Estimation of Dynamic Labor Demand Schedules Under Rational Expectations

Thomas J. Sargent

September 1977

Working Paper #: 93

University of Minnesota
and
Federal Reserve Bank
of Minneapolis

The views expressed herein are solely those of the author and do not necessarily represent the views of the Federal Reserve Bank of Minneapolis or the Federal Reserve System. The material contained is of a preliminary nature, is circulated to stimulate discussion, and is not to be quoted without permission of the author.

Robert Litterman most ably performed the rather involved calculations reported in this paper.
Both Keynes and various classical writers asserted that real wages would move countercyclically as employers moved along downward sloping schedules relating the employment-capital ratio to the real wage. Dunlop [1938] and Tarshis [1939] described evidence which they interpreted as failing to confirm a countercyclical pattern of real wage movements. That and subsequent evidence of a similar nature helped to stimulate attempts to describe aggregate employment and real wages by "disequilibrium models," the work of Barro and Grossman [1971] and Solow and Stiglitz [1968] being two prominent examples. However, most of that empirical evidence stemmed from fitting static regressions in attempts to test static theories of the demand for employment. The recent paper by Salih Neftci [1977], which computes long two-sided distributed lags between employment and real wages, indicates that there are complicated and statistically significant dynamic interactions between real wages and employment, at least in the post-World War II U.S. data.

This paper estimates a dynamic aggregative demand schedule for employment for post-war U.S. data. While the demand model makes employment depend inversely on the appropriate real wage, as does the static theory, a potentially rich dynamic structure is introduced into that dependence because firms are assumed to face costs of rapidly adjusting their labor force and so find it optimal to take into account future expected values of the real wage in determining their current employment. The model imposes overidentifying restrictions on the bivariate real wage-employment stochastic process and therefore can be tested given a single sample of data.

The model is formed by blending the costly adjustment model of Lucas [1967], Treadway [1969], and Gould [1969] with Lucas's static
model of overtime work and capacity [1970]. The model is formulated so that it delivers linear decision rules relating the demand for straight-time employment and overtime employment each to the real wage process. The model imposes the rational expectations hypothesis, since firms are supposed to use the true moments of the real wage process in forming forecasts. The rational expectations hypothesis is a main source of the overidentifying restrictions imposed by the model.

In addition hopefully to providing some new evidence in the Dunlop-Tarshis tradition, this paper illustrates a technology for maximum likelihood estimation of decision rules under the hypothesis that expectations are rational. That technology potentially has a variety of applications.
1. The Demand for Employment

The model is formed by taking Lucas's model of overtime work and capacity [1970] and amending it to permit potentially different adjustment costs to be associated with rapidly changing straight-time and overtime labor. It is widely asserted that it is much cheaper to adjust the overtime labor force quickly than it is to adjust the straight-time labor force; consequently, it is alleged that overtime labor responds rapidly to the market signals that the firm receives, while the straight-time labor force responds more sluggishly. The model is designed to represent this phenomenon and to provide a framework for estimating its dimensions and testing it.

I shall work with a representative firm, although as I shall remark below, the model can handle certain kinds of diversity across firms. Following Lucas, suppose that the representative firm faces the instantaneous production function

\[ y(t+\tau) = f(n(t+\tau), k(t+\tau)), \quad f_n, f_k, f_{nk} > 0; \quad f_{nn}, f_{kk} < 0 \]

\[ t=0, 1, 2, 3, \ldots \]

\[ \tau \in [0,1). \]

Here \( y(t+\tau) \) is the rate of output per unit time at instant \( t+\tau \), \( n(t+\tau) \) is the number of employees at instant \( t+\tau \), and \( k(t+\tau) \) is the stock of capital at \( t+\tau \). The length of the "day" is 1, so that \( t \) indexes days and \( \tau \) indexes moments within the day. The firm is assumed to have a constant capital stock over the day so that

\[ k(t+\tau) = k(t) = k_\tau \text{ for } \tau \in [0,1). \]
The firm is assumed to be able to hire workers for a straight-time shift of fixed length \( h_1 < 1 \) at the real wage \( w_t \) during day \( t \). During the overtime shift of length \( h_2 = 1 - h_1 \), the firm can hire all the labor it wants during day \( t \) at the real wage \( a w_t \), where \( a \approx 1.5 \) is an overtime premium. Thus, for the first \( h_1 \) moments of day \( t \) the firm must pay workers \( w_t \), while for the remaining \( h_2 \) moments it must pay \( a w_t \). Confronted with these market opportunities it is optimal for the firm to choose to set \( n(t+\tau) = n_{1t} \) for \( \tau \in [0,h_1] \) and \( n(t+\tau) = n_{2t} \) for \( \tau \in (h_1,1) \). That is, it is optimal for the firm to choose a single level of straight-time employment \( n_{1t} \) during \( t \) and a single level of overtime employment of \( n_{2t} \) during the day \( t \).

The firm's output over the "day" is then

\[
y_t = \int_0^1 y(t+\tau) d\tau = h_1 f(n_{1t}, k_t) + h_2 f(n_{2t}, k_t).
\]

I take several steps to specialize this setup further. First, to simplify things, I assume that capital is constant over time so that \( k_t \) can be dropped as an argument from \( f(\cdot, \cdot) \). (In the econometric work below, steps are taken to detrend the data prior to estimation partly in order to minimize the damage caused by this approximation.) Second, I assume a quadratic production function so that we write instantaneous output on the first and second shifts as

\[
f(n_{1t}, k_t) = (f_0 + a_{1t}) n_{1t} - \frac{f_1}{2} n_{1t}^2
\]

\[
f(n_{2t}, k_t) = (f_0 + a_{2t}) n_{2t} - \frac{f_1}{2} n_{2t}^2
\]
where \( f_0, f_1 > 0 \), and where \( a_{1t} \) and \( a_{2t} \) are exogenous stochastic processes affecting productivity of straight-time and overtime employment. We assume that \( E a_{1t} = E a_{2t} = 0 \). The stochastic processes \( a_{1t} \) and \( a_{2t} \) will be required to satisfy certain regularity conditions to be specified below.

The firm is assumed to bear daily costs of adjusting its straight-time labor force of \( \frac{d}{2}(n_{1t} - n_{1t-1})^2 \) and to bear daily costs of adjusting its overtime labor force of \( \frac{e}{2}(n_{2t} - n_{2t-1})^2 \). It is widely believed that it is substantially more expensive to adjust the straight-time labor force so that \( d \gg e \). The firm faces an exogenous stochastic process for the real wage of \( \{w_t\} \). The firm's straight-time and overtime wage bills are, respectively, \( w_t h_1 n_{1t} \) and \( a w_t h_2 n_{2t} \).

The firm chooses contingency plans for \( n_{1t} \) and \( n_{2t} \) to maximize its expected real present value:

\[
V_t = \sum_{j=0}^{\infty} b^j \left[ (f_0 + a_{1t+j} - w_{t+j}) h_1 n_{1t+j} - \frac{f_1}{2} h_1 n_{1t+j}^2 \right. \\
- \frac{d}{2}(n_{1t+j} - n_{1t+j-1})^2 + (f_0 + a_{2t+j} - a w_{t+j}) h_2 n_{2t+j} \\
- \left. \frac{f_1}{2} h_2 n_{2t+j}^2 - \frac{e}{2}(n_{2t+j} - n_{2t+j-1})^2 \right] \]

where \( n_{1t-1} \) and \( n_{2t-1} \) as well as the stochastic processes for \( w, a_1, \) and \( a_2 \) are given to the firm. Here \( b \) is a real discount factor that lies between zero and one. The operator \( E_t \) is defined by \( E_t x = E x | \Omega_t \) where \( x \) is a random variable, \( E \) is mathematical expectation, and \( \Omega_t \) is an information set available to the firm at time \( t \). I assume that \( \Omega_t \) includes at least \( \{n_{1t-1}, n_{2t-1}, a_{1t}, a_{1t-1}, \ldots, a_{2t}, a_{2t-1}, \ldots, w_t, w_{t-1}, \ldots\} \).
The firm is assumed to maximize (1) by choosing stochastic processes for \( n_t \) and \( n_{2t} \) from the set of stochastic processes that are (nonanticipative) functions of the information set \( \Omega_t \). (Below, I will further restrict the class of stochastic processes over which the optimization is carried out.) I assume that the stochastic processes \( \omega_t \), \( a_{1t} \), and \( a_{2t} \) are of exponential order less than \( \left( \frac{1}{b} \right)^j \), which means that for some \( K > 0 \)

\[
\begin{align*}
|E_t \omega_{t+j}| &< K\left( \frac{1}{b} \right)^j \\
|E_t a_{1t+j}| &< K\left( \frac{1}{b} \right)^j \\
|E_t a_{2t+j}| &< K\left( \frac{1}{b} \right)^j
\end{align*}
\]

for all \( t \) and all \( j \geq 0 \).

First-order necessary conditions for the maximization of (1) consist of a set of "Euler equations" and a pair of transversality conditions.\(^4\) The Euler equations for \( \{n_{1t}\} \) and \( \{n_{2t}\} \) are

\[
\begin{align*}
bE_{t+j} n_{1t+j+1} + \phi_1 n_{1t+j} + n_{1t+j-1} &= \frac{h_1}{d}(\omega_{t+j} - a_{1t+j} - f_0) \\
j &= 0, 1, 2, \ldots
\end{align*}
\]

(2)

\[
\begin{align*}
bE_{t+j} n_{2t+j+1} + \phi_2 n_{2t+j} + n_{2t+j-1} &= \frac{h_2}{e}(\omega_{t+j} - a_{2t+j} - f_0) \\
j &= 0, 1, 2, \ldots
\end{align*}
\]

(3)

where

\[
\begin{align*}
\phi_1 &= -\left( \frac{f_1 h_1}{d} + (1+b) \right) \\
\phi_2 &= -\left( \frac{f_2 h_2}{e} + (1+b) \right).
\end{align*}
\]
The transversality conditions are

\[ \lim_{T \to \infty} b^T e_n^T = \lim_{T \to \infty} b^T e_n^T 2T = 0. \]

To solve the Euler equations for the optimum contingency plans, first obtain the factorizations

\[ (1 + \frac{\phi_1}{b}z + \frac{1}{b^2}z^2) = (1-\delta_1 z)(1-\delta_2 z) \]
\[ (1 + \frac{\phi_2}{b}z + \frac{1}{b^2}z^2) = (1-\mu_1 z)(1-\mu_2 z). \]

Given the assumptions about the signs and magnitudes of the parameters composing \( b, \phi_1, \) and \( \phi_2, \) it follows that factorizations exist with 
\( 0 < \delta_1 < 1 < \frac{1}{b} < \delta_2 \) and \( 0 < \mu_1 < 1 < \frac{1}{b} < \mu_2. \) It then follows that solutions of the Euler equations that satisfy the transversality conditions and the initial conditions are given by

\[ \begin{align*}
(n_1^t) & = \delta_1 n_{1t-1} - \frac{\delta_1 h_1}{d} \sum_{i=0}^{\infty} \left( \frac{1}{\delta_2} \right)^i e_t (w_{t+1} + a_{1t+1} - f_0) \\
(n_2^t) & = \mu_1 n_{2t-1} - \frac{\mu_1 h_2}{e} \sum_{i=0}^{\infty} \left( \frac{1}{\mu_2} \right)^i e_t (aw_{t+1} + a_{2t+1} - f_0). 
\end{align*} \]

It can be verified directly that these solutions satisfy the Euler equations and the transversality conditions. The polynomial equation (5) implicitly defines \( \delta_1 \) and \( \delta_2 \) as functions of \( \frac{f_1 h_1}{d}. \) By studying this polynomial, \( \delta_1 \) it is possible to show that \( \delta_1 \) is a decreasing function of \( \frac{f_1 h_1}{d} \) and that \( \frac{1}{\delta_2} = b \delta_1. \) It follows that \( \delta_1 \) and \( \frac{1}{\delta_2} \) both increase with increases in the adjustment cost parameter \( d. \) Reference to equation (7a) then shows that increases in the adjustment cost parameter \( d, \) by increasing \( \delta_1 \) and \( \frac{1}{\delta_2}, \) decrease the speed with which the firm responds to the real wage and productivity signals that it receives. Similarly, \( \mu_1 \) and \( \frac{1}{\mu_2} \) are decreasing functions of \( \frac{f_1 h_2}{e} \), and \( \frac{1}{\mu_2} = b \mu_1. \)
Equations (7) are decision rules for setting $n_{1t}$ and $n_{2t}$ as linear functions of $n_{1t-1}$, $n_{2t-1}$, and the conditional expectations $E_t w_{t+i}$, $E_t a_{1t+i}$, and $E_t a_{2t+i}$, $i = 0, 1, 2, \ldots$. However, in general, these conditional expectations are nonlinear functions of the information in $\Omega_t$. Given particular stochastic processes for $w_t$, $a_{1t}$, and $a_{2t}$, equations (7) can be solved for decision rules expressing $n_{1t}$ and $n_{2t}$ as, in general, nonlinear functions of $\Omega_t$.

For the purposes of empirical work, it is convenient to restrict ourselves to the class of decision rules that are linear functions of $\Omega_t$. The optimal linear decision rules can be obtained by replacing the conditional mathematical expectations in (7) with the corresponding linear least squares projections on the information set $\Omega_t$. Accordingly, henceforth, I will interpret $E$ as the linear least squares projection operator.

To derive from (7) explicit decision rules for $n_{1t}$ and $n_{2t}$ as functions of $\Omega_t$, it is necessary further to restrict the stochastic processes $w_t$, $a_{1t}$, and $a_{2t}$. I assume that $a_{1t}$ and $a_{2t}$ are each first-order Markov processes for which

$$E_t a_{1t+i} = \rho_1 a_{1t} \quad i \geq 0$$

$$E_t a_{2t+i} = \rho_2 a_{2t} \quad i \geq 0$$

where $|\rho_1| < \frac{1}{b}$, $|\rho_2| < \frac{1}{b}$. I assume that $w_t$ is an $n^{th}$-order Markov process

$$w_t = v_0 + v_1 w_{t-1} + v_2 w_{t-2} + \ldots + v_n w_{t-n} + \xi_{3t}$$

where $\xi_{3t}$ is a least squares disturbance that satisfies $E_t \xi_{3t} = 0$. It is convenient to represent the $n^{th}$-order process

$$E_2^{3t} |\Omega_{t-1} = 0.$$
(10) as the \((n+1)\)-vector first-order Markov process

\[ x_t = Ax_{t-1} + \varepsilon_t \]

where

\[ X_t = \begin{bmatrix} w_t \\ w_{t-1} \\ w_{t-2} \\ \vdots \\ w_{t-n} \\ 1 \end{bmatrix} \quad \varepsilon_t = \begin{bmatrix} \varepsilon_{3t} \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix} \]

\[ A = \begin{bmatrix} v_1 & v_2 & \cdots & v_n & v_0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 & 1 \end{bmatrix} \]

We can write:

\[ x_{t+1} = Ax_t + \varepsilon_{t+1} \]

\[ x_{t+2} = A^2x_t + \varepsilon_{t+2} + A\varepsilon_{t+1} \]

\[ \vdots \]

\[ x_{t+j} = A^jx_t + \varepsilon_{t+j} + A\varepsilon_{t+j-1} + \cdots + A^{j-1}\varepsilon_{t+1} \]
Since $E_{t}x_{t+k} = 0$ for $k \geq 1$, we have
\[ E_{t}x_{t+j} = A^{j}x_{t}. \]

Assume that the eigenvalues of $A$ are distinct so that $A$ can be written as
\[ A = P\Lambda P^{-1} \]
where the columns of $P$ are the eigenvectors of $A$ and $\Lambda$ is the diagonal matrix whose elements are the eigenvalues of $A$. Then we have
\[ E_{t}x_{t+j} = PA^{j}Px_{t}. \]

Finally, let $c$ be the $1 \times (n+1)$ row vector $(1, 0, 0, \ldots, 0)$ so that $w_{t} = cx_{t}$. We thus have that
\[ (11) \quad E_{t}w_{t+j} = cPA^{j}P^{-1}x_{t}. \]

Substituting from (8) and (11) into (7a) gives
\[ n_{lt} = \delta_{1}n_{lt-1} - \frac{\delta_{1}h_{1}}{d}cP \sum_{i=0}^{\infty} \left( \frac{1}{\delta_{2}} \right)^{i}P^{-1}x_{t} \]
\[ + \frac{\delta_{1}h_{1}}{d} \left( \frac{f_{0}}{1 - \frac{1}{\delta_{2}}} \right) + \frac{\delta_{1}h_{1}}{d} \left( \frac{1}{\rho_{1}} \right)a_{lt}. \]

Let $\lambda_{i}$ be the $i^{th}$ element of $\Omega$. Since $\delta_{2} = \frac{1}{\delta_{1}b}$, we have that $|\lambda_{i}| = |\lambda_{i}\delta_{1}b| < 1$ by virtue of the assumption that $w_{t}$ is of exponential order less than $\frac{1}{b}$, i.e., that $|\lambda_{i}b| < 1$. Then the infinite sum above converges and we can write
\begin{align}
(12) \quad n_{1t} &= \delta_1 n_{1t-1} - \frac{\delta_1 h_1}{d} cP \left[ \frac{1}{1 - \frac{\lambda_1}{\delta_2}} \right]_{ii} P^{-1} x_t \\
&\quad + \frac{\delta_1 h_1}{d} \left( \frac{f_0}{1 - \frac{1}{\delta_2}} \right) + \frac{\delta_1 h_1}{d} \left( \frac{1}{1 - \rho_1} \right) a_{1t} \\
\end{align}

where $\left[ \frac{1}{\lambda_i} \right]_{ii}$ is a diagonal matrix with $(1 - \frac{\lambda_i}{\delta_2})$ as the $i^{th}$ diagonal element.

Let us write (12) as

\begin{align}
(13) \quad n_{1t} &= \delta_1 n_{1t-1} + a_{1t} + a_{2t} + \cdots + a_{n-1t} + a_{nt} + a_{0t} \\
&\quad + \frac{\delta_1 h_1}{d} \left( \frac{f_0}{1 - \frac{1}{\delta_2}} \right) + a_{1t} \\
\end{align}

where

\begin{align}
(14) \quad (a_1, a_2, \ldots, a_n, a_0) &= - \frac{\delta_1 h_1}{d} cP \left[ \frac{1}{1 - \frac{\lambda_1}{\delta_2}} \right]_{ii} P^{-1} \\
a_{1t}' &= \frac{\delta_1 h_1}{d} \left( \frac{1}{1 - \rho_1} \right) a_{1t}. \\
\end{align}

Proceeding in the same way, we can write the decision rule for $n_{2t}$ as

\begin{align}
(15) \quad n_{2t} &= \mu_1 n_{2t-1} + \beta_1 w_t + \beta_2 w_{t-1} + \cdots + \beta_{n-1} w_{t-n+1} + \beta_0 \\
&\quad + \frac{\mu_1 h_2}{e} \left( \frac{f_0}{1 - \frac{1}{\mu_1}} \right) + a_{2t}' \\
\end{align}

where

\begin{align}
(16) \quad (\beta_1, \beta_2, \ldots, \beta_n, \beta_0) &= - \frac{\mu_1 h_2}{e} cP \left[ \frac{1}{1 - \frac{\lambda_1}{\mu_1}} \right] P^{-1} \\
a_{2t}' &= \frac{\mu_1 h_2}{e} \left( \frac{1}{1 - \rho_2} \right) a_{2t}'. \\
\end{align}
Equations (14) and (16) succinctly summarize how the distributed lag coefficients, the \( \alpha \)'s and \( \beta \)'s, reflect the combination of forecasting (through the parameters of \( P \) and \( A \)) and optimization (through the parameters \( d, \delta, \) and \( \mu \)) elements. Clearly, the decision rules (13) and (15) are not invariant with respect to the stochastic process for real wages (8), a general characteristic of optimum decision rules whose far reaching implications for econometric policy evaluation have been stressed by Robert E. Lucas, Jr., [1976].

Since I will fit the model to data that are deviations from means and trends, I shall henceforth drop the constants from (13), (15), and (10). Substitute (10) for \( w_t \) and subtract \( \rho_{1}a'_{1t-1} \) from both sides of (13) to get

\[
\begin{align*}
(17) \quad n_{lt} &= (\delta_{1}+\rho_{1})n_{lt-1} - \rho_{1}\delta_{1}n_{lt-2} + (\alpha_{2}+\alpha_{1}v_{2}-\alpha_{2}\rho_{1})w_{t-1} \\
&\quad + (\alpha_{3}+\alpha_{1}v_{2}-\alpha_{2}\rho_{1})w_{t-2} + \ldots + (\alpha_{n}+\alpha_{1}v_{n-1}-\alpha_{n-1}\rho_{1})w_{t-n+1} \\
&\quad + (\alpha_{1}v_{n}-\alpha_{n}\rho_{1})w_{t-n} + [\alpha_{1}\xi_{3t} + (a'_{1t}-\rho_{1}a'_{1t-1})].
\end{align*}
\]

From our earlier assumptions, \( E_{t-1}[\alpha_{1}\xi_{3t} + (a'_{1t}-\rho_{1}a'_{1t-1})] = 0 \), so that (17) is the (vector) autoregression for \( n_{lt} \). In particular, we have

\[
(18) \quad E_{t-1}n_{lt} = (\delta_{1}+\rho_{1})n_{lt-1} - \rho_{1}\delta_{1}n_{lt-2} + (\alpha_{2}+\alpha_{1}v_{2}-\alpha_{2}\rho_{1})w_{t-1} \\
&\quad + (\alpha_{3}+\alpha_{1}v_{2}-\alpha_{2}\rho_{1})w_{t-2} + \ldots + (\alpha_{n}+\alpha_{1}v_{n-1}-\alpha_{n-1}\rho_{1})w_{t-n+1} \\
&\quad + (\alpha_{1}v_{n}-\alpha_{n}\rho_{1})w_{t-n}.
\]
Similarly, we have for $n_{2t}$

\begin{align*}
 n_{2t} &= (\mu_1 + \rho_2)n_{2t-1} - \rho_2 n_{2t-2} + (\beta_2^1 + \beta_1^1 \gamma_2^1 - \beta_2^2 \rho_2^2)w_{t-1} \\
 &\quad + (\beta_3^1 + \beta_1^1 \gamma_2^1 - \beta_2^2 \rho_2^2)w_{t-2} + \ldots + (\beta_n^1 + \beta_1^1 \gamma_{n-1}^1 - \beta_{n-1}^1 \rho_2^2)w_{t-n+1} \\
 &\quad + (\beta_1 \gamma_1 - \beta_n \rho_2^2)w_{t-n} + [\beta_1 t + (a_2^1 - \rho_1 a_2^1)].
\end{align*}

We can now write the complete three-variate vector autoregression for $n_{1t}$, $n_{2t}$, $w_t$ as

(a) $n_{1t} = (\delta_1 + \rho_1)n_{1t-1} - \rho_1 \delta_1 n_{1t-2} + (\alpha_2^1 + \alpha_1^1 \gamma_2^1 - \alpha_2 \rho_1^1)w_{t-1}$

\begin{align*}
 &\quad + (\alpha_3^1 \gamma_2 - \alpha_2^1 \rho_1)w_{t-1} + \ldots + (\alpha_n^1 \gamma_{n-1} - \alpha_{n-1}^1 \rho_1^1)w_{t-n+1} \\
 &\quad + (\alpha_1 \gamma_1 - \alpha_n \rho_1)w_{t-n} + u_{1t}
\end{align*}

(b) $n_{2t} = (\mu_1 + \rho_2)n_{2t-1} - \rho_2 \mu_1 n_{2t-2} + (\beta_2^1 + \beta_1^1 \gamma_2^1 - \beta_2^2 \rho_2^2)w_{t-1}$

\begin{align*}
 &\quad + (\beta_3^1 + \beta_1^1 \gamma_2^1 - \beta_2^2 \rho_2^2)w_{t-2} + \ldots + (\beta_n^1 + \beta_1^1 \gamma_{n-1}^1 - \beta_{n-1}^1 \rho_2^2)w_{t-n+1} \\
 &\quad + (\beta_1 \gamma_1 - \beta_n \rho_2^2)w_{t-n} + u_{2t}
\end{align*}

(c) $w_t = v_1 w_{t-1} + v_2 w_{t-2} + \ldots + v_n w_{t-n} + u_{3t}$

where

\[
 u_t = \begin{bmatrix} u_{1t} \\ u_{2t} \\ u_{3t} \end{bmatrix} = \begin{bmatrix} \alpha_1 \xi_3 + (a_1^1 \gamma_3^1 - \rho_1 \rho_1^1) \\ \beta_1 \xi_3 + (a_2^1 \gamma_3^1 - \rho_2 \rho_2^1) \\ \xi_3 \end{bmatrix} = \begin{bmatrix} n_{1t} - E_{t-1} n_{1t} \\ n_{2t} - E_{t-1} n_{2t} \\ w_t - E_{t-1} w_t \end{bmatrix}.
\]

Here $u_t$ is the vector of innovations, i.e., errors in predicting $(n_{1t}, n_{2t}, w_t)$ from past information. There are $(3n+4)$ regressors in (20), i.e.,
of which appear three times, and $n_{1t-1}$, $n_{1t-2}$, $n_{2t-1}$, and $n_{2t-2}$, each of which appears once. The free parameters of the model are $f_1$, $d$, $e$, $e_1$, $e_2$, $v_1$, ..., $v_n$, so that there are $(n+5)$ parameters to be estimated. As it turns out, the model is overidentified for $n>1$.

Collecting the equations that summarize the restrictions that the model imposes on the vector autoregression (20), we have

$$
egin{aligned}
\phi_1 &= -(\frac{f_1 h_1}{d} + (1+b)) \\
\phi_2 &= -(\frac{f_1 h_2}{e} + (1+b)) \\
(1 + \frac{\phi_1 z}{b z} + \frac{1}{b z^2}) &= (1-\delta_1 z)(1-\delta_2 z) \\
(1 + \frac{\phi_2 z}{b z} + \frac{1}{b z^2}) &= (1-\mu_1 z)(1-\mu_2 z) \\
(\alpha_1, \alpha_2, \ldots, \alpha_n, \alpha_0) &= -\frac{\delta_1 h_1}{d} c P[\frac{1}{1-\lambda_1 \delta_1}] il^{-1} \\
(\beta_1, \beta_2, \ldots, \beta_n, \beta_0) &= -\frac{\mu_1 h_2}{e} c P[\frac{1}{1-\lambda_1 \mu_1}] il^{-1} \\
A &= PAP^{-1}.
\end{aligned}
$$

Estimates of the free parameters $\phi = (f_1, d, e, e_1, e_2, v_1, \ldots, v_n)$ are obtained by using the method of maximum likelihood to estimate the vector autoregression (20), subject to (21). Let $\hat{u}_t = (\hat{u}_{1t}, \hat{u}_{2t}, \hat{u}_{3t})$ be the sample residual vector associated with the parameter values $\theta$. Under the assumption that $u_t$ is a trivariate normal vector with $E u_t u_t' = V$, the likelihood function of a sample of observations on the residuals extending over $t=1, \ldots, T$ is

$$
L(\theta) = (2\pi)^{-\frac{T}{2}} |V|^{-\frac{1}{2}} \exp(-\frac{1}{2} \sum_{t=1}^{T} \hat{u}_t' V^{-1} \hat{u}_t).
$$
As shown by Wilson [1973] and Bard [1974], maximum likelihood estimates of $\theta$ with $V$ unknown can be obtained by minimizing $|\hat{V}|$ with respect to $\theta$, where $\hat{V}$ is the sample covariance matrix of $u_t$,

$$\hat{V} = \frac{1}{T} \sum_{t=1}^{T} \hat{u}_t \hat{u}_t'.$$

The matrix $\hat{V}$ is the maximum likelihood estimator of $V$ (see Wilson [1973] or Bard [1974]).

Now consider the unconstrained version of the vector autoregression (20) in which each of the $(3n+4)$ regressors has its own free parameter. Let $L_u$ be the value of the likelihood function at its unrestricted maximum, i.e., the maximum obtained by permitting each of the $(3n+4)$ regressors to have its own free parameter. Let $L_r$ be the value of the likelihood under the restrictions (21). Then $-2 \log_e (L_r/L_u)$ is asymptotically distributed as $\chi^2(q)$ where $q = (3n+4) - (n+5)$ is the number of restrictions imposed by the theory. High values of the likelihood ratio lead to rejection of the restrictions that the theory imposes on the vector autoregression. Using the calculations of Wilson [1973, p. 80] or Bard [1974], it can be shown that the likelihood ratio is equal to

$$T \left\{ \log_e |\hat{V}_r| - \log_e |\hat{V}_u| \right\}$$

where $\hat{V}_r$ and $\hat{V}_u$ are the restricted and unrestricted estimates of $V$, respectively.
2. Alternative Estimation Strategies

It should be stressed that the vector autoregression (20) which builds in the cross-equation restrictions implied by the model has been obtained under the assumption (8) that the productivity shocks $a_{1t}$ and $a_{2t}$ are first-order Markov processes. The forms of the vector autoregressions (20) would be altered had we assumed other forms for the $a_{1t}$ and $a_{2t}$ processes, as the reader can verify by calculations parallel- ing those above.

An alternative estimation strategy is available that avoids the necessity to make specific assumptions about the forms of the stochastic processes for the disturbances $a_{1t}$ and $a_{2t}$, only requiring that these processes be covariance stationary. The alternative estimator requires instead that the $w_t$ process be strictly econometrically exogenous with respect to $n_{1t}$ and $n_{2t}$, in particular requiring that $E_{t-1}a_{1s} = E_{t-1}a_{2s} = 0$ for all $t$ and $s$. Under that assumption, the model (7a) and (7b) can readily be shown to place restrictions on the projections of $n_{1t}$ and $n_{2t}$, respectively, on the entire $\{w_s\}$ process. The structure of those restrictions parallels those worked out by Sargent [1977a] for a consumption function example. An asymptotically efficient estimator such as "Hannan's efficient estimator," which allows for complicated serial correlation patterns in the disturbances, could then be applied to estimating the projections with and without the restrictions imposed by the model.

This alternative estimation strategy gets along with much weaker assumptions about the serial correlation properties of the disturbance processes $\{a_{1t}\}$ and $\{a_{2t}\}$ at the cost of making somewhat more
stringent assumptions about the exogeneity of $w_t$, i.e., about the correlation between $w_t$ and the $a_i$s. The original estimator proposed that operates on (20) does assume that $\{w_t\}$ is a process that is not caused in Granger's [1969] sense by $n_{1t}$ or $n_{2t}$, i.e., that $E[w_t|w_{t-1}, w_{t-2}, \ldots, n_{1t-1}, n_{1t-2}, \ldots, n_{2t-1}, n_{2t-2}, \ldots] = E[w_t|w_{t-1}, w_{t-2}, \ldots]$. Now Sims' [1972] theorems assure us that if $w_t$ is not Granger-caused by $n_{1t}$ or $n_{2t}$, then there exists a statistical representation in which $w_t$ is strictly econometrically exogenous with respect to $n_{1t}$ or $n_{2t}$. However, this statistical representation need not correspond with the appropriate economic behavioral relationship. It is possible for $n_{1t}$ or $n_{2t}$ to fail to cause $w_t$, and yet for "instantaneous causality" to flow from $n_{1t}$ or $n_{2t}$ to $w_t$ so that $w_t$ is not strictly exogenous in the appropriate model. See Sargent [1977b] for an example of this phenomenon within the context of Cagan's model of hyperinflation. The "autoregressive estimator" based on (20) permits arbitrary correlation between the innovations to $n_{1t}$ or $n_{2t}$ and $w_t$ and makes no assumption about which pattern of instantaneous causality explains that correlation. On the other hand, the alternative "projection estimator" attributes all of that correlation to the workings of the demand schedules for $n_{1t}$ and $n_{2t}$, (7a) and (7b). For the present application, I prefer the estimator that makes the weaker assumption about the correlation between innovations to employment and the real wage.

The reader by now will have understood that optimizing, rational expectations models do not entirely eliminate the need for side assumptions not grounded in economic theory. Some arbitrary assumptions about the nature of the serial correlation structure of the disturbances and/or about strict econometric exogeneity are necessary in order to proceed with estimation.
Perhaps I should conclude this section by pointing to another source of arbitrariness, namely the latitude at our disposal in specifying the firm's optimization problem. For example, adding terms like \(-\frac{d^2}{2}(n_{1t-1} - n_{1t-2})^2\) to the firm's daily profits would lead to Euler equations that are fourth-order stochastic difference equations and would lead to decision rules that depend on two lagged values of employment. Such specifications would seem plausible and would lead to materially different restrictions than those above on vector autoregressions (or projections of \(n\) on \(w\), as the case may be). There are clearly limits set by the requirements of econometric identification on our ability to estimate such complicated adjustment cost parameterizations. Identification problems in such models have as yet received little attention at a general level.

The general theme of this section has been to issue a warning that rational expectations, optimizing models will not be able to save us entirely from the ad hoc assumptions and interpretations made in applied work. However, this is not to deny that the rational expectations hypothesis seems promising as a device for organizing restrictions on parameterizations of econometric models.
3. Parameter Estimates

The model was estimated using quarterly data on total civilian employment and a straight-time real wage index, with the period of observation for the dependent variables extending from 1947I through 1972IV. The variable $n_{1t}$ was in the first instance measured by the seasonally adjusted BLS series "Employees on Nonagricultural Payrolls, Private and Government." To get a measure of $n_{2t}$, the following procedure was used. I defined the variable $h_t$ to be average weekly hours, a series measured by the seasonally adjusted BLS series "Average Weekly Hours in Manufacturing." I then estimated total manhours by $h_t n_{1t}$. Finally, I measured $n_{2t}$ by

$$n_{2t} = \frac{\bar{h}_t n_{1t} - h_1 n_{1t}}{h_2}$$

where $h_1$ and $h_2$ were set a priori at 37 and 17, respectively.\(^{12/}\) The real wage $w_t$ was measured by deflating the seasonally unadjusted BLS series "Average Hourly Earnings: Straight-time Manufacturing Production Workers" by the seasonally unadjusted Consumer Price Index (1967=100).

I also created seasonally unadjusted measures of $n_{1t}$ and $n_{2t}$ by taking as a measure of $n_{1t}$ the seasonally unadjusted BLS series "Employees on Private Nonagricultural Payrolls" and then using the preceding procedure to create estimates of $n_{2t}$ by using the seasonally unadjusted average weekly hours series. The data are quarterly averages of monthly data. Notice that $h_1$ and $h_2$ are constants that are independent of time.

I begin by describing the estimates obtained using the seasonally adjusted employment series together with the seasonally unadjusted real
wage series. (Later I will describe the results obtained with the seasonally unadjusted series for all variables.) Before estimating the model, the data on \( n_{1t} \) and \( n_{2t} \) were each detrended by regressing them on a constant, linear trend, and trend squared, and then using the residuals from those regressions as the data for estimating the model. The data on \( w_t \) were formed as the residuals from a regression on a constant, linear trend, trend squared, and three seasonal dummies. Two reasons can be given for detrending in this way prior to fitting the model. First, the model ignores the effects of capital on employment, except to the extent that these can be captured by the productivity processes \( a_{1t} \) and \( a_{2t} \). Second, the theory predicts that any deterministic components of the employment and real wage processes will not be related by the same distributed lag model as are their indeterministic parts. Detrending prior to estimation is a device designed to isolate the indeterministic components. The real wage is measured in 1967 dollars, while employment is measured in millions of men.

Table 1 reports estimates of the vector autoregressions for \((n_{1t}, n_{2t}, w_t)\) both unconstrained and constrained by the model. Each set of estimates was obtained by the method of maximum likelihood. We have set \( n \) equal to four, a fourth-order autoregression being used to model the real wage process. This means that the likelihood ratio statistic is asymptotically distributed as chi-square with \( q = (3n+4) - (n+5) = 7 \) degrees of freedom. The likelihood ratio is 7.3172, which has a marginal confidence level of .6034. The model thus passes the likelihood ratio test of its overidentifying restrictions at the usual significance levels.

The parameter estimates for the model are reported in Table 2. The free parameters were \( f, d, e, \rho_1, \rho_2, v_1, v_2, v_3, \) and \( v_4 \) with b
being fixed at .95, $h_1$ at 37, and $h_2$ at 17. The decision rules associated with the maximum likelihood estimates are:

\[
n_{1t} = 0.5782n_{1t-1} - 1.3781w_t + 0.0580w_{t-1} + 0.1098w_{t-2} + 0.2929w_{t-3} + \alpha_{1t}
\]

\[
n_{2t} = 0.1979n_{2t-1} - 4.2723w_t - 0.0065w_{t-1} - 0.0217w_{t-2} + 0.1725w_{t-3} + \alpha_{2t}.
\]

Notice how both the shape of the distributed lag and the magnitude of the response to the real wage differs between straight-time and overtime employment. Overtime employment is more responsive to the real wage. Further, the straight-time adjustment cost parameter $d$ is estimated to be much larger than the overtime adjustment parameter $e$. That is why $n_{1t}$ depends more strongly on $n_{1t-1}$ than does $n_{2t}$ on $n_{2t-1}$.

Table 2 also reports the estimated covariance matrix of the innovations $V = \text{E} u_t u_t'$. Recall that

\[
\begin{bmatrix}
u_{1t} \\
u_{2t} \\
u_{3t}
\end{bmatrix} =
\begin{bmatrix}
1 & 0 & \alpha_1 \\
0 & 1 & \beta_1 \\
0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
\xi_{1t} \\
\xi_{2t} \\
\xi_{3t}
\end{bmatrix}
\equiv \mathbf{B} \xi_t
\]

where $\xi_{1t} = \alpha_{1t} - \rho_1 \alpha_{1t-1}$, $\xi_{2t} = \alpha_{2t} - \rho_2 \alpha_{2t-1}$, and where

\[
\mathbf{B} =
\begin{bmatrix}
1 & 0 & \alpha_1 \\
0 & 1 & \beta_1 \\
0 & 0 & 1
\end{bmatrix}, \quad \xi_t =
\begin{bmatrix}
\alpha_{1t} \\
\alpha_{2t} \\
\alpha_{3t}
\end{bmatrix}.
\]
Then, since $X_t = B^{-1} u_t$, the covariance matrix of $X_t$ can be estimated from $E X_t X_t' = B^{-1} V B^{-1}$, an estimate of which is also reported in Table 2. The correlation between the innovations to $a_{1t}$ and to $a_{2t}$, i.e., $\xi_{1t}$ and $\xi_{2t}$, is estimated to be .759. The correlation between the innovations to $a_{1t}$ and $\omega_t$, i.e., $\xi_{1t}$ and $\xi_{3t}$ is only .0215, while that between $\xi_{2t}$ and $\xi_{3t}$ is .0127. I had expected $\xi_{1t}$ and $\xi_{2t}$ to be even more highly correlated than they are.

As it happens, the estimates reported in Tables 1 and 2 correspond to the higher of two local maxima of the likelihood function which I found. The parameter estimates associated with the lower of these two local maxima are reported in Table 3. In view of the form the vector autoregression (20), it is not at all surprising that the likelihood function should exhibit multiple maxima. In particular, notice that the coefficients in (20) on $n_{1t-1}$, $n_{lt-2}$, $n_{2t-1}$, $n_{2t-2}$ are, respectively, $(\delta_1 + \rho_1)$, $-\delta_1 \rho_1$, $(\mu_1 + \rho_2)$, and $-\mu_1 \rho_2$. If it were not for the constraints across $\mu_1$ and the $\beta$'s and across $\delta_1$ and the $\alpha$'s and the appearance of $\rho_1$ and $\rho_2$ elsewhere on the right side of (20), the parameters $\delta_1$, $\rho_1$, $\mu_1$, and $\rho_2$ would not be identified, since it would be impossible to distinguish the effects of $\delta_1$ from $\rho_1$ and the effects of $\mu_1$ from $\rho_2$. The presence of lagged $\omega$'s on the right side of (20) and the aforementioned constraints resolve this identification problem but leave a vestige of it in the form of probable multiple peaks in the likelihood function with small samples. Comparing the parameter estimates in Tables 2 and 3 shows that Table 2 is a high $(\rho_1, \rho_2)$ - low $(\delta_1, \mu_1)$ solution, while Table 3 reports the high $(\delta_1, \mu_1)$ - low $(\rho_1, \rho_2)$ solution. Notice that for the Table 3 estimates, $\rho_1 + \delta_1 = 1.534$ and $\rho_1 \delta_1 = .538$, while for the Table 2 estimates, $\rho_1 + \delta_1 = 1.516$ while $\rho_1 \delta_1 = .542$. 
The presence of multiple maxima of the likelihood function means that caution is called for in interpreting the test statistics reported, since the asymptotic distribution on which the test is computed does not predict multiple maxima for the likelihood function and so does not provide a very good approximation for the sample size that we are studying. The presence of multiple maxima of the likelihood function also argues for starting the nonlinear estimation from several different initial parameter estimates. I followed this practice in each case reported below and, in each case, found another lower local maximum of the likelihood function in addition to the one reported below. In each case there was a high \((\delta_1, \mu_1)\) - low \((\rho_1, \rho_2)\) solution and a high \((\rho_1, \rho_2)\) low \((\delta_1, \mu_1)\) solution.

Table 4 reports the estimates associated with the higher likelihood of two maxima found for the seasonally unadjusted data with \(n=4\). In this case, the high \((\delta_1, \mu_1)\) - low \((\rho_1, \rho_2)\) solution had the higher likelihood. The estimates indicate \(d >> e\) and are qualitatively similar to those described above. The marginal confidence level is \(.6206\), which indicates that the sample does not contain strong evidence against the null hypothesis.

Table 5 reports estimates with the seasonally unadjusted data with \(n=8\). In this case, the likelihood function calls for high values of \(\delta_1\) and \(\mu_1\). The likelihood ratio statistic is now distributed asymptotically as chi-square with fifteen degrees of freedom under the null hypothesis that the model is correct. Once again the likelihood ratio statistic does not call for rejecting the model.

Table 6 shows the estimates obtained for the seasonally adjusted data for \(n=8\). It is interesting that with \(n=8\) the high \(\delta_1, \mu_1\) estimates
are the ones that maximize the likelihood function. The likelihood ratio test again fails to reject the model.

I also estimated a single-shift version of the model in which all of the terms in $n_{2t}$ are dropped from the objective function (1). The result is a model restricting only $n_{1t}$ and $w_t$ and consisting of equations (20a) and (20c). The right side of (20a) and (20c) contain $2n + 2$ regressors, while the model possesses the $(n+3)$ free parameters $d, f_1, \rho_1, v_1, \ldots, v_n$. So the model places $q = (2n+2) - (n+3)$ overidentifying restrictions on the bivariate vector autoregression (20a), (20c).

Tables 7 and 8 report the parameter estimates for the seasonally adjusted and unadjusted data. The seasonally adjusted data indicate that the overidentifying restrictions are marginally to be rejected at the .95 confidence level, while they are marginally rejected at the .90 confidence level for the seasonally unadjusted data. The parameter point estimates continue to indicate that adjustment costs exert an important influence on the demand for employment.
Conclusions

The simple contemporaneous correlations that formed the evidence in the original Dunlop-Tarshis-Keynes exchange, and also in much of the follow-up empirical work done to date, are not sufficient to rule on the question of whether the time series are compatible with a model in which firms are always on their demand schedules for employment. This is true according to virtually any dynamic and stochastic theory of the demand for employment. In this paper, I have tried to indicate one way in which the time series evidence can be brought to bear on the question in the context of a simple dynamic, stochastic model. The empirical results are moderately comforting to the view that the employment-real wage observations lie along a demand schedule for employment. It is important to emphasize that this view has content (i.e., imposes overidentifying restrictions) because I have a priori imposed restrictions on the orders of the adjustment cost processes and on the Markov processes governing disturbances. At a general level without such restrictions, it is doubtful whether the equilibrium view has content.
Table 1  
Vector Autoregressions, Seasonally Adjusted Data (n=4)

\[
\begin{array}{ccc}
\text{(20a)} & \text{Unconstrained} & \text{Constrained by (20)} \\
& & \\
\n_{t-1} & 1.5128 & 1.5159 \\
\n_{t-2} & -.5372 & -.5422 \\
\w_{t-1} & -2.0287 & .0225 \\
\w_{t-2} & -1.9667 & .0512 \\
\w_{t-3} & 2.9944 & .0969 \\
\w_{t-4} & -1.0538 & -.0343 \\
\end{array}
\]

\[
\begin{array}{ccc}
\text{(20b)} & & \\
& & \\
\w_{2t-1} & .9667 & .9730 \\
\w_{2t-2} & -.1596 & -.1534 \\
\w_{t-1} & -5.6155 & -.8117 \\
\w_{t-2} & .5847 & -.0297 \\
\w_{t-3} & 1.9400 & -.0988 \\
\w_{t-4} & -4.8444 & .6114 \\
\end{array}
\]

\[
\begin{array}{ccc}
\text{(20c)} & & \\
& & \\
\w_{t-1} & .9307 & .9635 \\
\w_{t-2} & -.0290 & .0031 \\
\w_{t-3} & .1146 & .0674 \\
\w_{t-4} & -.1907 & -.1744 \\
\end{array}
\]

\[|V| = 5.113199E-05 \quad T\{\log|V_r| - \log|V_u|\} = 7.3172 \]

Marginal confidence level = .6034

Let X be a random variable distributed chi-square seven degrees of freedom, and let x be the value of the computed test statistic. Then the marginal confidence level is defined as \(\text{Prob}\{X < x\}\) under the null hypothesis.
Table 2
Parameter Estimates
Seasonally Adjusted Data (n=4)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f_1$</td>
<td>0.2794</td>
</tr>
<tr>
<td>$d$</td>
<td>31.4283</td>
</tr>
<tr>
<td>$e$</td>
<td>1.4429</td>
</tr>
<tr>
<td>$\delta_1$</td>
<td>0.5782</td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>-1.3781</td>
</tr>
<tr>
<td>$\alpha_2$</td>
<td>0.0580</td>
</tr>
<tr>
<td>$\alpha_3$</td>
<td>0.1098</td>
</tr>
<tr>
<td>$\alpha_4$</td>
<td>0.2929</td>
</tr>
<tr>
<td>$\rho_1$</td>
<td>0.9377</td>
</tr>
<tr>
<td>$\rho_2$</td>
<td>0.7751</td>
</tr>
<tr>
<td>$\mu_1$</td>
<td>0.1979</td>
</tr>
</tbody>
</table>

$V = \begin{pmatrix}
0.9291E-01 & 0.2011E+00 & 0.1294E-02 \\
0.7746E+00 & 0.2089E-02 & \\
0.1939E-03 & 
\end{pmatrix}$

$\mathbf{B}^{-1} \mathbf{V} \mathbf{B}^{-1}' = \begin{pmatrix}
0.0968E+00 & 0.2106E+00 & 0.1561E-02 \\
0.7960E+00 & 0.5382E-03 & \\
0.1939E-03 & 
\end{pmatrix}$
Table 3
Seasonally Adjusted Data—Second Solution
Likelihood Equations

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f_1$</td>
<td>.03402</td>
</tr>
<tr>
<td>$d$</td>
<td>2367.94</td>
</tr>
<tr>
<td>$e$</td>
<td>10.4558</td>
</tr>
<tr>
<td>$\rho_1$</td>
<td>.5426</td>
</tr>
<tr>
<td>$\rho_2$</td>
<td>.1552</td>
</tr>
<tr>
<td>$\delta_1$</td>
<td>.9910</td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>-3.4140</td>
</tr>
<tr>
<td>$\alpha_2$</td>
<td>-.3800</td>
</tr>
<tr>
<td>$\alpha_3$</td>
<td>-.3483</td>
</tr>
<tr>
<td>$\alpha_4$</td>
<td>-.7619</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>-6.6658</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>.1880</td>
</tr>
<tr>
<td>$\beta_3$</td>
<td>.3530</td>
</tr>
<tr>
<td>$\beta_4$</td>
<td>-.3053</td>
</tr>
</tbody>
</table>

$$V = \begin{pmatrix} .9491E-01 & .2053E+00 & .1242E-02 \\ .7701E+00 & .1982E-02 & .1999E-03 \end{pmatrix}$$

$$B^{-1}V\mathbf{B}^{-1'} = \begin{pmatrix} .10572E+00 & .2249E+00 & .19241E-02 \\ .8054E+00 & .3315E-02 & .1999E-03 \end{pmatrix}$$

$$|V_r| = .56353E-05, |V_u| = .511320E-05$$

$$T\{\log|V_r| - \log|V_u|\} = 9.7221$$

Marginal confidence level = .7951
Table 4
Seasonally Unadjusted Data (n=4)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f_1$</td>
<td>.4358</td>
</tr>
<tr>
<td>$d$</td>
<td>3266.29</td>
</tr>
<tr>
<td>$e$</td>
<td>75.6748</td>
</tr>
<tr>
<td>$\rho_1$</td>
<td>.3849</td>
</tr>
<tr>
<td>$\rho_2$</td>
<td>.0909</td>
</tr>
<tr>
<td>$\delta_1$</td>
<td>.9512</td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>-.0110</td>
</tr>
<tr>
<td>$\alpha_2$</td>
<td>.0091</td>
</tr>
<tr>
<td>$\alpha_3$</td>
<td>.0114</td>
</tr>
<tr>
<td>$\alpha_4$</td>
<td>.0124</td>
</tr>
<tr>
<td>$v_1$</td>
<td>.8498</td>
</tr>
<tr>
<td>$v_2$</td>
<td>.1215</td>
</tr>
<tr>
<td>$v_3$</td>
<td>-.0219</td>
</tr>
<tr>
<td>$v_4$</td>
<td>-.1198</td>
</tr>
</tbody>
</table>

$$V = \begin{pmatrix} .1414E+00 & .2688E+00 & .1234E-02 \\ .8189E+00 & .1631E-02 \\ .2340E-03 \end{pmatrix}$$

$$B^{-1}V B^{-1} = \begin{pmatrix} .1414E+00 & .2695E+00 & .1237E-02 \\ .8208E+00 & .1762E-02 \\ .2340E-03 \end{pmatrix}$$

$|V_r| = .96438, |V_u| = .89475E-05$  
$T\{\log|V_r| - \log|V_u|\} = 7.4934$  
Marginal confidence level = .6206
Table 5
Seasonally Unadjusted Data (n=8)

\[ f_1 = .3721 \quad v_1 = .8388 \quad v_5 = .2718 \]
\[ d = 3266.29 \quad v_2 = .1512 \quad v_6 = -.1369 \]
\[ e = 75.6750 \quad v_3 = .0204 \quad v_7 = -.0306 \]
\[ \rho_1 = .39930 \quad v_4 = -.2786 \quad v_8 = -.0340 \]
\[ \rho_2 = .0522 \]
\[ \delta_1 = .9560 \quad \mu_1 = .7651 \]
\[ \alpha_1 = -.0282 \quad \beta_1 = -.7194 \]
\[ \alpha_2 = .0045 \quad \beta_2 = -.0316 \]
\[ \alpha_3 = .0092 \quad \beta_3 = .0653 \]
\[ \alpha_4 = .0108 \quad \beta_4 = .1045 \]
\[ \alpha_5 = .0040 \quad \beta_5 = -.0567 \]
\[ \alpha_6 = .0121 \quad \beta_6 = .1176 \]
\[ \alpha_7 = .0094 \quad \beta_7 = .0633 \]
\[ \alpha_8 = .0095 \quad \beta_8 = .0650 \]

\[ v = \begin{pmatrix} .1370E+00 & .2756E+00 & .9515E-03 \\ .8473E+00 & .1321E-02 \\ .2157E-03 \end{pmatrix} \]

\[ B^{-1}V_{B^{-1}}' = \begin{pmatrix} .1370E+00 & .2764E+00 & .9576E-03 \\ .8494E+00 & .1476E-02 \\ .2157E-03 \end{pmatrix} \]

\[ |V_r| = .83442E-05, \quad |V_u| = .79754E-05 \]

\[ T\{\log|V_r|-\log|V_u|\} = 4.3394 \]

Marginal confidence level = .0036
Table 6
Seasonally Adjusted Data (n=8)

\[
\begin{align*}
\begin{array}{cccc}
 f_1 & = & .2274 & v_1 & = & .9117 & v_5 & = & .0484 \\
 d & = & 2367.87 & v_2 & = & .0727 & v_6 & = & .0665 \\
e & = & 67.3950 & v_3 & = & .1036 & v_7 & = & -.1518 \\
 \rho_1 & = & .5632 & v_4 & = & -.2535 & v_8 & = & .0337 \\
 \rho_2 & = & .1532 & & & & & & \\
 \delta_1 & = & .9608 & u_1 & = & .8044 & & & \\
 \alpha_1 & = & -.5162 & \beta_1 & = & -1.0803 & & & \\
 \alpha_2 & = & -.0784 & \beta_2 & = & -.0305 & & & \\
 \alpha_3 & = & -.0484 & \beta_3 & = & .0385 & & & \\
 \alpha_4 & = & .0004 & \beta_4 & = & .1624 & & & \\
 \alpha_5 & = & -.1304 & \beta_5 & = & -.0614 & & & \\
 \alpha_6 & = & -.1178 & \beta_6 & = & -.0280 & & & \\
 \alpha_7 & = & -.0948 & \beta_7 & = & .0351 & & & \\
 \alpha_8 & = & -.1822 & \beta_8 & = & -.1180 & & & \\
\end{array}
\end{align*}
\]

\[
V = \begin{pmatrix}
.9443E-01 & .2156E+00 & .1152E-02 \\
.8021E+00 & .2463E-02 \\
.1789E-03 \\
\end{pmatrix}
\]

\[
B^{-1}V B^{-1} = \begin{pmatrix}
.9567E-01 & .2183E+00 & .1244E-02 \\
.8076E+00 & .2656E-02 \\
.1789E-03 \\
\end{pmatrix}
\]

\[
|V_r| = .48162E-05, \ |V_u| = .440883E-05
\]

\[
T\{\log |V_r| - \log |V_u|\} = 8.484
\]

Marginal confidence level = .0971
Table 7
Seasonally Adjusted, One-Shift Model

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f_1$</td>
<td>0.0059</td>
</tr>
<tr>
<td>$d$</td>
<td>3.4108</td>
</tr>
<tr>
<td>$\rho_1$</td>
<td>0.7834</td>
</tr>
<tr>
<td>$\lambda_1$</td>
<td>0.7934</td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>26.3880</td>
</tr>
<tr>
<td>$\alpha_2$</td>
<td>1.0662</td>
</tr>
<tr>
<td>$\alpha_3$</td>
<td>1.2488</td>
</tr>
<tr>
<td>$\alpha_4$</td>
<td>0.2356</td>
</tr>
</tbody>
</table>

$$
V = \begin{pmatrix}
0.8359E-01 & 1204E-02 \\
0.1986E-03 & 
\end{pmatrix}
$$

$$
|V_r| = 0.15154E-04, \quad |V_u| = 0.13983E-04
$$

$$
T\{\log|V_r| - \log|V_u|\} = 8.0402
$$

Marginal confidence level = 0.9548
Table 8
Seasonally Unadjusted, One-Shift Model

\[ f_1 = 0.0144 \quad v_1 = 0.8476 \]
\[ \delta = 5.1839 \quad v_2 = 0.0634 \]
\[ \rho_1 = 0.7002 \quad v_3 = -0.0388 \]
\[ \lambda_1 = 0.7417 \quad v_4 = -0.0147 \]
\[ \alpha_1 = -13.3288 \]
\[ \alpha_2 = -0.1070 \]
\[ \alpha_3 = 0.6930 \]
\[ \alpha_4 = 0.4663 \]

\[ V = \begin{pmatrix} 1.267E+00 & 1.224E-02 \\ \cdot & \cdot \end{pmatrix} \]

\[ |V_r| = 0.28511E-04, \quad |V_u| = 0.26757E-04 \]

\[ T(\log|V_r| - \log|V_u|) = 6.3492 \]

Marginal confidence level = 0.9042
Applications of related methods are contained in Sargent [1977a, 1977b].

Restrictions on the production function required to permit Lucas's static model to account for the cyclical behavior of labor productivity and real average hourly earnings were discussed by Sargent and Wallace [1974]. Adding differential costs for adjusting straight-time and overtime labor would widen the class of production functions that could lead to procyclical movements of average hourly earnings and labor productivity.

Optimization problems of this form are discussed by Holt, Modigliani, Muth, and Simon [1960], Graves and Telser [1971], and Kwakernaak and Sivan [1972]. The treatment here closely follows that of Sargent [1977c]. It would be straightforward to carry along n firms, each facing the same wage process and operating under the same functional form for its objective function (1), yet each having different values for the parameters f_0, f_1, d, and e. It would then be straightforward to aggregate the Euler equations and their solutions (7). Thus, assuming a representative firm is only a convenience, as the model admits a tidy theory of aggregation.

See Sargent [1977c], Chapters IX and XIV.

See Sargent [1977c].

See Sargent [1977c]. The solution (7) clearly exhibits the certainty-equivalence or separation property. That is, the same solution for n_1t and n_2t would emerge if we maximized the criterion formed by replacing (a_{1t+j},a_{2t+j},w_{t+j}) by (E_t a_{1t+j},E_t a_{2t+j},E_t w_{t+j}) and dropping the operator E_t from outside the sum in (1).

The condition that E_t \omega_{t-1} = 0 is equivalent with the condition that w_t is not caused, in Granger's [1969] sense, by n_2 or n_1.

The assumption that w_t is of exponential order less than (\frac{1}{b}) implies that the max |\lambda_i| < \frac{1}{b} where \lambda_i is the i^th element of \Lambda.

Here we are using that \sum_{i=0}^{\infty} (\frac{1}{\mu_2})^i \rho_1^i a_{1t} = \frac{1}{1 - \frac{1}{\mu_2} \rho_1} since |\rho_1| < \frac{1}{b} and |\mu_2| > \frac{1}{b}, so that the infinite sum converges.

Engineers directly obtain solutions of the form (13) by solving matrix Ricatti equations, e.g., see Kwakernaak and Sivan [1972].
Footnotes, continued

In their jargon, our system is not "controllable" but is "stabilizable" and "detectable" so that convergence of iterations on the Ricatti equation is assured. The stabilizability of our system depends on \( \{a_{1t}\} \), \( \{a_{2t}\} \), and \( \{w_t\} \) being of exponential order less than \( \left( \frac{1}{b} \right) \).

\(^{11/}\) The parameters \( f_0 \) and \( v_0 \) are dropped because the data are in the form of deviations from means and trend terms. The parameters \( b \), \( h_1 \), and \( h_2 \) will be fixed \textit{a priori}.

\(^{12/}\) That these values for \( h_1 \) and \( h_2 \) do not add to unity, as in the theoretical presentation of the model, amounts only to a harmless renormalization.

\(^{13/}\) With the seasonally unadjusted employment data, I first regressed each of \( n_{1t} \), \( n_{2t} \), and \( w_t \) against a constant, trend, trend squared, and three seasonal dummies and used the residuals from those regressions as the data.
References


------------------


------------------


References, continued


Sargent, Thomas J. "Notes on Macroeconomic Theory," University of Minnesota, manuscript, 1977c, Chapters IX and XIV.


