_t n_t$, where w is the (real) wage rate. If the household behaves competitively, then it is a price-taker in the capital and labor markets. In (5) the prices of consumption and investment goods are normalized to 1. Capital and labor income are taxed at rates τ_{kt} and τ_{nt}, respectively. T_t are lump-sum transfer payments from the government to the household in period t. The final source of income is depreciation allowances $\delta\tau_{kt}k_t$, where $0 \le \delta \le 1$ is the constant rate of capital depreciation.

The household owns the technology to convert investment and the current capital stock to next period capital. As in Kydland and Prescott (1982), capital takes four quarters to build. Total investment is a sum of investment in capital at various stages of production:

$$i_t = \phi_1 s_{t-3} + \phi_2 s_{t-2} + \phi_3 s_{t-1} + \phi_4 s_t, \qquad \phi_j \ge 0, \quad \sum_j \phi_j = 1 \tag{6}$$

where s_t is investment starts at time t. The parameters ϕ_j denote the fraction of resources allocated to projects j periods from completion. Thus, current investment consists of the value put in place during the first year of projects started in the current period, $\phi_4 s_t$, the value put in place during the second year of projects started in the previous period, $\phi_3 s_{t-1}$, and so on. In this case, the investment projects adding to the end-of-period capital stock at t are those started at $t-3$. Thus next period capital is

$$k_{t+1} = (1-\delta)k_t + s_{t-3} \tag{7}$$

where δ is the rate of depreciation of the capital stock in place.

Each period the firm chooses levels of output y_t, capital, k_t, and labor, n_t, so as to maximize profits:

$$y_t - r_t k_t - w_t n_t \tag{8}$$

subject to the technology

$$y_t = F(\lambda_t, k_t, n_t) \tag{9}$$

where F is a production function exhibiting constant returns-to-scale, λ_t is a stochastic technology shock, and r_t, w_t are input prices taken as given by the firm. Revenues are obtained from selling goods to the households and to the government.

The government levies taxes on factors of production to finance expenditures. Any revenue that is not used to finance current purchases is transferred to households in a lump-sum payment. Thus, real transfers are given by:

$$T_t = \tau_{kt} r_t K_t + \tau_{nt} w_t N_t - \tau_{kt} \delta K_t - G_t \tag{10}$$

where G_t is government expenditures (purchased at the same price as consumption and investment goods) and K_t, N_t, I_t are aggregate levels of capital, hours, and investment.[3] The government's debt is zero.

The stochastic processes governing technology shocks, government expenditures and tax rates are assumed to follow

$$v_{t+1} = a(L) v_t + b x_{2t} + \varepsilon_{t+1} \tag{11}$$

where $v_t = [\lambda_t, G_t, \tau_{kt}, \tau_{nt}]'$, L is a lag operator, $x_{2t} = [1,\, K_t,\, S_{t-1},\, S_{t-2},\, S_{t-3},\, H_t,\, I_t,\, N_t]'$ is a vector of variables that are not controllable by households, and ε_{t+1} is a vector white noise. The specification in (11) is very general. Constant values are accommodated by setting $a(L) = 0$, all columns of b but the first to zero, and $\varepsilon_t = 0$ for all t. Exogeneity of the processes in v_t is accommodated by setting $b = 0$. With b nonzero, specifications similar to that in Seater (1982) are possible. Seater (1982) finds that important determinants of the labor tax rate are federal government expenditures as a share of gross national product, per-capita gross national product, and the inflation rate. With (11), we can specify taxes as a function of government expenditures and its lags and (indirectly) output. The latter is

[3] Below, capital letters imply the aggregate level of the corresponding variable.

achieved by assuming that tax rates are functions of inputs to production. These variables are elements of x_{2t}. The specification in (11) does not accommodate nonlinear relationships between the elements of v_t. However, the procedure used to compute equilibria requires linear constraints. Any nonlinear functions are linearized. Thus, a linear specification is chosen.

Let X_t be the state vector for the household at time t. X includes current individual levels of capital, starts across projects, and accumulated labor, all aggregate quantities both private and public, and technology shocks. Partition $X_t = [X_{1t}', X_{2t}', X_{3t}']'$ into variables directly influenced by the individual household's controls, $X_{1t} = [k_t,\ s_{t-1},\ s_{t-2},\ s_{t-3},\ h_t]'$, all aggregate variables but investment and labor, $X_{2t} = [K_t,\ S_{t-1},\ S_{t-2},\ S_{t-3},\ H_t, ,\ \lambda_t, \ldots, \lambda_{t-l},\ G_t, \ldots, G_{t-1},\ \tau_{kt}, \ldots, \tau_{kt-l},\ \tau_{nt}, \ldots, \tau_{nt-l}]'$, and $X_{3t} = [I_t,\ N_t]'$, where l is the lag length of $a(L)$. Then the state of the economy is $[X_{2t}', X_{3t}']'$. The equations governing the variables in X_2 can be read off of equations (4),(6),(7), and (11) with individual levels replaced by aggregate levels. Functions for aggregate investment and hours, on the other hand, depend on the individual choice functions. These functions are not known *a priori*.

The household's problem can be formulated as a stationary discounted dynamic program with Bellman7s equation given by

$$V(X) = \max_{c,i,n}\Big(U(c + \psi G, \bar{H} - \alpha_0 n - \eta(1-\alpha_0)h) + \beta E\big[V(\tilde{X})|X\big]\Big) \tag{12}$$

subject to (4)-(7), the initial condition X_0, factor price and transfer functions, $r(X_2, X_3)$, $w(X_2, X_3)$, $T(X_2, X_3)$, and laws of motion for X_2, X_3. The firm maximizes its profits, (8), subject to (9), taking as given the price functions r and w. The government satisfies its sequence of budget constraints, (10).

Define $\mathtt{N} = [0, \bar{H}] \subset \mathbf{R}$, $\mathtt{X}_1 = \mathbf{R}_+^5$, $\mathtt{X}_2 = \mathbf{R}_+^5 \times \mathbf{R}^{4l}$, $\mathtt{X}_3 = \mathbf{R}_+ \times \mathtt{N}$, and Φ is a selector matrix such that $X_{1t} = \Phi X_{2t}$ implies $k_t = K_t$, $s_{t-j} = S_{t-j}$, $j = 1,2,3$, $h_t = H_t$, for all t. Assume that g_2 is a collection of transition functions for the variables in X_2: $X_{2t+1} = g_2(X_{2t}, X_{3t}, \varepsilon_{t+1})$.

Definition: A recursive *competitive equilibrium* is a collection of pricing functions, $r^* : \mathtt{X}_2 \times \mathtt{X}_3 \to \mathbf{R}$, $w^* : \mathtt{X}_2 \times \mathtt{X}_3 \to \mathbf{R}$, a transfer function, $T^* : \mathtt{X}_2 \times \mathtt{X}_3 \to \mathbf{R}$, policy functions, $c^* : \mathtt{X}_1 \times \mathtt{X}_2 \times \mathtt{X}_3 \to \mathbf{R}_+$, $i^* : \mathtt{X}_1 \times \mathtt{X}_2 \times \mathtt{X}_3 \to \mathbf{R}_+$, $n^* : \mathtt{X}_1 \times \mathtt{X}_2 \times \mathtt{X}_3 \to \mathtt{N}$, production plans, $y^* : \mathtt{X}_2 \times \mathtt{X}_3 \to \mathbf{R}_+$, $k^* : \mathtt{X}_2 \times \mathtt{X}_3 \to \mathbf{R}_+$, $\mathrm{n}^* : \mathtt{X}_2 \times \mathtt{X}_3 \to \mathtt{N}$, laws of motion $g_3^* : \mathtt{X}_2 \times \mathtt{X}_3 \times \mathbf{R}^{4l+5} \to \mathtt{X}_3$ and a value function $V^* : \mathtt{X} \to \mathbf{R}$ that satisfy *i.* Utility maximization: V^* solves (12) and c^*, i^*, n^* are optimal taking $r = r^*$, $w = w^*$, $T = T^*$, laws of motion for X_3 to be g_3^*, and $V = V^*$;

ii. Profit maximization: given prices r^*, w^*, the plans y^*, k^*, and n^* maximize the firm's profit function (8) subject to (9);

iii. Government budget constraints: (10) satisfied with $T = T^*$, $r = r^*$, $w = w^*$, $n = n^*$, $k = k^*$, and $X_1 = \Phi X_2$

iv. Market clearing:

$$\mathrm{n}^*(X_2, X_3) = n^*(\Phi X_2, X_2, X_3) \qquad \text{(labor)}$$

$$y^*(X_2, X_3) = c^*(\Phi X_2, X_2, X_3) + i^*(\Phi X_2, X_2, X_3) + X_2(l+7) \qquad \text{(goods)}$$

$$g_3^*(X_2, X_3, \varepsilon) = \begin{bmatrix} i^*(\Phi g_2(X_2, X_3, \varepsilon), g_2(X_2, X_3, \varepsilon), g_3^*(X_2, X_3, \varepsilon)) \\ n^*(\Phi g_2(X_2, X_3, \varepsilon), g_2(X_2, X_3, \varepsilon), g_3^*(X_2, X_3, \varepsilon)) \end{bmatrix}$$

The last term in the goods market clearing condition is government purchases which is the $(l+7)$th element of X_2. The last condition of (*iv*) says that the optimal investment and hours decision rules coincide with the laws of motion the household believes hold in the aggregate for investment and hours. The household has *perceptions* about the price functions and laws of motion governing aggregate capital, starts, labor, and investment. These perceptions have implications for decision rules. These decision rules combined with market-clearing conditions, in turn, have implications for the *realized* laws of motion of the aggregate quantities. In equilibrium, the realized functions are correctly anticipated.

3. The Linear-Quadratic Model

In general it is not possible to determine analytically the equilibria defined in Section 2. To compute equilibria numerically, there are several alternative methods that can be used.[4] The method used for this paper is described in detail in McGrattan (1990). As in Kydland and Prescott (1982), the nonlinear utility function U of (1) is replaced by an "approximate" quadratic function. This latter function is typically found by taking a second-order Taylor expansion of the utility function around the steady state.

If the model of Section 2 is fit to post-war U.S. time series, a likely outcome is a nonstationary system. Since the approximation method to be used requires that the underlying system have a steady state, some transformation of variables is necessary. It is

[4] See Taylor and Uhlig (1990) for a survey of algorithms for stochastic control problems.

assumed that C_t, I_t, G_t, Y_t, K_t, and S_{t-j}, $j = 1,2,3$, grow at the same geometric rate, μ. Over the post-war sample, growth rates for investment, output, government purchases, and capital are in the range of .49 to .63 percent per quarter. Hours of work and tax rates are assumed to be stationary.

The nonstationary variables at t, say G_t, are replaced by G_t/μ^t to transform the household's problem of Section 2 to a stationary one. To avoid two sets of notation, assume that all variables at t (except hours worked, past hours, and tax rates) in the "prime" equations to follow are detrended by μ^t. For the stationary model, replace (6) and (7) by

$$i_t = \frac{\phi_1}{\mu^3} s_{t-3} + \frac{\phi_2}{\mu^2} s_{t-2} + \frac{\phi_3}{\mu} s_{t-1} + \phi_4 s_t \tag{6'}$$

$$k_{t+1} = \frac{(1-\delta)}{\mu} k_t + \frac{1}{\mu^4} s_{t-3}. \tag{7'}$$

The transformed budget constraints for the household and government do not explicitly involve μ so (5) and (10) still hold in the transformed case. Assuming that output and capital stock grow at the same rate may not imply that the technology shock grows at rate μ. For the Cobb-Douglas specification, λ_t has to be adjusted by $\mu^{\theta t}$, where θ is labor's share of output. With y_t, λ_t, and k_t appropriately detrended, the production function in (9) remains the same. Also, assume that (11) is the specification for the stationary government expenditures and technology processes. The last change involves the discount factor of (1). In Section 5, the utility function is specified as $U(c,\ell) = (c^\gamma \ell^{1-\gamma})^\omega/\omega$ for the nonlinear, nonstationary economy. Since consumption is assumed to grow at rate μ and leisure is stationary, the transformed utility function at time t is $\mu^{\gamma\omega t} U(c_t/\mu^t, \ell_t)$. If we set the discount factor as $\beta\mu^{\gamma\omega}$, then the problem retains a constant discount factor. From here on, assume that the variables have been detrended unless otherwise specified. Also, assume (6′), (7′) replace (6), (7).

In equilibrium, rental rates for capital and wage rates are equated to the marginal products of capital and labor, respectively:

$$r_t = \partial F(\lambda_t, K_t, N_t)/\partial K_t \tag{13}$$

$$w_t = \partial F(\lambda_t, K_t, N_t)/\partial N_t. \tag{14}$$

If these functions and (10) are substituted into the household's budget constraint, then consumption can be written as a function of $x_t = [k_t,\ n_t,\ i_t,\ h_t,\ \lambda_t,\ K_t,\ N_t,\ I_t,\ G_t]$, assuming that U is strictly increasing in c and the budget constraint holds with equality. Call this function $c(x_t)$. Substitute $c(x_t)$ into the utility function and call the resulting function $\mathcal{U}(x_t)$. Then,

$$\begin{aligned}
U\big(c(x_t) + \psi G_t, \bar{H} - \alpha_0 n_t - \eta(1-\alpha_0)h_t\big) \\
&= \mathcal{U}(x_t) \\
&\simeq \mathcal{U}(\bar{x}) + \frac{\partial \mathcal{U}(x)}{\partial x}\Big|_{x=\bar{x}}(x_t - \bar{x}) + \frac{1}{2}(x_t - \bar{x})' \frac{\partial^2 \mathcal{U}(x)}{\partial x^2}\Big|_{x=\bar{x}}(x_t - \bar{x}) \\
&= X_t' Q X_t + u_t' R u_t + 2X_t' W u_t
\end{aligned} \tag{15}$$

where $\bar{x}$ is the steady state of the system, which exists because the model has been transformed to a stationary one, and $u_t = [i_t, n_t]'$. To get from the second equation in (15) to the third we use the fact that $x_t = \Delta_1 X_t + \Delta_2 u_t$ where Δ_1 and Δ_2 are the appropriate selector matrices.

The constraints of our control problem take the form

$$\begin{bmatrix} X_{1t+1} \\ X_{2t+1} \\ X_{3t+1} \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ 0 & A_{22} & A_{23} \\ 0 & A_{32} & A_{33} \end{bmatrix} \begin{bmatrix} X_{1t} \\ X_{2t} \\ X_{3t} \end{bmatrix} + \begin{bmatrix} B_1 \\ 0 \\ 0 \end{bmatrix} u_t + \begin{bmatrix} \epsilon_{1t+1} \\ \epsilon_{2t+1} \\ \epsilon_{3t+1} \end{bmatrix} \tag{16}$$

or $X_{t+1} = AX_t + Bu_t + \epsilon_{t+1}$.[5] In our case, ϵ_{1t} is a vector of zeros for all t and ϵ_{2t} is a function of ε_t. The matrices A_{32} and A_{33} are the coefficient matrices in the transition functions for X_3. They correspond to the laws of motion g_3 defined in Section 2. Thus, computing an equilibrium involves finding the values of A_{32} and A_{33} that impose $i_t = I_t$, $n_t = N_t$ (or $u_t = X_{3t}$).

If period t utility is given by the third expression in (15), then the household's problem can be formulated as a dynamic program with quadratic costs and linear constraints. The value function in the infinite horizon case is given by $X_t' P X_t$ where P is the fixed point of the Riccati difference equation. The optimal feedback rule for u_t in this case is linear (*e.g.* $u_t = -FX_t = -F_1 X_{1t} - F_2 X_{2t} - F_3 X_{3t}$) and depends on A_{32} and A_{33}. If the matrices A_{32} and A_{33}, which are taken as fixed by the household, satisfy $A_{3j} = -(F_1\Phi + F_2)A_{2j} - F_3 A_{3j}$, $j = 1, 2$, then an equilibrium is found. These conditions correspond to the condition (*iv*) of the definition for a recursive competitive equilibrium of Section 2. They imply that the laws of motion for aggregate hours and investment taken as given by the household coincide with those realized.

[5] In our case, $A_{12} = 0$ and $A_{13} = 0$ but the analytic gradients of the agent's policy function and of the likelihood function given in Appendix B assume only that $A_{21} = 0$, $A_{31} = 0$.

In Kydland and Prescott (1982), the competitive equilibrium can be obtained by solving the planner's problem. The planner maximizes the expected utility of the household subject to resource constraints such as the income identity. Since the planner knows that the individual's variables coincide with the aggregate variables, there is no distinction made. All constraints are known and there are no X_{3t} type states. Thus, in the distortion-free economy, the problem is reduced to computing F and P.

Given the decision rules for investment and hours, the equilibrium law of motion for our states can be expressed either in terms of X_t or X_{2t}:

$$\begin{aligned} X_{t+1} &= (A - BF)X_t + \epsilon_{t+1} \\ X_{2t+1} &= \big(A_{22} - A_{23}(I + F_3)^{-1}(F_1\Phi + F_2)\big)X_{2t} + \epsilon_{2t+1} \end{aligned} \tag{17}$$

where the second equation uses the fact that $u_t = X_{3t}$ in equilibrium. There are redundancies in the first equation of (17) since both individual and aggregate variables are included in X. However, either equation can be used for simulating time series given realizations of ϵ. Note that the only nonzero elements of ϵ_2 are in the equations for government purchases, the technology shock, and the tax rates.

4. Estimation Strategy[6]

To either system in (17) we add an equation relating observables, Z_t, and states and thus have the state-space representation:

$$\begin{aligned} X_{t+1} &= A^o X_t + \epsilon_{t+1} \\ Z_t &= CX_t + \xi_t \end{aligned} \tag{18}$$

where X_t is being used to mean either all states or all states in X_{2t} depending on the equation of (17) used, ξ_t is a vector of measurement errors which may be serially correlated, and Z_t is assumed to be stationary. If ξ_t is serially correlated, say $\xi_{t+1} = D\xi_t + \nu_t$, $E\nu_t\nu_t' = \Omega$, $E\nu_t\epsilon_t' = 0$, then we would replace the second equation of (18) with

$$\begin{aligned} z_t &= Z_{t+1} - DZ_t \\ &= (CA^o - DC)X_t + C\epsilon_{t+1} + \xi_{t+1} - D\xi_t \\ &= \bar{C}X_t + \bar{\nu}_t \end{aligned} \tag{19}$$

where

[6] This is the estimation strategy outlined in Harvey (1981) and Sargent (1989).

$$E\begin{bmatrix}\epsilon_{t+1}\\ \bar{\nu}_t\end{bmatrix}\begin{bmatrix}\epsilon_{t+1}\\ \bar{\nu}_t\end{bmatrix}' = \begin{bmatrix}\Sigma & \Sigma C'\\ C\Sigma & \Omega + C\Sigma C'\end{bmatrix}$$

The matrices A^o, C, and Σ are nonlinear functions of the parameters underlying preferences, technology, and fiscal policy. The matrices D, Ω are functions of the measurement error parameters. These parameters are elements of the $m \times 1$ vector Γ. If $[\epsilon_{t+1}, \bar{\nu}_t]$ is a normally distributed white noise process with covariance[7] matrix given in (19), then the maximum likelihood estimate is obtained by maximizing the following function with respect to Γ:

$$L(\Gamma) = -\frac{Tn}{2}\ln(2\pi) - \frac{T}{2}\ln|\Sigma_z| - \frac{1}{2}\text{trace}(\Sigma_z^{-1}S_{aa}) \tag{20}$$

where

$$S_{aa} = \frac{1}{T}\sum_{t=1}^{T}(z_t - \hat{z}_t)(z_t - \hat{z}_t)'$$
$$\hat{z}_t = E[z_t|z_{t-1}, z_{t-2}, \ldots, z_1, \hat{X}_0]$$

and n is the dimension of Z. To compute this function we need the prediction of z given past values and the covariance $\Sigma_z = E(z_t - \hat{z}_t)(z_t - \hat{z}_t)'$. Both are generated from the Kalman filter equations:

$$\begin{aligned}\hat{X}_{t+1} &= A^o\hat{X}_t + \mathcal{K}a_t\\ a_t &= z_t - \bar{C}\hat{X}_t\\ &= z_t - \hat{z}_t\end{aligned} \tag{21}$$

where

$$\begin{aligned}\hat{X}_t &= E[X_t|z_{t-1}, z_{t-2}, \ldots, z_1, \hat{X}_0]\\ \mathcal{K} &= (A^oS\bar{C}' + \Sigma C')\Sigma_z^{-1}\\ S &= A^oSA^{o\prime} + \Sigma - \mathcal{K}\Sigma_z\mathcal{K}'\\ \Sigma_z &= \bar{C}S\bar{C}' + \Omega + C\Sigma C'.\end{aligned}$$

The matrices $\mathcal{K}$ and S are the Kalman gain and $E(X_t - \hat{X}_t)(X_t - \hat{X}_t)'$, respectively.

---

[7] In the definition of equilibrium, consumption, investment, and hours choices are assumed to be positive. With $[\epsilon_{t+1}, \bar{\nu}_t]$ normally distributed and linear decision rules for investment and hours as specified in Section 3, there is no guarantee that these choices are positive. However, in practice, this is not a problem since means of consumption, investment, output, government purchases, capital stock, and hours are large relative to variances. See Figures 1-12.

Parameter	Value	Standard Error
δ	0.02260	0.00128
θ	0.5655	0.07279
α_0	0.000000321	0.1852
η	1.0	
γ	0.2733	0.02218
ω	-0.6989	0.6453
ψ	-0.01082	0.1717
β	0.99273	0.01955
ϕ_1	0.2543	0.01347
ϕ_2	0.05246	0.01093
ϕ_3	0.3201	0.01466
μ	1.0051	0.000209

Table 1. Structural Non-tax Parameter Values

Given values for the parameters in Γ, we compute the optimal matrix A^o and construct D, C, Σ, Ω. These matrices in turn are used to compute $\mathcal{K}$, $\mathcal{S}$, Σ_z. The sequence of innovations $\{a_t\}_{t=1}^{T}$ are obtained recursively via (21) given an initial value for $\hat{X}_0$ and a sequence of observations $\{Z_t\}_{t=1}^{T+1}$.

5. Results

5.1. Data, Estimation, Tests of the Model

The data are per-capita aggregate output (*i.e* consumption plus investment plus government purchases), investment, government purchases, hours worked, and capital stock for the United States over the sample 1947:1-1987:4.[8] The vector of observables is given by $Z_t = [Y_t/\mu^t, I_t/\mu^t, G_t/\mu^t, K_t/\mu^t, N_t]'$, where Y, I, G, K, and N are the data described in Appendix A. The data are detrended by μ to make the elements of Z stationary. As shown by Dhrymes (1970), this detrending implies that $L(\Gamma)$ must be adjusted by adding $\sum_{t=1}^{T} \ln(\mu^{-4t})$. Because output is a nonlinear function of the state X_t, we used a linear expansion of the function around the steady state for the first row of C of (18).

Starts, S_t, weighted past hours, H_t, and tax rates, τ_{kt} and τ_{nt} are assumed to be unobserved. Although there have been measurements of effective tax rates[9], there are

[8] See Appendix A for definitions and sources. A quasi-Newton method with a Broyden-Fletcher-Goldfarb-Shanno update was used for the hillclimbing routine. Analytical gradients were available and are given in Appendix B.

[9] See for example, Joines (1981), Seater (1982,1985), Barro (1983).

Parameter	Value	Standard Error
σ_1	0.00947	0.002803
σ_2	5.1713	0.4851
σ_3	-0.000989	0.009407
σ_4	-0.001303	0.001609
σ_5	9.5873	0.2820
σ_6	0.006626	0.02141
σ_7	0.002270	0.001598
σ_8	0.01224	0.04319
σ_9	-0.01416	
σ_{10}	0.008313	0.02805

Table 2. Values Corresponding to Variance-Covariances

large differences in the constructed series. In McGrattan (1989), different series were used for computing welfare costs and the results varied greatly across series. One reason for choosing maximum likelihood estimation over methods such as generalized method of moments is that with GMM, testing orthogonality conditions involving the tax processes requires observation of these series. The GMM procedure was used by Braun (1989) who estimates a model with distortionary taxation. Braun (1989) does not consider the welfare effects of the taxes and his results may not critically depend on his choice of series.

The following functional forms were chosen for production and utility:

$$\begin{aligned} F(\lambda, K, N) &= \lambda K^{1-\theta} N^{\theta}, \quad 0 < \theta < 1 \\ U(c, l) &= \frac{(c^{\gamma} l^{1-\gamma})^{\omega}}{\omega}, \quad 0 < \gamma < 1, \ \omega \leq 1. \end{aligned} \tag{22}$$

Given these functions and an initial guess for Γ, the equilibrium decision rules are computed at each iteration of the hillclimbing routine, and the method of Section 4 is used to construct the likelihood function. To get a good initial guess of $\hat{\Gamma}$ that maximizes the likelihood function, we first estimated the parameters of preferences and technology with GMM using tax series from Joines (1981) as the "observed" tax series. For initial coefficients of the vector autoregression describing the evolution of tax rates, the technology shock and government expenditures, we used least squares estimates. Any regression coefficients not significantly different from zero were initially set to zero for the maximum likelihood estimation.

In Tables 1-4, parameter estimates and standard errors are provided. Any parameters not reported have been set to zero with the exception of $\bar{H}$ which was set to 1304.5, the

Parameter	Value	Standard Error
Ω_{11}	10^{-8}	
Ω_{22}	10^{-8}	
Ω_{33}	477.49	
Ω_{44}	0.6528	0.2385
Ω_{55}	10^{-8}	
D_{33}	0.9243	0.02955

Table 3. Measurement Error Parameter Values

number of weeks in the quarter times 100 hours per week discretionary time. Standard errors are not reported for η, σ_9, Ω_{33}, and several coefficients in the tax rate equations of Table 4. The value of η was restricted to 1.0 in estimation because it hit the upper bound whenever free. The remaining parameters reported without standard errors were free during estimation but assumed fixed when the standard errors were computed. This was done because the information matrix was close to numerically singular with all parameters included. However, comparing the standard errors with all but η free to that under the assumption that σ_9, Ω_{33}, and some tax rate coefficients are not identified, there is very little difference with the exception of the standard errors on γ and ω. Thus we report errors for the subset of parameters in which we know the information matrix is invertible.

In Table 1, we give values for the parameters of the production function (θ), depreciation (δ), time-to-build (ϕ_j), the utility function (α_0, α, β, η, γ, ω), and the growth rate (μ). The estimate of labor's share in the production function, θ, is 0.566 with a standard error of 0.07. This point estimate is lower than estimates typically found for models with taxes absent. But if the tax and no-tax models are to match the same observed capital-labor ratios, the no tax model will overpredict labor's share. For most constructed tax rate processes, the effective marginal tax rate for capital exceeds that of labor over the post-war sample. (See Figures 11 and 12.) Thus, there is a larger difference in the before and after tax return to capital than the before and after tax wage rate. As a result, a lower level of capital labor ratio is found unless θ is smaller. Altug (1989), who estimated a version of Kydland and Prescott's (1982) model, got an estimate for this parameter of 0.70. The estimate found in this paper is consistent with Braun (1989) who estimates a similar model via GMM and finds labor's share to be .54.

The value of depreciation, δ, is 0.023 with a small standard error. In a model without the time-to-build assumption on capital, this value is pinned down by observations on

	Coefficients of $v_{t+1} = a(L)v_t + bx_{2t}$			
v_t	λ_{t+1}	G_{t+1}	τ_{kt+1}	τ_{nt+1}
1	0.00958 (0.0069)	2.316 (1.22)	-0.0227 (0.0816)	0.00415 (0.010)
λ_t	1.0995 (0.0329)		0.0175 (0.0611)	-0.00274 (0.023)
λ_{t-1}	0.0694 (0.0822)			
λ_{t-2}	-0.1448 (0.0758)			
λ_{t-3}	-0.03131 (0.02359)			
G_t		1.527 (0.0392)	-.00000177 (0.000018)	-0.0000025 (0.000018)
G_{t-1}		-0.4407 (0.0805)		
G_{t-2}		-0.0916 (0.0482)		
τ_{kt}			1.399 (2.151)	
τ_{kt-1}			-0.402 (2.147)	
τ_{nt}				1.055 (0.0346)
τ_{nt-1}				-0.0681 (0.0350)
K_t			1.17×10^{-8}	
I_t			-4.04×10^{-7}	
N_t				-0.000147
H_t				0.000125

Table 4. Technology, Spending, and Tax VAR Estimates

capital and investment. Here, the starts enters the capital equation but the steady state of starts is approximately equal to the steady state of investment, the difference being attributed to positive growth rates. Thus, there should be consensus across models for the rate of capital depreciation in the range of our estimate.

Three starting points were tried for the time to build parameters, ϕ_j: the GMM estimates, Altug's (1989) estimates, and Kydland and Prescott's choice of .25 for each. The resulting estimates suggest that an equal weight choice can be rejected and that most resources (70%) are put in in the beginning two periods of the project and very little is done in the third stage (5.2%). Using data from the U.S. Census on nonresidential construction completions, Taylor (1982) also found little evidence for equally weighted fractions.

The growth rate, μ, is estimated to be 1.0051 with a standard error of 0.0002. The quarterly rate of growth of .51 percent per quarter falls in the range of mean growth rates for output, investment, capital, and government expenditures. Our assumption of balanced growth imposes this growth rate on all of our observables except hours of work which is assumed to be stationary.

The parameters of the utility function (α_0, α, β, η, γ, ω) are also given in Table 1. The value of 1.0 for η implies that current leisure services depend on current leisure and leisure from the last period. Since the value for α_0 is very close to 0, this would further imply that only n_{t-1} enters the period t utility. These estimates differ from Altug (1989) and Eichenbaum, Hansen, and Singleton (1988) who found evidence for both n_t and n_{t-1} entering the period t utility.

The value of ψ which governs the effect of government on the household's utility is -.011. For $\psi < 0$, the marginal utility of the household is increased with an increase in government expenditures. However, given a standard error of .172, we cannot reject the specification of utility that does not depend on government consumption.

The point estimate for γ implies that the share of consumption in utility is .27 with a standard error of .022. To interpret this estimate, we follow Eichenbaum, Hansen, and Singleton (1988) and construct a crude estimate of γ as follows. We first note that the first order conditions of our nonlinear problem imply the following marginal conditions: $U_c(c_t, \ell_t) = U_\ell(c_t, \ell_t)w_t(1-\tau_{nt})$. Using the utility function of (22), solving for γ and substituting in steady state values gives $\gamma \approx c/(\ell w(1-\tau_n)+c)$. Using approximate sample averages[10] of C, N, Y, and τ_n equal to 1400, 300, 2600, and .2, respectively we have $\gamma \approx 0.26$. With $\tau_n = 0$, this crude estimate becomes 0.22 which is less than Kydland and Prescott's estimate of 1/3.

The value of ω is -0.699 with a standard error of 0.645, which gives some evidence of the preferences being different from logarithmic. However, the point estimate is just

[10] See Figures 1-12.

barely one standard deviation from $\omega = 0$.

While the standard error on the discount factor, β, is .019, the point estimate of .992 is economically meaningful. In many studies, an estimate for β greater than 1 has been found. Altug (1989) fixes β during estimation to avoid such problems.

The parameters related to the variance-covariance matrices Σ and Ω are given in Tables 2 and 3. It is assumed that

$$\Sigma = \begin{bmatrix} \sigma_1 & 0 & 0 & 0 \\ \sigma_2 & \sigma_5 & 0 & 0 \\ \sigma_3 & \sigma_6 & \sigma_8 & 0 \\ \sigma_4 & \sigma_7 & \sigma_9 & \sigma_{10} \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 & 0 & 0 \\ \sigma_2 & \sigma_5 & 0 & 0 \\ \sigma_3 & \sigma_6 & \sigma_8 & 0 \\ \sigma_4 & \sigma_7 & \sigma_9 & \sigma_{10} \end{bmatrix}' . \tag{23}$$

From Table 2, it is clear that the technology shock is not the only source of fluctuations. In Table 3, the setting of Ω_{ii}, $i = 1, 3, 5$ to 10^{-8} implies that the measurement error of capital, government purchases, and output are close to zero. The value of 10^{-8} is used to avoid numerical singularities. For hours of work the variance in the measurement error is relatively small. For investment, however, the measurement error process is picking up dynamics that the model (and, in particular, the tax on capital) is unable to.

In Table 4, we provide estimates for the coefficients of the v process. The series are highly persistent which may not be too surprising given the persistence of the observed series. (See Figures 1-12.) For stationarity, we required the eigenvalues of A^o to be less than 1 in absolute value (or less than some number strictly less than 1 for numerical reasons). None of the eigenvalue bounds was reached during estimation. In a first pass of this estimation problem, the processes in v_t were freely parameterized. As a result, many of the parameters of the vector autoregression were not identified and, hence, the specification was restricted for the second pass. The technology shock and government expenditures as reported in Table 4 are assumed to be functions of their own past values. Effects of taxes enter via the disturbances.

In McGrattan (1989), we found that the computation of the welfare costs was sensitive to specification of tax processes. In particular, the results differed significantly for cases in which the tax rates were stochastic and state-contingent or stochastic as they are here and when they were constant. Since some of the coefficients of the tax processes are not estimated accurately, we want to check to see if the likelihood value changes significantly under the assumption of constant tax rates. In Table 5, we report the likelihood value from the estimated parameters of Tables 1-4, $L(\Gamma)$, the likelihood value in which the coefficients

$L(\Gamma)$	-3414.77
$L(\tilde{\Gamma})$	-4061.10
λ_{LR}	1292.66

Table 5. Unconstrained and constrained Likelihood Values

in the equations of τ_{kt} and τ_{nt} are constrained to be zero with the exception of the constant term and the parameters of Σ, $L(\tilde{\Gamma})$, and the likelihood ratio $\lambda_{LR} = 2\big(L(\Gamma) - L(\tilde{\Gamma})\big)$ which is used to test the zero restrictions. The number of zero restrictions is 12 and λ_{LR} is asymptotically distributed as $\chi^2(12)$. With 12 degrees of freedom, a χ^2 value larger than 1292 has a zero probability. Thus, the value of λ_{LR} gives strong evidence against the specification of constant tax rates.

Two additional sets of restrictions could also be tested: (*i*) set $\sigma_j = 0$, $j = 3, 4, 6, 7, 8, 9, 10$, or (*ii*) set the constants in the τ_k and τ_n equations to zero. The first set of restrictions would make τ_{kt} and τ_{nt} constant and nonstochastic as in Judd (1987) and the second set of restrictions would give us the nested no-tax model. This latter case could be thought of as a version of Kydland and Prescott's (1982) model if government purchases were included and financed by lump-sum taxes. Note that these additional restrictions will result in lower likelihood values than that reported in Table 5 for $\tilde{\Gamma}$.

5.2. Time Series Implications

Following Sims (1980), we can use the vector autoregressive representation in (17) to determine the fraction of the variance of X attributable to each innovation. The variance-covariance matrix, $var(X)$, solves the equation

$$var(X) = A^o var(X) A^{o\prime} + \Sigma. \tag{24}$$

To decompose the variance of the state into fractions attributable to each of the four shocks in ϵ, first factor (via a Cholesky factorization) Σ as LL' where L is a lower triangular matrix and then for each element i of ϵ corresponding to the technology, government purchases, and tax rate shocks, replace Σ in (24) with $Le_{ii}L'$, where e_{ii} is a matrix of zeros with element (i,i) equal to 1. With Σ replaced by $Le_{ii}L'$ in (24), solve for the variances and the fixed point will be the fraction of the variance attributable to the ith variable. Since the ordering of the state vector affects the decomposition, the variances for two different orderings are reported.

First, we choose an orthogonalization that gives the technology shock its greatest role in terms of variability. The results are given in Table 6. The ordering of the variables is such that λ is first, G second, τ_k third, and τ_n fourth and, for example, the fraction of the variance of capital stock attributable to each shock is .68, .14, .18, and 0 respectively. In this case, the technology shock is capturing only about 70% of capital, investment, consumption, and output. Note also that only 35% of the variance of hours worked in this case is due to the technology shock. The results of Kydland and Prescott (1982) also indicate that the fluctuations in hours is underpredicted with only the technology shock included. With at least 30% of the variability in aggregate output, consumption, investment, and hours due to government expenditures or taxes, there is evidence that not all of the dynamics of the system can be explained by technology shocks.

The second ordering of the state vector gives the technology shock its smallest role. The results are reported in Table 7. When λ is ordered last, it captures only about 30% of the variance in many of the aggregate quantities with τ_k and G making up much of the remainder. While comparisons can also be made across different orderings of the government expenditures and tax shocks, our intent in reporting Tables 6 and 7 was to point out the importance of incorporating the public sector. However, a few remarks concerning how the tax rate and government spending processes affect the time series might be of some help.

In testing the potential of their model to fit a subset of second moments, Kydland and Prescott (1982) choose the variance of output so as to match that particular standard deviation to U.S. gross national product. In doing so, they attribute all of the variance in output to the technology shock. They find that the standard deviation of series such as hours of work are significantly underpredicted. Their model hours has a standard deviation of 1.05 (after taking logarithms and detrending) while the same statistic for the data they report is 2.0. By restricting the shocks to enter only the production function, they also have difficulties resolving certain correlations such as that between hours and real wages (or productivity). The observed correlation between hours and wages is approximately zero. As Christiano and Eichenbaum (1988) note, what is needed to match this observation is some mechanism for affecting labor supply. They do this by introducing government purchases in a growth model. This can also be done with taxation on factors of production. As McGrattan (1989) shows, with taxes included, the correlation between hours and wages is not significantly different from zero.

In duplicating the exercise of Kydland and Prescott (1982) with both the estimates of Section 5.1 and in McGrattan (1989) we find that variability in investment and capital is increased significantly with increases in taxes on capital and that variability in consumption, hours, and output is increased significantly with increases in taxes on labor. One explanation for an increase in the volatility of hours worked with the inclusion of taxes is given by the equation of τ_n in Table 4. The labor tax falls with an increase in the technology shock. When there is a positive shock to the economy's production opportunities, agents substitute labor for leisure. From the τ_n equation, it is also clear that this positive effect on labor is further increased by a decrease in the labor tax. With a decrease in the labor tax rate, agents work more. The opposite occurs with a negative shock. Greenwood and Huffman (1989) and Braun (1989) also find that increases in tax rates imply increases in volatility. Cassou (1990) reports increases in volatility with increases in income taxes but not with increases in corporate taxes.

5.3. Welfare Implications

In addition to studying the effect that distortionary taxes have on aggregate fluctuations, we examine their effect on expected lifetime utility and revenues. That is, we compute the change in utility and distortionary revenues due to a change in one of the tax rates. This is done with the estimated parameters of Tables 1-4. Also, it is necessary to introduce a new state variable. Let τ_t^s be an additive shock in the equations determining τ_{kt} or τ_{nt}. If the changes in tax shocks are assumed to be permanent, the transition function for this variable is given by

$$\tau_{t+1}^s = \tau_t^s \tag{25}$$

which can be appended to the vector autoregression in v_t. Assume that our state vector, X_t, includes this tax shock. Also assume that the matrices A, B, Q, R, and W of Section 3 have been appropriately changed. Note that when $\tau_0^s = 0$, this new system results in exactly the same decision rule and time series as the old system.

To compute the change in utility and the change in distortionary revenues due to changes in the taxes, it is necessary to first specify the utility and revenue functions. The utility function is the discounted sum of the function given in (15) and can be written as

z	% Variance of z Attributable to Shocks to			
	λ	G	τ_k	τ_n
λ	1.00	0.00	0.00	0.00
K	0.68	0.14	0.18	0.00
G	0.23	0.77	0.00	0.00
τ_k	0.64	0.13	0.23	0.00
τ_n	0.05	0.06	0.60	0.29
I	0.67	0.15	0.18	0.00
N	0.35	0.20	0.12	0.34
C	0.68	0.15	0.15	0.02
Y	0.71	0.12	0.16	0.01

Table 6. Variance Decomposition Based on Ordering $\lambda, G, \tau_k, \tau_n$

z	% Variance of z Attributable to Shocks to			
	τ_k	τ_n	G	λ
τ_k	0.29	0.01	0.45	0.26
K	0.24	0.01	0.46	0.29
τ_n	0.51	0.38	0.08	0.03
G	0.15	0.22	0.64	0.00
λ	0.01	0.03	0.53	0.43
I	0.25	0.01	0.46	0.28
N	0.26	0.09	0.17	0.47
C	0.21	0.05	0.44	0.30
Y	0.22	0.01	0.46	0.31

Table 7. Variance Decomposition Based on Ordering $\tau_k, \tau_n, G, \lambda$

the function $V(X_0)$ where

$$\begin{aligned} V(X_0) &= E_0 \sum_{t=0}^{\infty} \beta^t (X_t' Q X_t + u_t' R u_t + 2X_t' W u_t) \\ &= X_0' P X_0 + \frac{\beta}{1-\beta} \text{trace}(P\Sigma). \end{aligned} \tag{26}$$

P is the solution to the Riccati equation and Σ is the covariance matrix of ϵ_t.

The present discounted value of future distortionary revenues is given by

$$\begin{aligned} R(X_0) &= E_0 \sum_{t=0}^{\infty} \beta^t p_t^0 (\tau_{kt} \frac{\partial F(\lambda_t, K_t, N_t)}{\partial K_t} K_t + \tau_{nt} \frac{\partial F(\lambda_t, K_t, N_t)}{\partial N_t} N_t - \delta \tau_{kt} K_t) \\ &= E_0 \sum_{t=0}^{\infty} \beta^t \mathcal{R}(X_t) \end{aligned} \tag{27}$$

where $p_t^0 = U_c(c_t + \psi G_t, \bar{H} - \alpha_0 n_t - \eta\ (1-\alpha_0)h_t)$, and $E_0[\beta^t p_t^0 c_t]$ is the date zero value of date t consumption.[11]

As in the case of the utility function, some approximation is necessary to compute the present value in (27). Assume that $\mathcal{R}(X_t) \simeq X_t' M X_t$. This quadratic function is obtained by taking a second-order Taylor expansion around the steady state. Since X_t contains a constant term, the constant and linear terms of the Taylor expansion are incorporated in M. In this case, we have

$$
\begin{aligned}
R(X_0) &= E_0 \sum_{t=0}^{\infty} \beta^t X_t' M X_t \\
&= \sum_{t=0}^{\infty} \beta^t X_0'(A^{o\prime})^t M (A^o)^t X_0 + \sum_{j=1}^{\infty} \beta^j \sum_{t=0}^{\infty} \beta^t E_0 \epsilon_j' (A^{o\prime})^t M (A^o)^t \epsilon_j \\
&= X_0' \Big(\sum_{t=0}^{\infty} \beta^t (A^{o\prime})^t M (A^o)^t \Big) X_0 + \sum_{j=1}^{\infty} \beta^j \sum_{t=0}^{\infty} \beta^t \mathrm{trace}\Big((A^{o\prime})^t M (A^o)^t \Sigma \Big) \\
&= X_0' \Big(\sum_{t=0}^{\infty} \beta^t (A^{o\prime})^t M (A^o)^t \Big) X_0 + \frac{\beta}{1-\beta} \mathrm{trace}\Big(\sum_{t=0}^{\infty} \beta^t (A^{o\prime})^t M (A^o)^t \Sigma \Big)
\end{aligned}
\tag{28}
$$

since $X_{t+1} = A^o X_t + \epsilon_{t+1}$ and $E\epsilon_s \epsilon_t' = 0, t \neq s$. If we define S_r to be $\sum_{t=0}^{\infty} \beta^t (A^{o\prime})^t M (A^o)^t$, then

$$
S_r = \beta A^{o\prime} S_r A^o + M. \tag{29}
$$

Solving (29) for S_r, we can then write $R(X_0)$ as

$$
R(X_0) = X_0' S_r X_0 + \frac{\beta}{1-\beta} \mathrm{trace}(S_r \Sigma). \tag{30}
$$

Following Judd (1987), we define the marginal deadweight loss ($\mathcal{M}_{dwl}$) from a permanent change in a tax rate to be the ratio of the derivative of lifetime utility with respect to the tax shock to the derivative of the present value of distortionary revenues with respect to the tax shock:

$$
\begin{aligned}
\mathcal{M}_{dwl} &= \frac{\partial V(X_0)}{\partial \tau_0^s} / \frac{\partial R(X_0)}{\partial \tau_0^s} \\
&= \frac{e' P X_0}{e' S_r X_0}
\end{aligned}
\tag{31}
$$

[11] The pricing kernal p_t^0 is related to Arrow-Debreu prices as follows. Let ρ_t^0 be the time 0 state-contingent price of consumption at t contingent on the history $\epsilon^t = (\epsilon_1, \epsilon_2, \ldots, \epsilon_t)$ and X_0 and let $f_t(\epsilon^t)$ be the density function of ϵ^t. Then $p_t^0(\epsilon^t, X_0) = \rho_t^0(\epsilon^t, X_0)/(\beta^t f_t(\epsilon^t))$.

and the welfare cost of this tax change is $-\mathcal{M}_{dwl}$. In (31), e is a vector of zeros with a 1 in the position corresponding to τ_0^s. Because of the computational procedure used in this paper, we can analyze changes only near a particular steady state which is why we compute derivatives of value functions and revenue functions. Greenwood and Huffman (1989), on the other hand, use a technique for solving the dynamic program and simulate time series from nonlinear decision rules. Thus, they can analyze the effects of a major tax reform but estimation is infeasible.

In Table 8, we report estimates of the welfare costs for the estimated tax processes from Table 4, $\hat{\tau}_k$ and $\hat{\tau}_n$ and several alternatives. All other parameters used for these calculations are obtained in Tables 1-4. The first row in Table 8 are the costs associated with the predicted tax rate series that have transition functions given in Table 4. These series are those labelled "predicted" in Figures 11 and 12. The welfare cost of a permanent change in the capital tax rate is 88 cents per dollar of distortionary revenue raised.[12] The labor tax has an efficiency cost of 13 cents per dollar revenue. These estimates fall in the range of McGrattan7s (1989) and Judd's (1987) estimates. They also support Judd's (1987) claim that "the excess burden of permanent capital taxation substantially exceeds that of permanent wage taxation."

The constant tax rates .583 and .092 for the second calculation are the steady state values of τ_k and τ_n given the estimates of Tables 1-4. That is, the second tax rates are the elements of $\bar{X}$ corresponding to τ_k and τ_n with parameters of preferences and technologies set to their estimated values. From Figure 11, we see that .583 is the maximum value for the capital tax rate over the sample. From Figure 12, we see that .092 is the minimum value for the labor tax rate over the sample. Thus, the steady state values are not equal to the sample means. This is due in part to setting $\hat{X}_0 = \bar{X}$ when estimating.[13] Since initial capital is the low point of the sample and initial hours worked is high, the steady state tax rate on capital is high and the the steady state tax rate on labor is low. (See Figures 7, 9, 11, and 12.)

The third set of tax rates are the estimates found with zero restrictions imposed on the tax rate equations. These restrictions are imposed in computing the constrained likelihood

---

[12] "Dollar" refers to the value in terms of date 0 consumption.

[13] An attempt was made to use the initial observations on Z to avoid steady states not being approximately equal to sample means. The result of this exercise was large initial innovations because initial values chosen for unobserved states were not consistent with those set for the observed states.

Current rates		Welfare Costs for τ_k	Welfare Costs for τ_n
$\tau_k = \hat{\tau}_k$,	$\tau_n = \hat{\tau}_n$	0.879	0.131
$\tau_k = .583$,	$\tau_n = .092$	0.503	0.094
$\tau_k = .475$,	$\tau_n = .35$	0.400	0.223
$\tau_k = .583$,	$\tau_n = .3$	0.539	0.201
$\tau_k = .3$,	$\tau_n = .092$	0.167	0.061
$\tau_k = .583$,	$\tau_n = \hat{\tau}_n$	0.473	0.129
$\tau_k = \hat{\tau}_k$,	$\tau_n = .092$	1.019	0.095

Table 8. Welfare Costs of Permanent Changes in τ_k, τ_n

value described in Section 5.1. Note that these values (.475, .35) come closer to the sample means of $\hat{\tau}_k$ and $\hat{\tau}_n$. Two additional sets of constant tax rates are also used. The high tax rate case has $\tau_k = .583$, $\tau_n = .3$ and the low tax rate case has $\tau_k = .3$, $\tau_n = .092$. Finally, we analyze the cases with one rate stochastic and the other set to its steady state value.

Comparing the costs for the estimated tax rates to those with rates at their steady state values, we find the costs are underestimated by the latter. The cost of the capital tax for the estimated rates is almost twice that of the constant rates. The same is true for constant rates closer to the sample means. Comparing the different constant tax cases, we see that the cost increases with the rates (or level of distortions). This is also the case for Greenwood and Huffman (1989) and Judd (1987) who both specify constant rates. More evidence of the sensitivity of the results to specification of tax processes is found when comparing the first case with the last two. This is especially true for the costs due to the tax on capital.

The estimates of the costs of capital taxation also vary widely across nonconstant specifications of the tax processes. Although not reported, we found with estimates from the first pass of our estimation procedure (*i.e.* with λ_t and G_t processes more freely parameterized) that the welfare calculations implied a permanent increase in τ_{kt} actually decreased the present value of revenues. That is, an increase in τ_{kt} caused a decline in investment and output and, hence, a decline in income from renting capital and labor. The welfare cost was negative. In McGrattan (1989), for several cases with state-contingent taxes similar results are obtained. Also, the calculations of the effect of labor taxation are found to be more robust to specification of parameters and tax processes. With these results, there is some apprehension in concluding that the burden of permanent capital taxation exceeds that of permanent labor taxation. Also, the efficiency losses from labor or income taxes may be underpredicted in this model with the human capital element of labor ignored. Driffill and Rosen (1983) predict greater costs of labor taxation when agents can accumulate human capital.

6. Conclusions

The effects of distortionary tax policies are studied in the context of a dynamic recursive stochastic equilibrium model. The model is a modified version of the distortion-free economy studied by Kydland and Prescott (1982). The presence of distortions in our model requires abandoning their method of computing an equilibrium, which exploited the optimality of the equilibrium of their distortion-free economy. In the spirit of Kydland and Prescott (1982), we study a linear quadratic approximation to the distorted economy and use methods designed to compute equilibria for such approximate economies.

Estimates of the parameters underlying the model are obtained via maximum likelihood with United States post-war data. Given the particular tax policies implied by the estimates, the model is used to study the time-series and welfare implications. Some of the predictions of Kydland and Prescott (1982) can be improved by the inclusion of taxes. It is shown that government expenditures and tax rate shocks have a significant effect on the variance of most of the variables in the model. The variance of output attributable to a technology shock is estimated to be at most 70%.

The estimates of the welfare cost of capital and labor taxation are 88 cents and 13 cents per dollar of revenue, respectively. Specifying taxes as constant versus state-contingent are important for this result. For the capital tax rate, the choice of the transition function may also be critical. That is, the estimated costs of capital taxation are a function of the zero restrictions imposed during estimation. The results for labor taxation are more robust.

The methods described in this paper can be applied to formulate and estimate a variety of recursive equilibrium models with externalities and distortions. The present paper serves partly to illustrate the feasibility of using these methods to study artificial economies with relatively large state spaces. In future work, our hope is to compute the equilibrium directly from the nonlinear model to see how robust are the results from using linear-quadratic approximation. We would also like to relax the assumption of a representative agent so as to explore the distributional effects of taxation across agents with different preferences and investment opportunities.

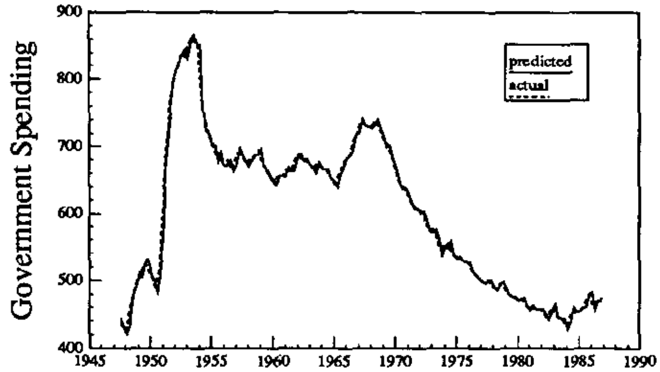

Figure 1. Predicted and actual government spending.

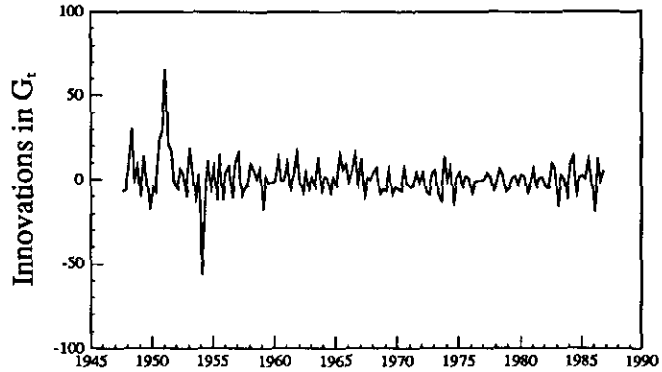

Figure 2. Innovations in government spending.

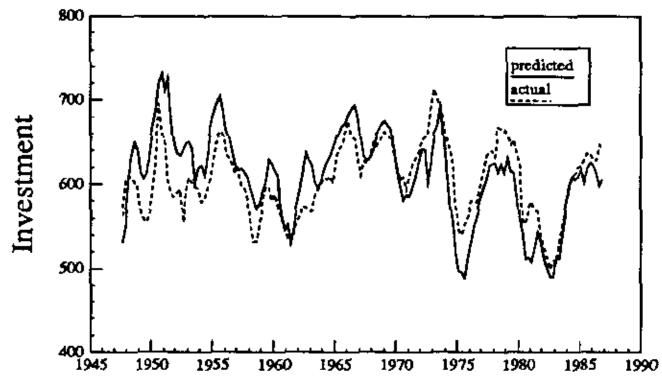

Figure 3. Predicted and actual investment.

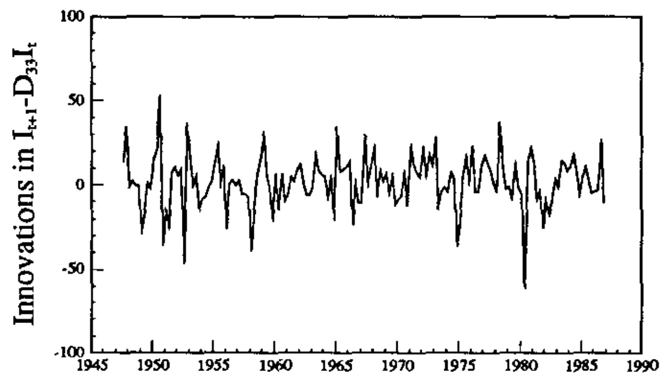

Figure 4. Innovations in investment.

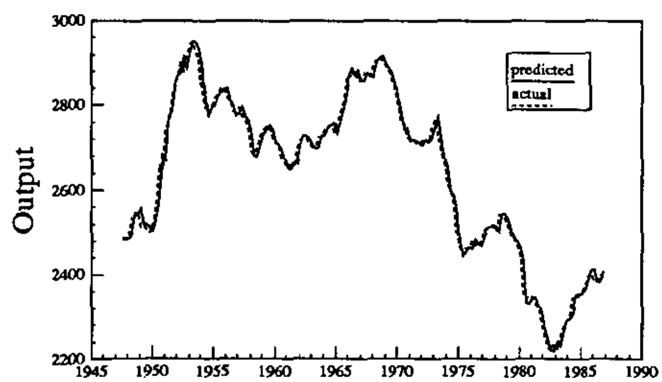

Figure 5. Predicted and actual output.

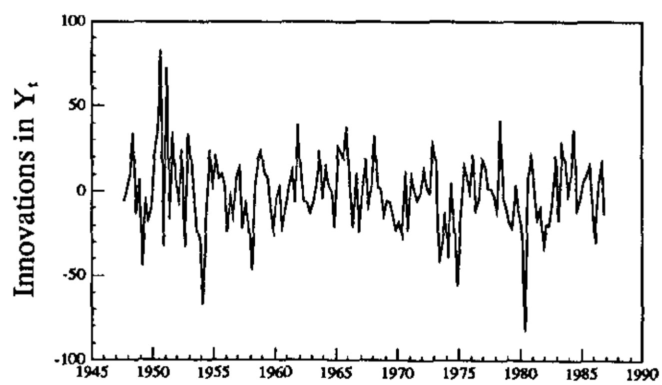

Figure 6. Innovations in output.

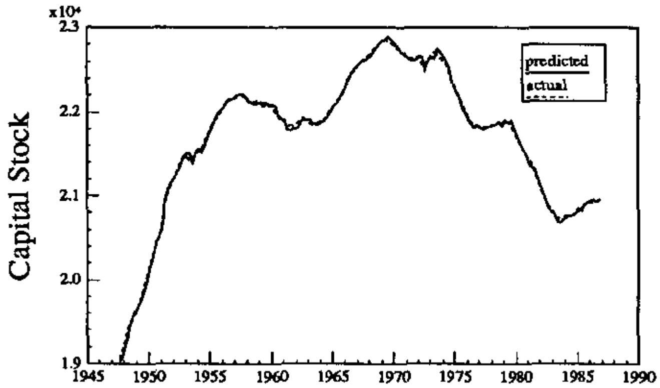

Figure 7. Predicted and actual capital stock.

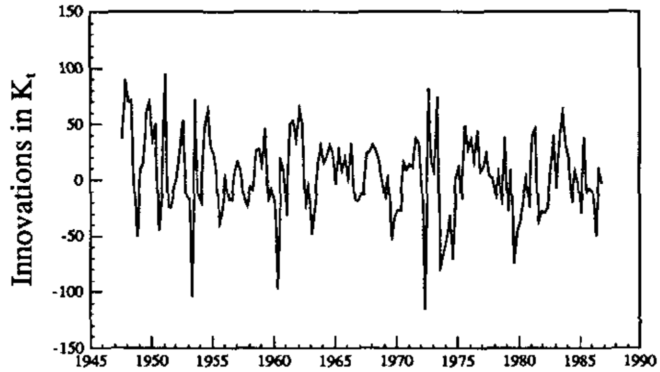

Figure 8. Innovations in capital stock.

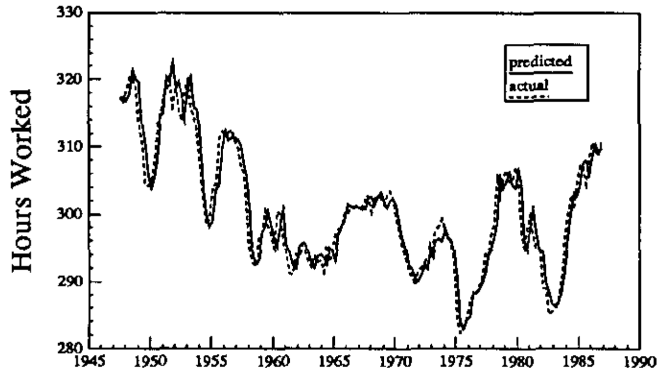

Figure 9. Predicted and actual hours worked.

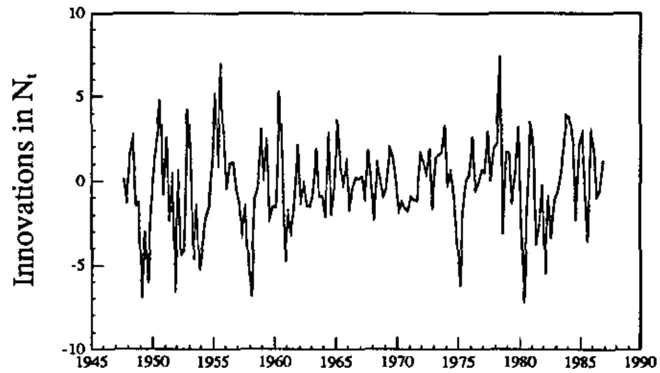

Figure 10. Innovations in hours worked.

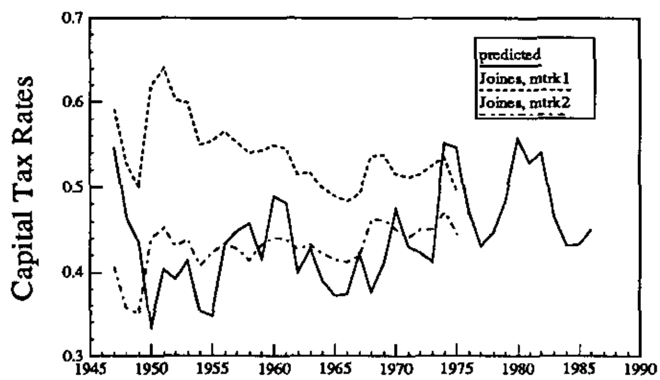

Figure 11. Predicted and measured capital tax rates (annual).

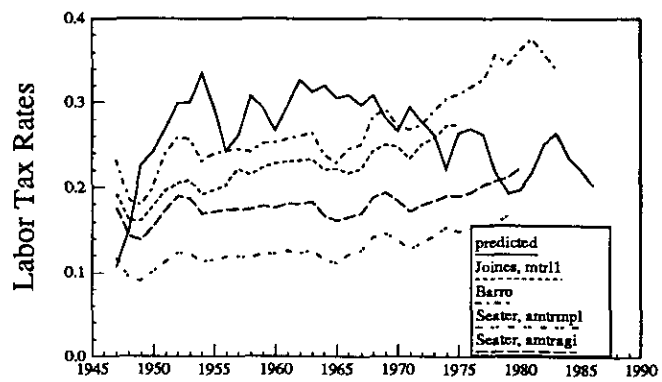

Figure 12. Predicted and measured labor tax rates (annual).

Appendix A

The data used in this study are real aggregate data of the United States for sample 1947:1-1987:4. All annual series (*i.e.* capital and tax rates) are log-linearly interpolated to obtain quarterly observations. The final numbers were obtained by dividing the series listed by the population series given below.

i. C_t: personal consumption expenditures of nondurable goods and services (Source: *National Income and Product Accounts*, Table 1.2 or Citibase variables GCN82, GCS82)

ii. I_t: private fixed investment plus personal consumption expenditures of durable goods (Source: *National Income and Product Accounts*, Table 1.2 Citibase variables GIF82, GCD82)

iii. G_t: government purchases of goods and services (Source: *National Income and Product Accounts*, Table 1.2 or Citibase variable GGE82)

iv. N_t: total manhours employed per week (Source: U.S. Department of Labor, BLS, *The Employment Situation–Household Survey* or Citibase variable LHOURS)

v. K_t: constant-dollar net stock of fixed private capital plus net stock of durable goods (Source: John Musgrave, "Fixed Reproducible Tangible Wealth in the U.S.: Revised Estimates," *Survey of Current Business*, January, 1986, Table 4, row 1, and Table 20, column 1 (with updates in *Survey of Current Business*, issues: August, 1987 and October, 1988)

vi. τ_{kt}: effective marginal tax on capital income

a. Source: rates constructed by Joines (1981), Table 3, columns 2-5 (MTRKj, $j = 1, 2, 3, 4$)

b. Source: rates constructed by Seater (1982), Table 2, column 7

vii. τ_{nt}: effective marginal tax on labor income

a. Source: rates constructed by Barro (1986), Table 2, column 6 ($\bar{\tau}$)

b. Source: rates constructed by Joines (1981), Table 2, columns 2-5 (MTRLj, $j = 1, 2, 3, 4$)

c. Source: rates constructed by Seater (1985), Table 2, columns 2,3,6,7 (AMTRAGI, AMTRGNP, AMTRHI, AMTRMPL)

viii. Population measure: civilian noninstitutional population, 16 years and older (Source: U.S. Department of Labor, BLS, *The Employment Situation* or Citibase variable P16)

Appendix B

In this appendix, we provide analytic gradients for the likelihood function given in (37).[14] The formulas provided assume that the control problem solved is of the form maximize $\sum \beta^t(X_t'QX_t + u_t'Ru_t + 2X_t'Wu_t)$ subject to the constraints (16). It is also assumed that the gradients ∇A_{11}, ∇A_{12}, ∇A_{13}, ∇B_1, ∇C, $\nabla \Sigma$, $\nabla \Omega$, $\nabla \bar{x}$, $\partial^j \mathcal{U}(x,p)/\partial x^j$, $j = 1,2,3$, $\partial \mathcal{U}(x,p)/\partial p$, and $\partial^j \mathcal{U}(x,p)/\partial x^j \partial p$, $j = 1,2$ are known, where column i of ∇A is equal to $vec(\partial A/\partial \Gamma_i)$, $\Gamma = [_1, \ldots, \Gamma_m]'$. The function $\mathcal{U}$ is the same one given in (15) but we have made the parameter argument, p, explicit. The vector $x = \Delta_1 X + \Delta_2 u$ are those states and controls entering the return function. To simplify expressions for gradients, we will also make use of the following definitions:

$$
\begin{aligned}
\mathcal{R} &= (R + \beta B_1' P_{11} B_1)^{-1} \\
\mathcal{B} &= \beta \mathcal{R} B_1' \\
\mathcal{W}_j &= \mathcal{R}(\beta B_1' P_{11} A_{1j} + W_j'), \quad j = 2,3 \\
\mathcal{Q}_j &= Q_{1j} + \beta A_1^{o\prime} P_{11} A_{1j} - F_1' W_j', \quad j = 2,3 \\
F_1 &= (R + \beta B_1' P_{11} B_1)^{-1}(\beta B_1' P_{11} A_{11} + W_1') \\
F_j &= (R + \beta B_1' P_{11} B_1)^{-1}(\beta B_1'(P_{11} A_{1j} + P_{12} A_{2j} + P_{13} A_{3j}) + W_j'), \quad j = 2,3 \qquad (B1) \\
P_{1j} &= Q_{1j} + \beta (A_{11} - B_1 F_1)'(P_{11} A_{1j} + P_{12} A_{2j} + P_{13} A_{3j}) - F_1' W_j', \quad j = 1,2,3 \\
A_1^o &= A_{11} - B_1 F_1 \\
\Psi &= -(I + F_3)^{-1}(F_1 \Phi + F_2) \\
\bar{x} &= \Delta_1 \bar{X} + \Delta_2 \bar{u} \\
&= D_1 \bar{X}_1 + D_2 \bar{X}_2 + D_3 \bar{X}_3 + \Delta_2 \bar{u}.
\end{aligned}
$$

Analytical derivatives of the likelihood function require $\partial A^o/\partial \Gamma_i$, $\partial C/\partial \Gamma_i$, $\partial \Sigma/\partial \Gamma_i$, $\partial \Omega/\partial \Gamma_i$, $\partial D/\partial \Gamma_i$ for each free parameter Γ_i. The first of these derivatives is found by analyzing the formulations from the control problem. Since, $A^o = A - BF$, we need derivatives for A, B, F. It is assumed that gradients for all partitions but A_{32}, A_{33} are known for A, and that ∇B_1 is known. Thus, we must derive ∇A_{32}, ∇A_{33}, ∇F. These matrices are functions of the matrices of the return function which in turn are functions

[14] Derivations of these gradients are not provided but are available upon request.

of $\mathcal{U}(x)$ when evaluated at the steady state. Thus we need:

$$
\begin{aligned}
\nabla q &= \nabla_x \frac{\partial^2 \mathcal{U}(x,p)}{\partial x^2}\bigg|_{x=\bar{x}} \nabla \bar{x} + \nabla \frac{\partial^2 \mathcal{U}(x,p)}{\partial x^2}\bigg|_{x=\bar{x}} \\
\nabla R &= \frac{1}{2}(\Delta_2' \otimes \Delta_2')\,\nabla q \\
\nabla Q_{1j} &= \frac{1}{2}(D_j' \otimes D_1')\,\nabla q \\
\nabla W_j &= \frac{1}{2} I_{nk}(D_j' \otimes \Delta_2')\,\nabla q
\end{aligned}
\tag{B2}
$$

where column i of $\nabla_x A$ is equal to $vec(\partial A/\partial x_i)$ and I_{mn} is a matrix of zeros and ones defined by $vec(A') = I_{mn} vec(A)$ for any $m \times n$ matrix A. The subscripts n and k on I in (B2) and below indicate the dimension of the vector X_1 and u respectively. Let $f(M) = .5(M' \nabla b - M' \partial^2 \mathcal{U}(x,p)/\partial x^2 \nabla x + (x' \otimes M') \nabla q)$, where

$$
\nabla b = \frac{\partial^2 \mathcal{U}(x,p)}{\partial x^2}\bigg|_{x=\bar{x}} \nabla \bar{x} + \frac{\partial \mathcal{U}(x,p)}{\partial x \partial p}\bigg|_{x=\bar{x}}. \tag{B3}
$$

If element i of X_2 is 1, then add $f(D_1)$ to the rows of ∇Q_{12} corresponding to this element and $f(\Delta_2)$ to the rows of $\nabla W_2'$ corresponding to this element.

In computing F_1 we solved the Riccati matrix for the first partition and thus need the gradients:

$$
\begin{aligned}
\nabla P_{11} = {} & [I - \beta(A_1^o \otimes A_1^o)]^{-1}\{\nabla Q_{11} + \beta[(I \otimes A_1^{o\prime} P_{11}) + (A_1^{o\prime} P_{11} \otimes I) I_{nn}]\,\nabla A_{11} \\
& - \beta[(F_1' \otimes A_1^{o\prime} P_{11}) + (A_1^{o\prime} P_{11} \otimes F_1') I_{nk}]\,\nabla B_1 + (F_1' \otimes F_1')\,\nabla R \\
& - [(F_1' \otimes I) + (I \otimes F_1') I_{nk}]\,\nabla W_1\} \\
\nabla F_1 = {} & \beta(I \otimes \mathcal{R} B_1' P_{11})\,\nabla A_{11} + [\beta(A_1^{o\prime} P_{11} \otimes \mathcal{R}) I_{nk} - (F_1' \otimes \mathcal{B} P_{11})]\,\nabla B_1 \\
& - (F_1' \otimes \mathcal{R})\,\nabla R + (I \otimes \mathcal{R}) I_{nk}\,\nabla W_1 + (A_1^{o\prime} \otimes \mathcal{B})\,\nabla P_{11}.
\end{aligned}
\tag{B4}
$$

Note that these derivatives should correspond to those derived by Zadrozny (1988) and would be used with A, B, Q, W, P and F replacing A_{11}, B_1, Q_{11}, W_1, P_{11}, F_1 in a distortion-free economy.

For ∇F_2, ∇F_3, we require a few intermediate computations:

$$
\begin{aligned}
\nabla\mathcal{R} &= -(\mathcal{R}\otimes\mathcal{R})\,\nabla R - [(\mathcal{B}P_{11}\otimes\mathcal{R})I_{nk} + (\mathcal{R}\otimes\mathcal{B}P_{11})]\,\nabla B_1 - \beta^{-1}(\mathcal{B}\otimes\mathcal{B})\,\nabla P_{11} \\
\nabla\mathcal{B} &= (B_1'\otimes I)\,\nabla\mathcal{R} + (I\otimes\mathcal{R})I_{nk}\,\nabla B_1 \\
\nabla\mathcal{W}_j &= \beta((A_{1j}'P_{11}B_1 + W_j)\otimes I)\,\nabla\mathcal{R} + \beta(A_{1j}'P_{11}\otimes\mathcal{R})I_{nk}\,\nabla B_1 + (A_{1j}'\otimes\mathcal{B})\,\nabla P_{11} \\
&\quad + (I\otimes\mathcal{B}P_{11})\,\nabla A_{1j} + (I\otimes\mathcal{R})\,\nabla\mathcal{W}_1' \\
w_j &= \nabla\mathcal{W}_j + (A_{2j}'(P_{12}+P_{13}\Psi)'\otimes I)\,\nabla\mathcal{B} + \big(I\otimes\mathcal{B}(P_{12}+P_{13}\Psi)\big)\,\nabla A_{2j} \\
\nabla\mathcal{Q}_j &= \nabla Q_{1j} + \beta(A_{1j}'P_{11}\otimes I)I_{nn}\,\nabla A_{11} - \beta(A_{1j}'P_{11}\otimes F_1')I_{nk}\,\nabla B_1 \\
&\quad - [\beta(A_{1j}'P_{11}B_1\otimes I) + (W_j\otimes I)]I_{nk}\,\nabla F_1 + \beta(A_{1j}'\otimes A_1^{o\prime})\,\nabla P_{11} \qquad (B5) \\
&\quad + \beta(I\otimes A_1^{o\prime}P_{11})\,\nabla A_{1j} - (I\otimes F_1')I_{nk}\,\nabla\mathcal{W}_j \\
\nabla A_1^o &= \nabla A_{11} - (I\otimes F_1)\,\nabla B_1 - (B_1'\otimes I)\,\nabla F_1 \\
q_j &= \nabla\mathcal{Q}_j + \beta(A_{2j}'(P_{12}+P_{13}\Psi)'\otimes I)I_{nn}\,\nabla A_1^o + \big(I\otimes\beta A_1^{o\prime}(P_{12}+P_{13}\Psi)\big)\,\nabla A_{2j} \\
\nabla F_j &= \omega_j + (A_{2j}'\otimes\mathcal{B})(I - ((A_{22}+A_{23}\Psi)'\otimes\beta A_1^{o\prime} - P_{13}((I+F_3)^{-1}\mathcal{B})))^{-1}\cdot \\
&\quad \Big[q_2 + (\Psi'\otimes I)q_3 - (\Psi'\otimes P_{13}(I+F_3)^{-1})w_3 - (\Phi'\otimes P_{13}(I+F_3)^{-1})\,\nabla F_1 \\
&\quad - (I\otimes P_{13}(I+F_3)^{-1})w_2\Big]
\end{aligned}
$$

for $j = 2,3$. Combining the submatrices of F and computing the gradients for ∇A_{3j}, we have

$$
\begin{aligned}
\nabla F &= \begin{bmatrix} \nabla F_1 \\ \nabla F_2 \\ \nabla F_3 \end{bmatrix} \\
\nabla\Psi &= -(\Psi'\otimes(I+F_3)^{-1})\,\nabla F_3 - (\Phi'\otimes(I+F_3)^{-1})\,\nabla F_1 - (I\otimes(I+F_3)^{-1})\,\nabla F_2 \qquad (B6) \\
\nabla A_{3j} &= (A_{22}'\otimes I)\,\nabla\Psi + (I\otimes\Psi)\,\nabla A_{2j}.
\end{aligned}
$$

The derivatives of ∇A_{3j}, $j = 2,3$, use the fact that $X_{3t} = \Psi X_{2t}$. Thus, some modification of these derivatives would be required in the case that equilibrium conditions took another form than $u_t = X_{3t}$.

Given ∇A^o, ∇C, ∇D, $\nabla\Sigma$, $\nabla\Omega$, we can then derive $\partial L/\partial\Gamma_i$.[15] In the formulas provided we require the following sums: $S_{aa} = T^{-1}\sum_{t=1}^{T} a_t a_t'$, $S_{\hat{X}a} = T^{-1}\sum_{t=1}^{T}\hat{X}_t a_t'$,

[15] Note that A^o can be obtained for the large system with state vector X_t or the small system with state vector X_{2t}. For both we require gradients for all partitions of F.

$S_{\hat{X}\varphi} = T^{-1}\sum_{t=1}^{T-1}\hat{X}_t\varphi_{t+1}'$, $S_{a\varphi} = T^{-1}\sum_{t=1}^{T-1}a_t\varphi_{t+1}'$, $S_{Za} = T^{-1}\sum_{t=1}^{T}Z_t a_t'$, and $S_{Z\varphi} = T^{-1}\sum_{t=1}^{T-1}Z_t\varphi_{t+1}'$ where φ_t is generated by

$$\begin{aligned}\varphi_t &= (A^o - \mathcal{K}\bar{C})'\varphi_{t+1} + \bar{C}\Sigma_z^{-1}a_t \\ \varphi_T &= \bar{C}\Sigma_z^{-1}a_T\end{aligned} \tag{B7}$$

and the remaining variables are as defined in Section 4. In addition, to simplify notation, we define M_1, M_2, M_3 as follows:

$$\begin{aligned}M_1 &= \Sigma_z^{-1}(I - S_{aa}\Sigma_z^{-1}) \\ M_2 &= \bar{C}'M_1\bar{C} - 2\Sigma_z^{-1}S_{a\varphi}(A^o - \mathcal{K}\bar{C}) \\ M_3 &= (A^o - \mathcal{K}\bar{C})'M_3(A^o - \mathcal{K}\bar{C}) + .5(M_2 + M_2').\end{aligned} \tag{B8}$$

Given these definitions and the those of Section 4, we obtain the following expression for the derivative of L with respect to some parameter Γ_i:

$$\begin{aligned}\frac{\partial L}{\partial \Gamma_i} = {}& 2\mathrm{trace}\Big[\frac{\partial A^o}{\partial \Gamma_i}\Big\{(-S_{\hat{X}\varphi} - S\Sigma_z^{-1}\bar{C}'S_{a\varphi})(I - \mathcal{K}C) + S\bar{C}'M_1C - (S_{\hat{X}a} + S(A^o - \mathcal{K}\bar{C})'S_{a\varphi}') \\ & \quad \Sigma_z^{-1}C + S(A^o - \mathcal{K}\bar{C})'M_3 - SAo'M_3\mathcal{K}C\Big\}\Big] \\ & + \mathrm{trace}\Big[\frac{\partial \Sigma}{\partial \Gamma_i}\Big\{C'(M_1 + \mathcal{K}'M_3\mathcal{K})C - 2C'\Sigma_z^{-1}S_{a\varphi}(I - \mathcal{K}C) + M_3 - 2C'\mathcal{K}'M_3\Big\}\Big] \\ & + 2\mathrm{trace}\Big[\frac{\partial C'}{\partial \Gamma_i}\Big\{\Sigma C'(M_1 + \mathcal{K}'M_3\mathcal{K}) + (A^oS\bar{C}' + \Sigma C')\Sigma_z^{-1}S_{a\varphi}\mathcal{K} - A^oS(A^o - \mathcal{K}\bar{C})' \\ & \quad \cdot(S_{a\varphi}'\Sigma_z^{-1} - S_{a\varphi}'\Sigma_z^{-1}D + M_3\mathcal{K}) - \Sigma(I - C'K')S_{a\varphi}'\Sigma_z^{-1} \\ & \quad - S\bar{C}'(M_1 + \Sigma_z^{-1}S_{a\varphi}\mathcal{K})D + A^oS\bar{C}'M_1 - A^oS_{\hat{X}a} + \Sigma_z^{-1} + S_{\hat{X}a}\Sigma_z^{-1}D \\ & \quad - S_{\hat{X}\varphi}\mathcal{K}D + A^oS_{\hat{X}\varphi}\mathcal{K} - \Sigma M_3\mathcal{K} + S(A^o - \mathcal{K}\bar{C})'M_3\mathcal{K}D\Big\}\Big] \\ & + \mathrm{trace}\Big[\frac{\partial \Omega}{\partial \Gamma_i}\Big\{2\Sigma_z^{-1}S_{a\varphi}\mathcal{K} + M_1 + \mathcal{K}'M_3\mathcal{K}\Big\}\Big] \\ & + 2\mathrm{trace}\Big[\frac{\partial D}{\partial \Gamma_i}\Big\{-CS\bar{C}'(M_1 + \mathcal{K}'M_3\mathcal{K} + \Sigma_z^{-1}S_{a\varphi}\mathcal{K}) - (S_{Za} - CS_{\hat{X}a})\Sigma_z^{-1} \\ & \quad - CS_{\hat{X}\varphi}\mathcal{K} + CS((A^o - \mathcal{K}\bar{C})'S_{a\varphi}'\Sigma_z^{-1} + A^{o\prime}M_3\mathcal{K}) + S_{Z\varphi}\mathcal{K}\Big\}\Big] \\ & - 2\mathrm{trace}\Big[\frac{\partial \bar{X}}{\partial \Gamma_i}\varphi(t_0)/T\Big].\end{aligned} \tag{B9}$$

The computations of the gradients for matrices corresponding to the control are extensions of Zadrozny (1988) and the gradient of the likelihood extends the work of Wilson and Kumar (1982). Zadrozny (1988) did not assume any distortions were present and Wilson and Kumar (1982) did not assume growth in the observables, a nonzero or parameter-dependent initial state, or serially correlated measurement errors.

References

[1] Altug, Sumru (1989), "Time to Build and Aggregate Fluctuations: Some New Evidence," *International Economic Review*, 30:889-920.

[2] Barro, Robert J. and Chaipat Sahasakul (1983), "Average Marginal Tax Rates from Social Security and the Individual Income Tax" *Journal of Business*, 59:555-566.

[3] Becker, Robert A. (1985), "Capital Income Taxation and Perfect Foresight," *Journal of Public Economics*, 26:147-167.

[4] Benhabib, Jess, Richard Rogerson, and Randall Wright, (1990) "Homework in Macroeconomics", NBER working paper #3344.

[5] Braun, R. Anton (1989), "The Dynamic Interaction of Distortionary Taxes and Aggregate Variables in Postwar U.S. Data", mimeo, University of Virginia.

[6] Cassou, Steven (1990), "Welfare Costs of Stabilization", mimeo, State University of New York at Stony Brook.

[7] Christiano, Lawrence J. (1988), "Why Does Inventory Investment Fluctuate So Much," *Journal of Monetary Economics*, 21:247-280.

[8] Christiano, Lawrence J. and Martin Eichenbaum (1988), "Is Theory Really Ahead of Measurement? Current Real Business Cycle Theories and Aggregate Labor Market Fluctuations," mimeo, Federal Reserve Bank of Minneapolis and Northwestern University.

[9] Dhrymes, Phoebus J. (1970), *Econometrics: Statistical Foundations and Applications*, New York: Harper and Row.

[10] Driffill, John E. and Harvey S. Rosen (1983), "Taxation and Excess Burden: A Life Cycle Perspective," *International Economic Review*, 24:671-683.

[11] Eichenbaum, Martin S., Lars Peter Hansen, Kenneth J. Singleton (1988) "A Time Series Analysis of Representative Agent Models of Consumption and Leisure Choice Under Uncertainty," *Quarterly Journal of Economics*, 51-78.

[12] Greenwood, Jeremy and Gregory Huffman (1989), "Tax Analysis in a Dynamic Stochastic Model: On Measuring Harberger Triangles and Okun Gaps", mimeo, University of Western Ontario.

[13] Hansen, Gary D. (1985), "Indivisible Labor and the Business Cycle," *Journal of Monetary Economics*, 16:309-27.

[14] Harvey, A. C. (1981), *Time Series Models*, New York: Halsted.

[15] Joines, Douglas H. (1981), "Estimates of Effective Marginal Tax Rates on Factor

Incomes", *Journal of Business*, 54:191-226.

[16] Judd, Kenneth L. (1987), "The Welfare Cost of Factor Taxation in a Perfect-Foresight Model," *Journal of Political Economy*, 95:675-709.

[17] Kydland, Finn E. and Edward C. Prescott (1982), "Time to Build and Aggregate Fluctuations," *Econometrica*, 50:1345-1370.

[18] Long, John B. and Charles Plosser (1983), "Real Business Cycles," *Journal of Political Economy*, 91:39-69.

[19] McGrattan, Ellen R. (1989), "Computation and Application of Equilibrium Models with Distortionary Taxation," Ph.D. Thesis, Stanford University.

[20] McGrattan, Ellen R. (1990), "Computation of Equilibria in Linear-Quadratic Models with Distortions or Externalities," mimeo, Duke University.

[21] Prescott, Edward C. (1986), "Theory Ahead of Business Cycle Measurement," *Federal Reserve Bank of Minneapolis Quarterly Review*, 10:9-22.

[22] Sargent, Thomas, J. (1989), "Two Models of Measurement and the Investment Accelerator," *Journal of Political Economy*, 97:251-287.

[23] Seater, John J. (1982), "Marginal Federal Personal and Corporate Income Tax Rates in the U.S., 1909-1975," *Journal of Monetary Economics*, 10:361-381.

[24] Seater, John J. (1985), "On the Construction of Marginal Federal Personal and Social Security Tax Rates in the U.S.," *Journal of Monetary Economics*, 15:121-135.

[25] Sims, Christopher A. (1980), "Macroeconomics and Reality," *Econometrica*, 48:1-48.

[26] Taylor, John B. (1982) "The Swedish Investment Funds System as a Stabilization Policy Rule," *Brookings Papers on Economic Activity*, 57-99.

[27] Taylor, John B. and Harold Uhlig (1990), "Solving Nonlinear Stochastic Growth Models: A Comparison of Alternative Solution Methods," *Journal of Business and Economic Statistics*, 8:1-18.

[28] Vaughan, David R. (1970), "A Nonrecursive Algebraic Solution for the Discrete Riccati Equation," *IEEE Transactions on Automatic Control*, vol. AC-15, pp. 597-599, 1970.

[29] Wilson, David A. and Alok Kumar (1982), "Derivative Computations for the Log Likelihood Function," *IEEE Transactions on Automatic Control*, AC-27:230-232.

[30] Zadrozny, Peter (1988), "Analytic Derivatives for Estimation of Discrete-Time, Linear Quadratic, Dynamic, Optimization Models," *Econometrica*, 56:467-472.