ELSEVIER Journal of Monetary Economics 38 (1996) 139–160

JOURNAL OF Monetary ECONOMICS

Overtime, effort, and the propagation of business cycle shocks

George J. Hall

Economic Research Department, Federal Reserve Bank of Chicago, Chicago, IL 60604, USA

(Received December 1994; final version received March 1996)

Abstract

This paper presents and estimates a variant of Hansen and Sargent's (1988) real business cycle model with straight time and overtime. The model presented has only one latent variable, the state of technology, yet it does a better job propagating and magnifying shocks than the labor hoarding models which incorporate unobserved effort. This model, as well as a version of Burnside, Eichenbaum, and Rebelo's (1993) labor hoarding model, is estimated using maximum likelihood. The maximum likelihood parameter estimates are compared to those using GMM.

Key words: Business cycle fluctuations; Dynamic general equilibrium; Shock propagation; Maximum likelihood

JEL classification: E32; C51

1. Introduction

This paper estimates a dynamic general equilibrium real business cycle model of the U.S. economy incorporating straight time and overtime. This model is a hybrid of Hansen and Sargent's (1988) model with straight time and overtime and Burnside, Eichenbaum, and Rebelo's (1993) labor hoarding model. A version of Burnside, Eichenbaum, and Rebelo's model is also studied. These two

I thank John Cochrane, Charles Evans, Robert Lucas, David Marshall, Ellen McGrattan, Ned Prescott, Argia Sbordone, Thomas Tallarini, and Mark Watson for helpful conversations. I thank Martin Eichenbaum, Lars Hansen, and Thomas Sargent for their encouragement, support, and comments. This paper has benefited from comments made by an anonymous referee and seminar participants at the University of Chicago and the Federal Reserve Banks of Chicago and New York. All errors are mine.

PII S0304-3932(96)01270-6

models are analyzed to study how different assumptions about labor market rigidities affect the real business cycle model's ability to propagate and magnify shocks. This paper concludes that although the model presented has only a single latent variable, the state of technology, it magnifies and propagates business cycle shocks better than Burnside, Eichenbaum, and Rebelo's (1993) labor hoarding model which incorporates the unobservable variable effort.

A well-known and often repeated criticism of standard real business cycle (RBC) models is their dependence on implausibly large aggregate technology shocks to account for the variability in output.[1] Another way of phrasing this criticism is that standard RBC models possess weak propagation mechanisms. That is, if these models could plausibly magnify and spread over time the effect of the shocks, they would be less dependent on large aggregate shocks to produce sufficiently volatile time series. Consequently a challenge to business cycle theorists is to construct models that amplify shocks and realistically propagate shocks over the business cycle.

In response to this challenge, Burnside and Eichenbaum (1995) study two potential channels of propagation: time-varying effort and time-varying capital utilization. Their model has the ability to substantially magnify and propagate shocks. One of the two models studied here is a special case of their model; in this model capital utilization is fixed to focus attention on the implications for shock transmission inherent in the time-varying effort assumption. Thus, this model is just Burnside, Eichenbaum, and Rebelo's (1993) labor hoarding model although the issues addressed are from Burnside and Eichenbaum (1995). This paper studies in detail the propagation mechanism of the labor hoarding channel and proposes a complementary channel – the differentiation between straight time and overtime.

The model presented here builds on the model of Hansen and Sargent (1988). As in their model, the model presented here incorporates a version of Lucas's (1970) instantaneous production function in which straight time and overtime are not perfect substitutes. However, it differs from Hansen and Sargent's (1988) model in two ways. First, a government sector and a government shock are added to the model. Second, firms must commit to the number of workers they will employ before observing any shocks to the economy; once the shocks are observed, firms can adjust the number of employees working overtime. In other words, the cost of adjusting the contemporaneous number of workers after observing the shocks is infinite.

The model presented can also be viewed as a variation on the model of Burnside, Eichenbaum, and Rebelo (1993), henceforth BER. In both models, firms must commit to the number of workers employed before observing any shocks to the economy. In BER's model, firms can adjust the work effort after observing the shocks. In the model presented here, firms can adjust the number

[1] See, for example, Cochrane (1994) and Cogley and Nason (1995).

of persons working overtime after observing the shocks. The advantage of the model presented here is that unlike effort which cannot be measured, overtime employment is observed and measured.

Using economically interpretable parameter values obtained by maximum likelihood, the model presented here delivers greater magnification and propagation of shocks over the business cycle than BER's labor hoarding model. This success is achieved under the constraint that all variables in this model, with the exception of technology, are observed. Moreover, the model does as well as BER's model in matching the first- and second-moment properties of the data. This result challenges the quantitative importance of unobserved effort as a propagation and amplification mechanism.

The remainder of this paper is organized as follows. In Section 2, the model is presented. In the third section, the solution and estimation procedures are discussed. In Section 4, empirical results are calculated and assessed. In the final section, some concluding remarks are made.

2. The economy

Consider an economy with a continuum on [0,1] of identical infinitely lived agents who have preferences over consumption of a single nondurable good, C_t, and leisure, L_t. They maximize:

$$\mathrm{E}\sum_{t=0}^{\infty}\beta^t[\log C_t + v\log L_t], \tag{1}$$

where $0 < \beta < 1$ is their subjective discount factor and v is strictly greater than zero.

Agents are endowed with T units of time each period which can be divided between labor and leisure; consequently $0 < L_t \leq T$. Furthermore L_t is restricted to take only one of three values. Let h_1 and h_2 be the lengths of a 'straight time shift' and an 'overtime shift', respectively. Let $0 < h_1 < h_1 + h_2 < T$. Therefore:

$$L_t = \begin{cases} T & \text{if the agent is unemployed,} \\ T - h_1 & \text{if the agent works only the straight time shift,} \\ T - h_1 - h_2 & \text{if the agent works both the straight time and overtime shifts.} \end{cases}$$

Proceeding as in Hansen (1985) and Rogerson (1988) lotteries are employed to convexify the commodity space. Assume π_{1t} and π_{2t} are the probability of working just the straight time shift and the probability of working both shifts, respectively. Hence $1 - \pi_{1t} - \pi_{2t}$ is the probability of being unemployed. An agent's expected single period utility is then:

$$\pi_{1t}[\log C_t + v\log(T - h_1)] + \pi_{2t}[\log C_t + v\log(T - h_1 - h_2)]$$
$$+(1 - \pi_{1t} - \pi_{2t})[\log C_t + v\log(T)]. \tag{2}$$

Define N_{2t} to be the fraction of agents who work both shifts (overtime employment) and N_{1t} to be the fraction of agents who work the straight time shift (total employment). So N_{1t} equals $\pi_{1t} + \pi_{2t}$, N_{2t} equals π_{2t}, and the agent's utility function, (1), can be written

$$\mathrm{E}\sum_{t=0}^{\infty} \beta^t[\log C_t - a_1(N_{1t} - N_{2t}) - a_2N_{2t} - a_0(1 - N_{1t})], \tag{3}$$

where $a_0 = -v\log(T)$, $a_1 = -v\log(T - h_1)$, and $a_2 = -v\log(T - h_1 - h_2)$. This preference specification is from Hansen and Sargent (1988).

Agents have access to an instantaneous Cobb–Douglas production technology such that total output produced over the two shifts during period t is

$$Q_t = \exp(\lambda_t)(\gamma^t)^\alpha K_t^{1-\alpha}[h_1N_{1t}^\alpha + h_2N_{2t}^\alpha]. \tag{4}$$

Here K_t is the capital stock, λ_t is the state of technology, and γ is the growth rate of exogenous labor-augmenting technological progress. Labor's share, α, is restricted such that $0 < \alpha < 1$.

Total output is allocated each period to private consumption, C_t, government consumption, G_t, and investment, I_t,

$$C_t + G_t + I_t \leq Q_t. \tag{5}$$

Productive capital depreciates each period at the rate $0 < \delta \leq 1$, so

$$K_{t+1} = (1 - \delta)K_t + I_t. \tag{6}$$

Technology has the following law of motion:

$$\lambda_{t+1} = \mu_\lambda + \rho_\lambda\lambda_t + \sigma_\lambda w_{\lambda t+1}, \tag{7}$$

where $0 < \rho_\lambda < 1$ and $\{w_{\lambda t}\}$ is a sequence of i.i.d. normal random variables with mean zero and variance one.

Government consumption evolves according to

$$g_{t+1} = \mu_g + \rho_g g_t + \sigma_g w_{gt+1}, \tag{8}$$

where $g_t = \log(G_t/\gamma^t)$ and $0 < \rho_g < 1$. $\{w_{gt}\}$ is a sequence of i.i.d. normal random variables with mean zero and variance one and is orthogonal to innovations in technology.

The timing of this economy differs from the one presented in Hansen and Sargent (1988). In this model N_{1t} must be chosen before, instead of after, (λ_t, G_t)

are known; N_{2t} and K_{t+1} are chosen after observing the shocks. Formally, let the initial information set, $\mathscr{I}_0$, be generated by the set of initial conditions, $\{\lambda_0, G_0, K_0, N_{10}\}$. Let the information set $\mathscr{I}_t$ be generated by $\{\lambda_0, G_0, K_0, N_{10}\}$ and $\{(w_{\lambda s}, w_{gs}): s = 1,2,\ldots,t\}$, so $\mathscr{I}_t$ consists of all measurable functions of $\{\lambda_0, G_0, K_0, N_{10}\} \cup \{(w_{\lambda s}, w_{gs}): s = 1,2,\ldots,t\}$. The social planner's choice of $\{K_{t+1}, N_{1t+1}, N_{2t}\}$ is a function of the elements of $\mathscr{I}_t$.

If markets are complete, the decentralized competitive equilibrium corresponds to the solution of a social planning problem. In this case, the planning problem is to choose a set of stochastic processes $\{K_{t+1}, N_{1t+1}, N_{2t}\}_{t=0}^{\infty}$ to maximize (3) subject to (4)–(8) given initial conditions $\{\lambda_0, G_0, K_0, N_{10}\}$. The set $\{K_{t+1}, N_{1t+1}, N_{2t}\}$ is required to be a function of the elements of $\mathscr{I}_t$. This definition of the social planning problem completes the description of the model; this model is referred to as the overtime model.

For comparison purposes, a version of Burnside, Eichenbaum, and Rebelo's (1993) labor hoarding model is presented. In their model, firms commit to the number of workers employed before observing any shocks to the economy; after observing the shocks, firms can adjust the work effort of their employees. Higher work effort increases output but lowers the agents' utility. Formally the planner chooses a set of stochastic processes $\{K_{t+1}, N_{1t+1}, e_t\}_{t=0}^{\infty}$ to maximize

$$\mathrm{E}\sum_{t=0}^{\infty} \beta^t \big(N_{1t}[\log C_t + \upsilon \log(T - \xi - e_t h_1)] + (1 - N_{1t})[\log C_t + \upsilon \log(T)] \big), \tag{9}$$

subject to

$$Q_t = \exp(\lambda_t)(\gamma^t)^{\alpha} K_t^{1-\alpha} h_1 [e_t N_{1t}]^{\alpha} \tag{10}$$

and (5)–(8). The set $\{K_{t+1}, N_{1t+1}, e_t\}$ is restricted to be a function of the elements of $\mathscr{I}_t$. In this model N_{1t} is the fraction of agents who work a single shift of length h_1. Two additional variables are introduced in BER's model. The date t level of employee work effort is denoted by e_t; ξ is a fixed cost of going to work. The only difference between this model and the one presented in BER is that the production function in (10) is instantaneous. This model is referred to as the effort model.

3. Approximation, estimation, and data

Since it is not possible to obtain an analytic solution to either of these models, Christiano's (1988) log-linear-quadratic approximation is used to obtain an approximate solution to the social planner's problem.[2] In the overtime model the

[2] The effort model is solved in a similar fashion.

endogenous variables do not converge to a nonstochastic steady state but to a steady state growth path. Therefore the economy is transformed so that it has a steady state. Let $q_t = \log(Q_t/\gamma^t)$, $c_t = \log(C_t/\gamma^t)$, $i_t = \log(I_t/\gamma^t)$, $k_t = \log(K_t/\gamma^t)$, $n_{1t} = \log(N_{1t})$, and $n_{2t} = \log(N_{2t})$. Define the state vector, $x_t = (1, \lambda_t, g_t, k_t, n_{1t})$, the control vector, $u_t = (n_{1t+1}, n_{2t}, k_{t+1})$, and the shock vector, $w_t = (w_{\lambda t}, w_{gt})$. Let $\theta = (\alpha, \beta, \delta, \gamma, \mu_\lambda, \mu_g, \rho_\lambda, \rho_g, h_1, h_2, T, v, \sigma_\lambda, \sigma_g)$ be a vector of parameters. Finally, let $z_t = [x_t', u_t']$.

The approximation proceeds as follows. First, Eqs. (3)–(8) are written in terms of the lower-case variables. Second, constraints (4)–(6) are substituted into the objective function. Third, a second-order Taylor expansion of the objective function around the steady state of z_t is taken; this yields a log-linear quadratic version of the planner's problem. The solution to this approximate planner's problem generates time-invariant decision rules that express u_t as a linear function of x_t. This procedure implies the mapping:

$$x_{t+1} = \mathscr{A}_0 x_t + \mathscr{C} w_{t+1}, \tag{11}$$

where the elements of $\mathscr{A}_0$ are nonlinear functions of θ.

Let y_t denote a vector of observer variables: $y_t = [q_t\ c_t\ i_t\ g_t\ k_t\ h_t\ n_{1t}\ n_{2t}]'$. Let Y_t denote the corresponding vector of untransformed (upper-case) observer variables. All observer variables are written as linear functions of the state variables

$$y_t = \mathscr{G} x_t. \tag{12}$$

Total hours, h_t, is defined such that $\exp(h_t) = H_t \equiv h_1 N_{1t} + h_2 N_{2t}$. Since q_t, c_t, i_t, and h_t are nonlinear functions of the state variables, Taylor approximations are used. The remaining four variables of y_t are in z_t.

If measurement error is added to Eq. (12), then the linear system (11) and (12) can be used to compute a Gaussian likelihood function. Economically sensible parameter values for the entire vector θ that maximize the likelihood function can then be deduced from the data. If the form of the distribution of the errors is specified, the maximum likelihood (ML) parameter estimates are the most efficient estimates obtainable. While other estimation methods may use just a select set of moments to compute parameter values that match the data, ML incorporates all the information in the model by working with the complete distribution. This estimation procedure is discussed in more detail in Anderson, Hansen, McGrattan, and Sargent (1995).

Assume the vector y_t is observed with measurement error. Replace (12) with

$$y_t = \mathscr{G} x_t + v_t, \tag{13}$$

where v_t is measurement error such that $v_t = \mathscr{D} v_{t-1} + \eta_t$, $E\eta_t = 0$, $E\eta_t\eta_t' = \mathscr{R}$, and $E\eta_t v_t = 0$. Assume $\mathscr{R}$ is a diagonal matrix. Define the following: let $\bar{y}_t \equiv y_{t+1} - \mathscr{D} y_t$ and $\bar{\mathscr{G}} \equiv \mathscr{G}\mathscr{A}_0 - \mathscr{D}\mathscr{G}$; let $\hat{E}[Y|X]$ denote the linear least squares projection of Y onto X; finally, let $\mathscr{K}$ and Σ denote the steady state 'Kalman gain' and 'state-covariance matrix' of the time-invariant Kalman filter. Applying the time-invariant Kalman filter to the linear system described above attains the corresponding innovations representation:

$$\hat{x}_{t+1} = \mathscr{A}_0\hat{x}_t + \mathscr{K}u_t,$$
$$\bar{y}_t = \bar{\mathscr{G}}\hat{x}_t + u_t,$$

where $u_t = y_{t+1} - \hat{E}[y_{t+1}|y_t,\ldots,y_1,\hat{x}_1]$, $\hat{x}_t = \hat{E}[x_t|y_t,\ldots,y_1,\hat{x}_1]$, and $Eu_tu_t' \equiv \Omega = \bar{\mathscr{G}}\Sigma\bar{\mathscr{G}}' + \mathscr{R} + \mathscr{G}\mathscr{C}\mathscr{C}'\bar{\mathscr{G}}'$. Hence the Gaussian log-likelihood function for $\{y_t\}_{t=1}^{\mathscr{T}}$ conditioned on $\hat{x}_1$ is

$$\log L(\theta) = -(\mathscr{T} - 1)\log 2\pi p - \tfrac{1}{2}(\mathscr{T} - 1)\log|\Omega| - \tfrac{1}{2}\sum_{t=1}^{\mathscr{T}-1} u_t'\Omega^{-1}u_t, \qquad (14)$$

where p is the number of series in y_t.

This formulation of the log-likelihood function is predicated on the assumption that the initial condition, $\hat{x}_1$, is estimated from an infinite history of y's: $\hat{x}_1 = \hat{E}[x_1|y_0, y_{-1},\ldots]$. In this study, $\hat{E}[x_1|y_0, y_{-1},\ldots]$ is approximated by $\hat{E}[x_1|y_0, y_{-1}, x_{-1}]$ where x_{-1} is set to a vector whose elements are: the number one, the unconditional mean (steady-state level) of technology, and the beginning-of-sample values for g, k, and n_1. This initialization implies that the parameter values reported in Section 4.1 maximize an approximation to (14).

The data used in this analysis are quarterly, seasonally adjusted, aggregate real data of the United States for the sample 1955:Q1 to 1992:Q4.[3] In order to make the data consistent with the lower-case variables in the model, all data are converted to per capita terms using the civilian, noninstitutional population, 16 years and older. Each series is appropriately detrended, and logarithms are taken.

The capital stock series, K_t, is the net stock of fixed private capital plus the net stock of durable goods reported in the *Survey of Current Business*; they are annual data interpolated to quarterly data. Investment, I_t, is private fixed investment plus the personal consumption expenditure on durable goods. Private consumption, C_t, is the sum of the imputed service flow from the stock of consumer durables computed by Brayton and Mauskopf (1985) and the personal consumption expenditures on nondurable goods and services. G_t is government purchases. Unless otherwise noted, these series are from the NIPA. Output, Q_t, is the sum of C_t, G_t, and I_t.

Total employment, N_{1t}, is the number of employed persons at work in nonagricultural industries who worked 35 hours and over a week; overtime employment, N_{2t}, is the number of employed persons at work in nonagricultural industries who worked 41 hours and over a week. Total hours, H_t, is total employment multiplied

[3] Unless otherwise stated, the data were obtained from the Federal Reserve Bank of Chicago's database. Data on total employment and overtime employment from 1955 to 1976 were provided by Gary Hansen.

by the average weekly hours worked in nonagricultural industries.[4] These series are from the BLS's household survey and reported in *Employment and Earnings*. These seasonally adjusted monthly data are averaged when aggregating to quarterly data.

4. Empirical results

This section assesses the quantitative implications of the two models. First, the parameter values are computed by maximum likelihood; these parameter values are compared to estimates computed using alternative estimation strategies. Second, the ability of the models, evaluated at the computed parameter values, to match the first and second moments of the data is measured. Third, the propagation mechanisms embodied within each model are studied along the dimensions discussed in Burnside and Eichenbaum (1995).

4.1. Parameter values

The estimated models are identical to the theoretical models described above with one exception; in the estimated models, government, g_t, is allowed to grow at a different rate than the other series. Its trend is denoted γ_g. To facilitate comparisons across the two models, three parameters are fixed prior to estimation; γ is set to 1.0044, γ_g is set to 1.0007, and T is set to 1369 hours per quarter. The value for γ is chosen by separately regressing q_t, c_t, i_t, and k_t on a constant and a linear time trend. The coefficients on the time trend are constrained to be equal across all four regressions. Likewise γ_g is computed by regressing g_t on a constant and a linear time trend. The value for the time endowment, T, corresponds to 15 hours per day.

Several assumptions about the measurement error process are also made prior to the estimation. At first the matrix $\mathscr{D}$ was assumed to be diagonal; however, to avoid estimating a unit root in the autoregressive coefficient for the measurement error of capital, the following assumption from Christiano (1988) is made. It is assumed that the original capital stock is measured without error; but since investment is measured with error, the measurement error of capital is a weighted sum of the past measurement errors of investment. This implies:

$$v_{t+1}^k = \frac{1-\delta}{\gamma} v_t^k + \frac{\gamma - 1 + \delta}{\gamma} v_t^i, \tag{15}$$

[4] There is an inconsistency in the construction of H_t; the average weekly hours series includes persons working fewer than 35 hours a week, while N_{1t} excludes these part-time workers. Ideally, the average weekly hours series should be computed excluding all workers who work fewer than 35 hours a week. Unfortunately, the data necessary to construct a series of average hours worked conditional on working 35 hours or more a week are only available back to 1976.

Table 1
Measurement error parameter estimates

Parameter	Effort model		Overtime model	
	Value	S.E.	Value	S.E.
$\mathscr{D}(1,1)$	0		0	
$\mathscr{D}(2,2)$	0.999		0.999	
$\mathscr{D}(3,3)$	0.980	0.011	0.977	0.008
$\mathscr{D}(4,4)$	0.980	0.024	0.983	0.024
$\mathscr{D}(6,6)$	0.550	0.157	0.757	0.204
$\mathscr{D}(7,7)$	0.576	0.166	0.821	0.170
$\mathscr{D}(8,8)$			0.946	0.037
$\sqrt{\mathscr{R}(1,1)}$	0		0	
$\sqrt{\mathscr{R}(2,2)}$	0.0036	0.0006	0.0033	0.0005
$\sqrt{\mathscr{R}(3,3)}$	0.0109	0.0014	0.0104	0.0014
$\sqrt{\mathscr{R}(4,4)}$	0.0064	0.0020	0.0064	0.0017
$\sqrt{\mathscr{R}(5,5)}$	0		0	
$\sqrt{\mathscr{R}(6,6)}$	0.0212	0.0049	0.0208	0.0065
$\sqrt{\mathscr{R}(7,7)}$	0.0185	0.0047	0.0179	0.0052
$\sqrt{\mathscr{R}(8,8)}$			0.0209	0.0020

Series ordered: q_t, c_t, i_t, g_t, k_t, h_t, n_{1t}, n_{2t}

where v_t^k and v_t^i are the time t measurement errors on capital and investment, respectively.[5] Unfortunately, this assumption causes the approximated likelihood function to reach a constrained maximum when the autoregressive coefficient of the measurement error process for consumption, $\mathscr{D}(2,2)$, is one. Therefore $\mathscr{D}(2,2)$ is set to 0.999. Output is assumed to be measured without error, and $\mathscr{D}(1,1)$ is set to 0.[6] To ensure that the estimated variances are positive, the square root of the diagonal terms of $\mathscr{R}$ are estimated and reported. See Table 1.

Table 2 presents the maximum likelihood point estimates and standard errors of the preference and technology parameters for the effort and the overtime models.[7] Several of the point estimates for the effort model are virtually identical to those found by Burnside, Eichenbaum, and Rebelo [BER] (1993). For

[5] This assumption implies that the model is singular. Consequently, the standard deviation of the innovations to the measurement error on capital, $\sqrt{\mathscr{R}(5,5)}$, is set to 1×10^{-5} rather than exactly zero. This assumption also implies that the standard errors on δ are implausibly small.

[6] This assumption corresponds to the Heywood solution in the factor analysis literature. See Sargent and Sims (1977) and Altug (1989). For numerical reasons, $\sqrt{\mathscr{R}(1,1)}$ is set to 1×10^{-5} rather than exactly zero.

[7] Standard errors are also reported for all the estimates with the exception of δ; δ is free during estimation but fixed when computing the standard errors. This was done to avoid inverting a near singular matrix when computing the standard errors.

Table 2
Preference and technology parameter estimates

	Effort model		Overtime model	
Parameter	Value	S.E.	Value	S.E.
α	0.656	0.093	0.689	0.027
β	0.983	0.012	0.988	0.003
δ	0.0225		0.0225	
μ_λ	−0.0145	0.0309	−0.0149	0.0112
μ_g	0.0886	0.0858	0.0887	0.0859
ρ_λ	0.979	0.006	0.973	0.005
ρ_g	0.987	0.012	0.987	0.012
h_1	511.4	10.8	451.1	48.5
h_2			158.9	47.5
v	7.004	5.706	6.670	0.609
ξ	93.56	110.20		
σ_λ	0.0057	0.0006	0.0060	0.0004
σ_g	0.0089	0.0014	0.0092	0.0013

example, in the effort model, the estimate of labor's share in the production function is 0.656; BER estimate labor's share to be 0.654.[8] The estimate for ρ_λ in the effort model is 0.979, while BER compute 0.982. Likewise the point estimate for ρ_g is 0.987; this is identical to BER's estimate. This replication of their results is discussed further in the following subsection.

In the effort model three parameters are estimated that are fixed by BER. First is the discount factor, β. The point estimate, 0.983, implies a high annual risk-free interest rate of 8.7%. Nevertheless, the point estimate is less than one and within one standard error of the values usually assumed. Second is the shift length, h_1. BER fix the shift length so steady-state effort is one. In the model presented here, the point estimate is economically interpretable ($h_1 = 511.4$) and corresponds to 39.3 hours per week. The third parameter is the fixed cost to working, ξ. BER state that their results are insensitive to choices of ξ between 20 and 120. These results support that statement; the point estimate for ξ is 93.56 with a standard error of 110.20.

The parameter estimates for the overtime model are similar to those in the effort model. The estimate of labor's share is 0.689, a slightly higher number than often used. The point estimate of β is 0.988 – the same number King, Plosser, and Rebelo (1988) use. Both models estimate δ to be 0.0225. The point estimates for the shift parameters h_1 and h_2 are 451.1 and 158.9, respectively. These point estimates correspond to a straight time shift of 34.7 hours a week and an overtime shift of 12.2 hours a week. Since the definition of total employment

[8] The BER parameter estimates are for their Labor Hoarding I – Overidentified Model which are presented on p. 256 of their paper.

excludes persons working less than 35 hours a week, the point estimate of h_1 is a little troubling. Not too much should be made of the point estimate however; its standard error is large, 48.5.

The bottom two rows of Table 2 report the parameter estimates for σ_λ and σ_g, the standard deviation of the technology innovations and government innovations, respectively. For both models, the standard deviation of the technology innovations is estimated to be about 0.006; likewise, for both models, the standard deviation of the government innovations is estimated to be about 0.009. Both models match the data with shocks of the same volatility.

The standard errors reported here suggest that both models have difficulty matching the hours and employment series. On the one hand, the effort model provides a tight parameter estimate of the shift length; the standard error on h_1 is 10.8. On the other hand, there are very large standard errors associated with the preference parameters v and ξ. In contrast, applying ML to the overtime model yields estimates for the shift lengths with large standard errors but a tight estimate on v.

This paper as well as studies such as McGrattan (1994) and McGrattan, Rogerson, and Wright (1995) are examples in which maximum likelihood estimation of a stochastic growth model give sensible point estimates to the economically interpretable parameters. These examples are in contrast to some previous studies which had trouble computing economically interpretable parameter estimates using maximum likelihood. For example Altug (1989) and Christiano (1988) must fix β to avoid a point estimate greater than one.

4.2. Maximum likelihood, GMM, and the effort model

The estimation above raises the following question: why are several of the parameter estimates for the effort model similar to those obtained by Burnside, Eichenbaum, and Rebelo (1993) via GMM? The subsection makes two points. First, the means of the data contain most of the information needed to compute sensible point estimates for a subset of θ. Second, if the effort model is specified correctly then both maximum likelihood and BER's estimation strategies are consistent; so it should not be surprising that both methods lead to similar point estimates.

Both BER's estimation strategy and maximum likelihood exploit the information coming from the means of the data. Specifically, BER employ the following moment conditions to estimate α, δ, and v:

$$\mathrm{E}\left[\beta^{-1} - \frac{C_t}{C_{t+1}}\left((1-\alpha)\frac{Q_{t+1}}{K_{t+1}} + (1-\delta)\right)\right] = 0, \tag{16}$$

$$\mathrm{E}\left[\delta - \left(1 + \frac{I_t}{K_t} - \frac{K_{t+1}}{K_t}\right)\right] = 0, \tag{17}$$

$$\mathrm{E}[H_t - H_{ss}] = 0, \tag{18}$$

where H_{ss} denotes the steady state of hours worked. In Eq. (18), the role the means play in BER's estimation strategy is explicit; the value of v is determined by minimizing the difference between the model-determined mean (steady state) of H and the sample mean. Some experimentation with Eqs. (16) and (17) demonstrates that if the growth rates of C_t and K_t are known, sensible parameter estimates for α and δ can be obtained from the means of data. Recall that BER fix β prior to their estimation.

Maximum likelihood exploits all the moment conditions implied by the model. The role the first moments play can be illustrated explicitly if the log-likelihood function, (14), is approximated in the frequency domain. Let $\mu(\theta)$ and $S_y(\theta,\omega_j)$ denote the model-determined mean (steady state) and spectral density matrix. The periodogram of the data is denoted by $I_y(\omega_j)$. The log-likelihood function can then be approximated by

$$\begin{aligned}\log L(\theta) = &-\tfrac{1}{2}(\mathcal{T} + \mathcal{T}p)\log 2\pi - \sum_{j=1}^{\mathcal{T}/2+1} \log|S_y(\theta,\omega_j)| \\ &- \sum_{j=1}^{\mathcal{T}/2+1} \mathrm{trace}[S_y(\theta,\omega_j)^{-1} I_y(\omega_j)] \\ &- \frac{\mathcal{T}}{2}\left\{\left[\frac{1}{\mathcal{T}}\sum_{t=1}^{\mathcal{T}} y_t - \mu(\theta)\right] S_y(\theta,0)^{-1}\left[\frac{1}{\mathcal{T}}\sum_{t=1}^{\mathcal{T}} y_t - \mu(\theta)\right]\right\},\end{aligned} \tag{19}$$

where p is the number of series contained in y. Note that the fourth term is simply the product of a constant $(-\mathcal{T}/2)$ and a GMM criterion where the moment conditions are the differences between the means of the data and the model-determined means; the weighting matrix is the model-determined spectral density matrix at frequency zero.

To measure the importance of the information coming from the first moments of the data, the following is done. First, point estimates for α, β, δ, μ_g, h_1, and v that minimize the GMM criterion in the fourth term of (19) are reported in column of 3 of Table 3.[9] During the minimization the weighting matrix and the remaining parameters in θ are fixed at their maximum likelihood values. The standard errors associated with these GMM point estimates are reported in column 5.[10] Note that these GMM standard errors are not functions of the data; they are functions of just $\partial\mu/\partial\theta$ and $S_y(\theta,0)$. Second, the maximum likelihood point

[9] The effort model implies restrictions on seven data series; however one of the means of these series is redundant (i.e., consumption – since the resource constraint holds by definition). The first-order conditions of the model imply that point estimates for δ, μ_g, h_1, and three of the four parameters, $\{\alpha, \beta, \mu_\lambda, v\}$ can be obtained using just the means of the data if the other parameters are known.

[10] The parameter δ is free during both estimations, but is fixed when computing the standard errors. The asymptotic standard errors for δ are implausibly small.

Table 3
Asymptotic standard errors

Parameter	ML p.e.	GMM p.e.	ML s.e.	GMM s.e.	Hausman
α	0.656	0.657	0.0010	0.0153	0.19
β	0.983	0.982	0.0015	0.0040	0.55
δ	0.0225	0.0229			
μ_g	0.0886	0.0890	0.00070	0.00080	0.76
h_1	511.44	511.36	2.6639	2.6643	3.36
v	7.004	7.043	0.124	0.161	0.14

estimates for these six parameters, originally reported in Table 2, are reproduced in column 2 of Table 3. Asymptotic ML standard errors are computed using a simulated a time series of y_t of length 100,000; these standard errors are reported in column 4. Finally, Hausman specification tests are performed parameter-by-parameter for five of these parameters. These statistics are distributed χ^2 with one degree of freedom and are reported in column 6.

From the results reported in Table 3, it is clear that the point estimates from maximum likelihood and the point estimates from GMM are very similar. Both the asymptotic ML and the asymptotic GMM standard errors, relative to the point estimates, are small. From the GMM estimation above, one can see that the first moments of the data contain enough information to compute sensible estimates for a subset of θ. Note that while BER choose to fix β, a sensible point estimate can be obtained using just the information from the means of the data.

The maximum likelihood estimates are efficient; they incorporate all the information from the data. The GMM estimates incorporate just the information from the means. What is the efficiency loss from ignoring the higher moments in the data? Not much. For h_1 the standard errors are virtually identical. This should not be surprising given how the hours worked series is constructed; it is just the product of average weekly hours worked and total employment. So h_1 in the effort model is just pinned by the ratio of H_t to N_t. For β and v, the ratio of the asymptotic GMM standard errors to the asymptotic ML standard errors is about 1.4. For μ_g this ratio is about 1.15. There does appear to be some efficiency gains to exploiting the higher moments when estimating α; the ratio of its asymptotic GMM standard error to its asymptotic ML standard error is about 2.3.

If the model is specified correctly, then the moments selected to estimate the model should not effect the consistency of the point estimates, only the efficiency. Since both the ML and GMM point estimates are similar, this bodes well for the specification of the model. The null hypothesis that the model is specified correctly cannot be rejected at the 5% level for any of the five parameter-by-parameter Hausman test statistics reported in column 6. It should be noted however that if the ML and GMM standard errors for δ are computed, they imply that δ is estimated very tightly (up to the eighth decimal place); nevertheless,

Table 4
Means and standard deviations of the implied and U.S. time series

The unconditional means of the implied series are their steady-state values; the unconditional standard deviations are the square roots of the diagonal elements of the matrix $\mathscr{G}\text{var}(x)\mathscr{G}'$. The reported moments for the U.S. data are the sample moments from 1955:Q3 to 1992:Q4 for the logged and detrended (lower-case) time series.

	Effort model		Overtime model		United States	
Series	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.
Output, q	8.347	0.039	8.328	0.038	8.331	0.048
Consumption, c	7.743	0.043	7.718	0.039	7.818	0.028
Investment, i	6.783	0.075	6.783	0.081	6.747	0.078
Government, g	6.938	0.056	6.916	0.058	6.967	0.070
Capital, k	10.397	0.042	10.397	0.041	10.347	0.030
Hours, h	5.330	0.023	5.329	0.023	5.341	0.048
Total employment, n_1	−0.907	0.023	−0.909	0.023	−0.896	0.038
Overtime employment, n_2			−1.875	0.024	−1.898	0.098

the GMM and ML point estimates differ in the fourth decimal place. This result suggests that either the law of motion for capital, Eq. (6), is misspecified or the assumption that the measurement error on capital is a weighted sum of the measurement error of investment, Eq. (15), is a poor one.

4.3. *First and second moments*

This subsection assesses the ability of the two models evaluated at the ML parameter estimates to match the first- and second-moment properties of the data. Table 4 reports the unconditional means and standard deviations of the models' implied time series and the actual U.S. data.

Table 4 shows that both models imply similar first- and second-moment properties. Not surprisingly maximum likelihood chooses parameter values such that both models do extremely well matching the means of each series (with the exception of consumption). Both models do less well matching the standard deviations. Since output is assumed to be measured without error, it is reassuring that both models match its mean; however, both models underpredict its standard deviation. Both models do a better job matching the mean and standard deviation of investment than consumption, and both models underpredict the volatility of hours and employment. It is difficult to use the implied second-moment properties of the two models to differentiate between them.

The most dramatic weakness of the overtime model is its inability to explain the relative volatility of total employment to overtime employment. The overtime employment data are over twice as volatile as the total employment data; however, the model predicts overtime employment and total employment to have almost the same standard deviations. In the absence of adjustment costs and the one-period

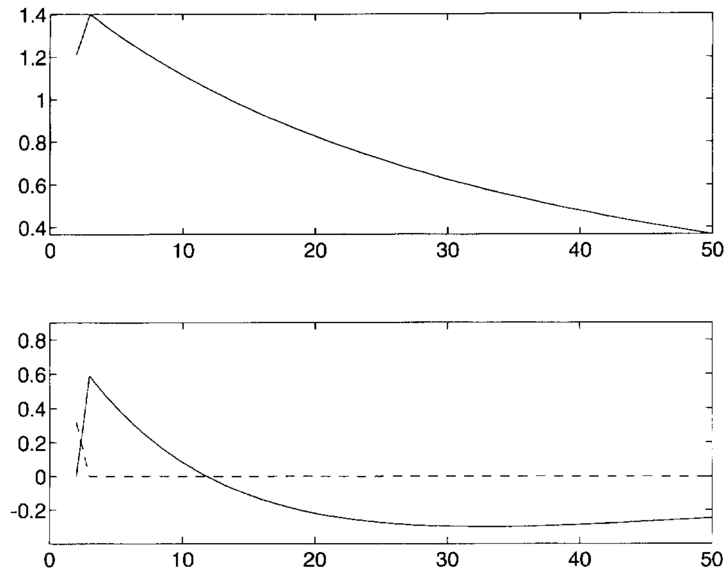

Fig. 1. Effort model – impulse responses from a unit shock to technology. Top – output; bottom – employment (solid) and effort (dashed).

labor rigidity, the overtime model implies that the ratio of n_{1t} to n_{2t} does not vary over the cycle. Augmenting the model with an adjustment cost term to changes in total employment can lead to the implication that overtime employment is over twice as volatile as total employment. Adjustment costs are absent in the overtime model to sharpen the contrast between the two models.

4.4. Propagation and magnification of shocks

To study how shocks are transmitted within the two models, consider the impulse response functions of the linear system described by (11) and (12) for the two models. These impulse responses are plotted in Figs. 1 and 2. These figures plot responses to a positive unit shock to λ. Both models produce 'hump-shaped' impulse response functions.

To understand how the social planner responds to shocks in the effort model, recall that although effort and employment are perfect substitutes in production (see Eq. (10)), they enter the utility function differently. Because of the employment lotteries, utility is linear over employment while curved over effort; thus in equilibrium, it is relatively more costly (in terms of utility) to adjust labor over the intensive margin (effort) than over the extensive margin (employment). In the absence of the one-period labor rigidity, it would be optimal for the planner to set employment so that effort remains at its steady-state level.

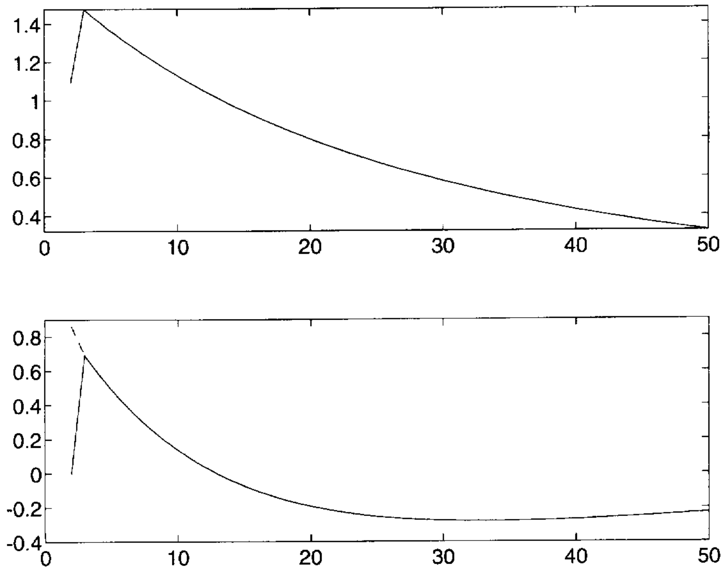

Fig. 2. Overtime model – impulse responses from a unit shock to technology. Top – output; bottom – total employment (solid) and overtime employment (dashed).

Of course, in the effort model the planner must condition $\{K_{1t+1}, N_{1t+1}, e_t\}$ on the date t information set, $\mathscr{I}_t$. When a positive technology shock occurs, the marginal product of labor increases. But the only contemporaneous labor margin the social planner can adjust in response to this increase in productivity is effort. Thus output increases initially from the direct impact of the technology shock and the increased worker effort. See Fig. 1. This shock is persistent; so after observing the shock, the social planner increases future employment such that effort returns its steady-state level. Since, at the steady state, increasing employment reduces an agent's expected utility less than increasing the intensity of work effort, the subsequent deviation of output from its nonstochastic steady-state level is greater than the initial deviation. As the shock decays, the income effect from the increased capital stock dominates the substitution effect from the increased (but declining) marginal product of labor; thus employment declines below its steady-state level before returning to its steady-state level.

In the overtime model, the intuition behind its hump-shaped impulse response function is slightly different. In this model, N_{1t} and N_{2t} are not perfect substitutes in production. Output is curved over both N_{1t} and N_{2t}; but utility is linear in both types of employment. Hence, in the absence of the one-period labor rigidity, the first-order conditions imply that the ratio of N_{1t} to N_{2t} would always remain constant at its steady-state level.

When a positive technology shock occurs the marginal product of employment on both shifts increases. Since the social planner faces a fixed stock of total employees and capital for one period, the planner can only adjust overtime employment immediately. Thus, when the shock impacts, output produced on the first shift increases only because of the rise in the state of technology; but output on the second shift (as well as aggregate output) increases beyond what is simply due to the rise in technology because of the increase in overtime employment. See Fig. 2.

After observing the shock the social planner sets all future total and overtime employment such that the marginal product of employment on each shift is equated to the utility loss of employing an additional worker on that shift. The highly persistent technology shock continues to raise symmetrically the marginal product of employment on both shifts; so the planner increases total and overtime employment proportionally such that their steady-state ratio holds. Since it lowers an agent's expected utility more to produce output on the second shift than on the first shift, this increase in the first shift's output more than offsets the slight decrease in the second shift's output (caused by the decay of the shock and the substitution of production across shifts). Thus output deviates from its nonstochastic steady-state level more during the subsequent period than during the initial period; this generates a hump-shaped response of output to a technology shock.

The differences in the labor market response to a technology shock within each model causes differences in the shock transmission properties of each model. For the effort model, a 1% shock to technology increases output 1.20% in the initial period and 1.37% in the subsequent period. For the overtime model, a 1% shock to technology increases output only 1.09% in the initial period but 1.48% in the following period. The relative increase in output from the first to the second period is much larger in the overtime model than in the effort model. Moreover the maximal response of output to a technology shock is larger in the overtime model than in the effort model.

In Table 4 one can see that both models imply an unconditional standard deviation of output of about 0.038. Table 5 reports the unconditional standard deviation of the state of technology, std(λ), for each model. For the overtime model it is 0.0260, while for the effort model it is 0.0277. One measure of amplification is the unconditional standard deviation of output divided by the unconditional standard deviation of λ. This measure is reported in the second row of Table 5. It implies that both models possess strong magnification mechanisms. The magnification mechanism within the effort model leads to a 42% increase in the volatility of output; for the overtime model the increase is slightly larger, 48%.

Another way to measure the propagation mechanisms is to analyze the autocorrelation of output growth. Cogley and Nason (1995) demonstrate that, while output growth in the data displays considerable autocorrelation, standard real business cycle models predict this autocorrelation to be white noise. Watson (1993)

Table 5
Magnification of shocks

	Effort model	Overtime model
std(λ)	0.0277	0.0260
std(q)/std(λ)	1.42	1.48

Table 6
Autocorrelation of output growth

Autocorrelations	1	2	3	4	5	6	7	8
Effort model	0.17	−0.02	−0.02	−0.02	−0.02	−0.02	−0.02	−0.02
Overtime model	0.30	−0.03	−0.03	−0.03	−0.03	−0.03	−0.03	−0.02
U.S. data	0.42	0.28	0.17	0.07	0.04	0.06	−0.03	−0.10

shows that standard RBC models cannot explain the peak of the spectrum of output growth at business cycle frequencies.

Table 6 presents the two models' implications for the autocorrelation of output growth. Both models imply a positive correlation at lag 1.[11] The lag 1 correlation coefficient for the overtime model is almost twice the size of the coefficient for the effort model. However, both models underestimate the autocorrelation of output growth.

It is straightforward to consider each model's implication on output growth in the frequency domain. In Figs. 3 and 4 the spectrum of output growth is plotted for each model. The spectrum of output growth for the data is also plotted. Both models imply spectra of the same shape; both models succeed in magnifying the shocks (which are white noise) such that a hump in the spectrum at business cycle frequencies is produced. This hump is more pronounced for the overtime model than for the effort model. Although both models dramatically overestimate the variance of output growth, particularly at high frequencies, this hump in the spectrum is a success for both models.

Why does the overtime model, relative to the effort model, have a smaller initial response but a larger maximal response to a technology shock? Or more generally, why does the overtime model do a better job propagating and magnifying shocks than the effort model? To this end, it is useful to reinterpret the effort model as a variable shift length model.[12] In this reinterpreted effort model, all employed agents work a single shift of length $h_1 e_t^{\alpha}$, but the social planner can choose the

[11] If adjustment costs to changing total employment are incorporated, both models can imply positive correlation coefficients for lags greater than one.

[12] I thank an anonymous referee for this suggestion.

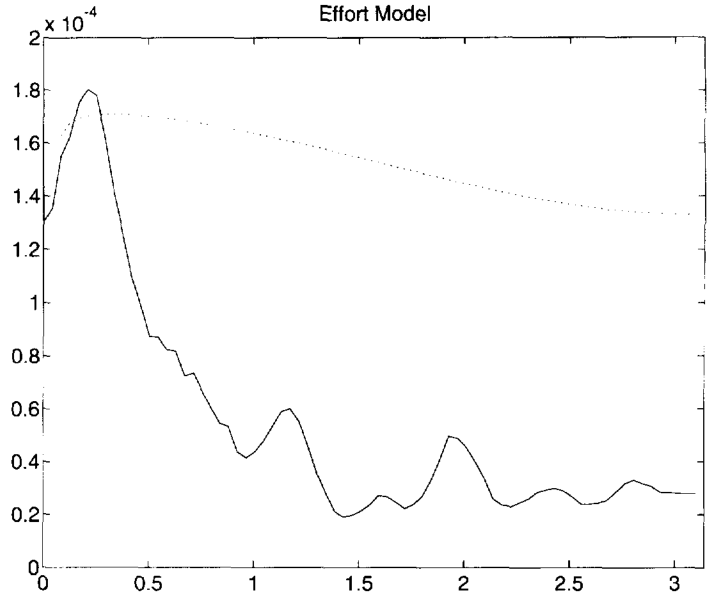

Fig. 3. The spectrum of output growth in the effort model (dotted) and the data (solid).

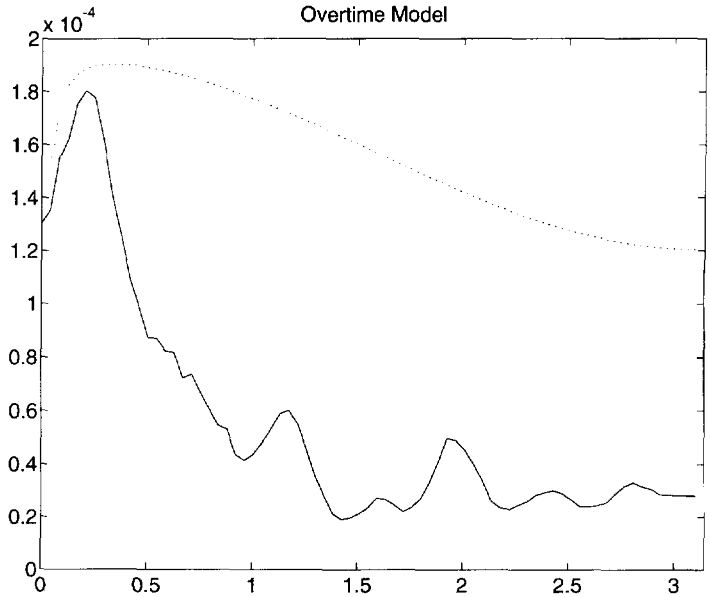

Fig. 4. The spectrum of output growth in the overtime model (dotted) and the data (solid).

shift length. Furthermore, all employed agents must work the same length shift and must work ξ nonproductive hours. The social planner in the overtime model does not face these last two constraints.

Consider again the impulse response exercises described above. When a technology shock occurs in the reinterpreted effort model, the social planner can immediately increase output by increasing the length of the shift; of course, increasing the shift length decreases the employed agent's utility by decreasing the amount of leisure each employed agent consumes. Likewise when a technology shock occurs in the overtime model, the social planner can increase output on the second shift by increasing N_{2t}; increasing N_{2t} decreases the utility of the agents who must now work both shifts rather than just the straight time shift. Because utility is curved in leisure for both models, the disutility of marginally increasing the shift length of those engaged in work in the effort model is less than the utility cost of increasing the shift length of a fraction of employed agents from h_1 to $h_1 + h_2$ in the overtime model. Moreover the marginal product of an additional unit of e_t in the effort model is greater than the marginal product of an additional worker on the second shift in the overtime model. Thus, the effort model generates a larger initial response to a technology shock than the overtime model.

After the initial period following the shock, e_t in the effort model returns to its steady-state level; so the reinterpreted effort model becomes, in effect, a fixed shift model. In these subsequent periods, the planner can take better advantage of a higher state of technology in the overtime model than in the effort model. The planner in the overtime model can always choose to have all employed agents work either both shifts or just the straight-time shift, so adding the constraint that all employed agents must work the same length shift cannot make agents better off. More importantly though, in the overtime model employed agents do not work any unproductive hours. Since utility is curved over hours of work, this absence of a fixed cost implies that the utility cost of increasing N_{1t} is less in the overtime model than in the effort model. Consequently the overtime model delivers a larger maximal output response to a technology shock than the effort model.

5. Conclusion

This paper formulates and estimates a real business model which differentiates between straight time and overtime. This model has only one latent variable, the state of technology, yet it does a better job propagating and magnifying shocks than Burnside, Eichenbaum, and Rebelo's (1993) labor hoarding model with its second latent variable, effort.

These two models do have their shortcomings. First, both models underestimate the variance of hours worked and employment. Second, both models underestimate the autocorrelation of output growth. Third, both models are unable to generate as pronounced a hump in the spectrum of output growth as is observed in the data. These shortcomings highlight the need for additional mechanisms within the models through which shocks can be propagated.

Finally firms and agents adjust their labor allocations along both the effort margin and the straight time/overtime margin. These margins are complementary adjustment mechanisms. But since both models ignore one of these margins, both models overstate the role of the other. Furthermore Bresnahan and Ramey (1994) demonstrate that for automobile assembly plants, the two most important margins along which firms adjust their labor input are shutting down the plant for a week at a time and opening and closing a second shift. Both the overtime and the effort models omit these nonconvex margins. Measuring the role these nonconvex margins play in providing a mechanism for the transmission of business cycle shocks is an issue for future research.

References

Altug, S., 1989, Time to build and aggregate fluctuations: Some new evidence, International Economic Review 30, 889–920.

Anderson, E., L. Hansen, E. McGrattan, and T. Sargent, 1995, Mechanics of forming and estimating dynamic linear economies, Manuscript (University of Chicago, Chicago, IL).

Brayton, F. and E. Mauskopf, 1985, The MPS model of the United States economy (Board of Governors of the Federal Reserve System, Division of Research and Statistics, Washington, DC).

Bresnahan, T. and V. Ramey, 1994, Output fluctuations at the plant level, Quarterly Journal of Economics 109, 593–624.

Burnside, C. and M. Eichenbaum, 1995, Factor hoarding and the propagation of business cycle shocks, Manuscript (Northwestern University, Evanston, IL).

Burnside, C., M. Eichenbaum, and S. Rebelo, 1993, Labor hoarding and the business cycle, Journal of Political Economy 101, 245–273.

Christiano, L., 1988, Why does inventory investment fluctuate so much?, Journal of Monetary Economics 21, 247–280.

Cochrane, J., 1994, Shocks, Carnegie–Rochester Conference Series on Public Policy 41, 295–364.

Cogley, T. and J. Nason, 1995, Output dynamics in real business cycle models, American Economic Review 85, 492–511.

Hansen, G., 1985, Indivisible labor and the business cycle, Journal of Monetary Economics 16, 309–327.

Hansen, G. and T. Sargent, 1988, Straight time and overtime in equilibrium, Journal of Monetary Economics 21, 281–308.

King, R., C. Plosser, and S. Rebelo, 1988, Production, growth and business cycles: I. The basic neoclassical model, Journal of Monetary Economics 21, 195–232.

Lucas, R., 1970, Capacity, overtime, and empirical production functions, American Economic Review Papers and Proceedings 60, 23–27.

McGrattan, E., 1994, The macroeconomic effects of distortionary taxation, Journal of Monetary Economics 33, 573–601.

McGrattan, E., R. Rogerson, and R. Wright, 1995, An equilibrium model of the business cycle with household production and fiscal policy, Staff report 191 (Federal Reserve Bank of Minneapolis, Minneapolis, MN).

Rogerson, R., 1988, Indivisible labor, lotteries and equilibrium, Journal of Monetary Economics 21, 3–16.

Sargent, T. and C. Sims, 1977, Business cycle modeling without pretending to have too much a priori economic theory, in: C. Sims, ed., New methods in business cycle research: Proceedings from a conference (Federal Reserve Bank of Minneapolis, Minneapolis, MN) 45–109.

Watson, M., 1993, Measures of fit for calibrated models, Journal of Political Economy 101, 1011–1041.