PACS File #: 2520

Estimation of Dynamic Labor Demand Schedules Under Rational Expectations

Thomas J. Sargent

April 1978

Staff Report #: 27

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.

Robert Litterman most ably performed the rather involved calculations reported in this paper.

Abstract

A dynamic linear demand schedule for labor is estimated and tested. The hypothesis of rational expectations and assumptions about the orders of the Markov processes governing technology impose over-identifying restrictions on a vector autoregression for straight-time employment, overtime employment, and the real wage. The model is estimated by the full information maximum likelihood method. The model is used as a vehicle for re-examining some of the paradoxical cyclical behavior of real wages described in the famous Dunlop-Tarshis-Keynes exchange.

Journal of Political Economy

University of Minnesota, Minneapolis, Minnesota, U.S.A.

Both Keynes and various classical writers asserted that real wages would move countercyclically as employers moved along downward sloping demand 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 much subsequent evidence on the question, which is reviewed and extended by Bodkin [1969], consisted mostly of simple contemporaneous regressions between real wages and some measure of the stage of the business cycle. By and large that evidence was regarded as rejecting the view that the data can be described as observations falling along an aggregate demand schedule for employment. This view of the evidence in large measure stimulated 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.

This paper aims to provide a framework for reexamining some of this evidence within the context of a stochastic and dynamic aggregate demand schedule for labor. The old evidence is simply not decisive because the view that the aggregate data lie along the type of demand schedule considered in this paper places no restrictions on the simple contemporaneous regressions in the studies summarized by Bodkin [1969]; however, under certain conditions, that view does place restrictions on aggregate real wages and employment as a vector stochastic process. The plan of this paper is to extract and test these implications.

This paper starts from the findings of the recent paper by Salih Neftci [1977], which computed long two-sided distributed lags between aggregate employment and real wages for post-World War II data for the

U.S. Neftci found that there were complicated and economically significant dynamic interactions between real wages and employment and that there was much stronger evidence for Granger [1969] causality flowing from real wages to employment than for Granger causality in the reverse direction. Further, the influence of real wages on employment was predominantly negative.

To represent Neftci's findings in a slightly different form than he did, Table 1 reports estimates of a fourth-order bivariate autoregression for quarterly aggregate measures of real wages w and employment n_1, both seasonally unadjusted. The theory of vector autoregressions and moving averages is reviewed briefly in the appendix. The data are a straight-time wage index in manufacturing divided by the consumer price index measured in 1967 dollars, and number of employees on nonagricultural payrolls, measured in millions of men. The data are described more in Section 3 below. The F-statistic pertinent for testing the null hypothesis that lagged real wages have zero coefficients in the vector autoregression for employment has a marginal significance level of .091. The F-statistic pertinent for testing the hypothesis that lagged levels of employment have zero coefficients in the vector autoregression for the real wage has a marginal significance level of .869.[1/] This pattern is consistent with Neftci's finding much stronger evidence of Granger causality extending from real wages to employment than in the other direction.

Table 2 reports estimates of the moving average representation implied by the autoregressions in Table 1. Table 2 depicts the matrix of responses to one standard deviation innovations in the real wage and employment, respectively. A one standard deviation innovation in employment leads to a strong, sustained increase in employment and a small (relative to the response to its own innovation) sustained decrease in the real wage. A one standard deviation innovation in the real wage leads to a sustained and sizable decrease in employment and a sustained and sizable increase in the real wage. The response of employment to the real wage innovation is of the same order of magnitude as it is to its own innovation, in contrast to the response of the real wage to the employment innovation. The magnitude of the estimated response of employment to real wage innovations seems of substantial economic significance. Notice how, qualitatively, the real wage innovation acts like a disturbance to labor supply while the employment innovation acts as we would expect an innovation to the demand for labor to act.

Tables 3 and 4 report two alternative decompositions of the variances of the k-step ahead forecast errors of the (n_1, w) process into parts attributable to variance in the "orthogonalized innovations" in employment and the real wage. As indicated in the appendix, these decompositions are not unique, which accounts for the two tables. However, since the innovations in employment and the real wage in Table 1 have only a moderate correlation of .2442, the differences between the decompositions in Tables 2 and 3 are bound to be modest, as they are. The tables reveal that a substantial percentage (40 or 48) of the 35-quarter ahead forecast error variance in employment (which approximates the steady-state variance in the indeterministic part of employment) is accounted for by innovations in the real wage. Only a small percentage (1 or 6) of the 35-quarter ahead forecast error variance in the real wage is accounted for by the innovation in employment.

Two characteristics of these results are particularly important for purposes of this study. First, there do appear to be some complicated dynamic interactions between aggregate employment and these real wage data that might be susceptible to analysis with a dynamic model of the demand for employment. Second, these data seem to be consistent with the assumption that the real wage is not Granger-caused by employment. This assumption, which will be imposed below, substantially simplifies the modeling task.

The plan of this paper is to estimate a dynamic aggregative demand schedule for employment for postwar 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 a vector of stochastic processes composed of employment, a measure of overtime employment, and the real wage. The aim is to test the adequacy of these overidentifying restrictions.

The model is formed by blending the costly adjustment model of Lucas [1967], Treadway [1969], and Gould [1968] 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.[2/]

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.[3/] 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

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

$$t=0,\ 1,\ 2,\ 3,\ \ldots$$

$$\tau\varepsilon[0,1).$$

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 τ indexes moments within the day. The firm is assumed to have a constant capital stock over the day so that

$$k(t+\tau) = k(t) \equiv k_t \text{ for } \tau\varepsilon[0,1).$$

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 pw_t, where $p \approx 1.5$ 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 pw_t. Confronted with these market opportunities it is optimal for the firm to choose to set $n(t+\tau) = n_{1t}$ for $\tau\varepsilon[0,h_1]$ and $n(t+\tau) = n_{2t}$ for $\tau\varepsilon(h_1,1)$. 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 and write instantaneous output on the first and second shifts as

$$f(n_{1t},k) = (f_0+a_{1t})n_{1t} - \frac{f_1}{2}n_{1t}^2$$

$$f(n_{2t},k) = (f_0+a_{2t})n_{2t} - \frac{f_1}{2}n_{2t}^2$$

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. I assume that $Ea_{1t} = Ea_{2t} = 0$. 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 $d >> e$. 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 $pw_t h_2 n_{2t}$.

The firm chooses contingency plans for n_{1t} and n_{2t} to maximize its expected real present value[4/]

$$(1) \qquad v_t = E_t \sum_{j=0}^{\infty} b^j [(f_0+a_{1t+j}-w_{t+j})h_1 n_{1t+j} - \frac{f_1}{2} h_1 n_{1t+j}^2 - \frac{d}{2}(n_{1t+j}-n_{1t+j-1})^2 + (f_0+a_{2t+j}-pw_{t+j})h_2 n_{2t+j} - \frac{f_1}{2} h_2 n_{2t+j}^2 - \frac{e}{2}(n_{2t+j}-n_{2t+j-1})^2]$$

$$f_0,\ f_1,\ d,\ e > 0,\ \ p > 1,\ \ 0 < b < 1$$

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 $E_t x \equiv Ex|\Omega_t$ where x is a random variable, E is the mathematical expectation operator, and Ω_t is an information set available to the firm at time t. I assume that Ω_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_1 and n_2 from the set of stochastic processes that are (nonanticipative) functions of the information set Ω_t. (Below, I will further restrict the class of stochastic processes over which the optimization is carried out.) I assume that the stochastic processes w_t, a_{1t}, and a_{2t} are of exponential order less than $(\frac{1}{b})$, which means that for some K > 0 and some x such that $1 \leq x < 1/b$,

$$|E_t w_{t+j}| < K(x)^{j+t}$$

$$|E_t a_{1t+j}| < K(x)^{j+t}$$

$$|E_t a_{2t+j}| < K(x)^{j+t}$$

for all t and all $j \geq 0$. I further assume that all random variables have finite first- and second-order moments.

First-order necessary conditions for the maximization of (1) consist of a set of "Euler equations" and a pair of transversality conditions.[5/] The Euler equations for $\{n_{1t}\}$ and $\{n_{2t}\}$ are

$$bE_{t+j}n_{1t+j+1} + \phi_1 n_{1t+j} + n_{1t+j-1} = \frac{h_1}{d}(w_{t+j} - a_{1t+j} - f_0)$$

$$j=0,\ 1,\ 2,\ \ldots$$

(2)

$$bE_{t+j}n_{2t+j+1} + \phi_2 n_{2t+j} + n_{2t+j-1} = \frac{h_2}{e}(pw_{t+j} - a_{2t+j} - f_0)$$

$$j=0,\ 1,\ 2,\ \ldots$$

where

$$\phi_1 = -(\frac{f_1 h_1}{d} + (1+b))$$

(3)

$$\phi_2 = -(\frac{f_2 h_2}{e} + (1+b)).$$

The transversality conditions are

$$(4) \qquad \lim_{T\to\infty} b^T E_t n_{1t+T} = \lim_{T\to\infty} b^T E_t n_{2t+T} = 0.$$

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

$$(5) \qquad (1 + \frac{\phi_1}{b}z + \frac{1}{b}z^2) = (1-\delta_1 z)(1-\delta_2 z)$$

$$(6) \qquad (1 + \frac{\phi_2}{b}z + \frac{1}{b}z^2) = (1-\mu_1 z)(1-\mu_2 z).$$

Given the assumptions about the signs and magnitudes of the parameters composing b, ϕ_1, and ϕ_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 6/

$$(7) \quad \text{(a)} \quad n_{1t} = \delta_1 n_{1t-1} - \frac{\delta_1 h_1}{d} \sum_{i=0}^{\infty} (\frac{1}{\delta_2})^i E_t(w_{t+i} - a_{1t+i} - f_0)$$

$$\text{(b)} \quad n_{2t} = \mu_1 n_{2t-1} - \frac{\mu_1 h_2}{e} \sum_{i=0}^{\infty} (\frac{1}{\mu_2})^i E_t(pw_{t+i} - a_{2t+i} - f_0).$$

It can be verified directly that these solutions satisfy the Euler equations and the transversality conditions. The polynomial equation (5) implicitly defines δ_1 and δ_2 as functions of $\frac{f_1 h_1}{d}$. By studying this polynomial,7/ it is possible to show that δ_1 is a decreasing function of $\frac{f_1 h_1}{d}$ and that $\frac{1}{\delta_2} = b\delta_1$. It follows that δ_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 δ_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, μ_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 Ω_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 Ω_t.

For the purposes of empirical work, it is convenient to restrict ourselves to the class of decision rules that are linear functions of Ω_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 Ω_t. Accordingly, henceforth, in all forecasting formulas, I will replace the mathematical expectation operator E by the linear least squares projection operator $\hat{E}$.[8/]

To derive from (7) explicit decision rules for n_{1t} and n_{2t} as functions of Ω_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

$$\hat{E}_t a_{1t+i} = \rho_1^i a_{1t} \qquad i \geq 0$$

(8)

$$\hat{E}_t a_{2t+i} = \rho_2^i a_{2t} \qquad i \geq 0$$

where $|\rho_1| < \frac{1}{b}$, $|\rho_2| < \frac{1}{b}$. That is, I assume that a_{1t} and a_{2t} are generated by the stochastic processes

$$a_{1t} = \rho_1 a_{1t-1} + \tilde{\xi}_{1t}$$

(9)

$$a_{2t} = \rho_2 a_{2t-1} + \tilde{\xi}_{2t}$$

where $\tilde{\xi}_{1t}$ and $\tilde{\xi}_{2t}$ are least squares residuals with finite variances and $\hat{E}\tilde{\xi}_{1t}|\Omega_{t-1} = \hat{E}\tilde{\xi}_{2t}|\Omega_{t-1} = 0$. Although (9) permits $\tilde{\xi}_{1t}$ and $\tilde{\xi}_{2t}$ to be arbitrarily correlated contemporaneously, it does in effect rule out correlation between them at any nonzero lags. I assume that w_t is an n^{th}-order Markov process

(10) $$w_t = v_0 + v_1 w_{t-1} + v_2 w_{t-2} + \dots + v_n w_{t-n} + \xi_{3t}$$

where ξ_{3t} is a least squares disturbance that satisfies $\hat{E}_{t-1}\xi_{3t} \equiv \hat{E}\xi_{3t}|\Omega_{t-1} = 0$. The condition that $\hat{E}\xi_{3t}|\Omega_{t-1} = 0$ means that ξ_{3t} is serially uncorrelated <u>and</u> that w_t is not caused in Granger's [1969] sense, by n_1 or n_2. That the lack of Granger causality from n_1 or n_2 to w is a workable approximation for the data to be studied here is supported by the empirical results of Neftci [1977] which are summarized above. It is convenient to represent the n^{th}-order process (10) as the (n+1)-vector first-order Markov process

$$x_t = Ax_{t-1} + \varepsilon_t$$

where

$$x_t = \begin{bmatrix} w_t \\ w_{t-1} \\ w_{t-2} \\ \cdot \\ \cdot \\ \cdot \\ w_{t-n} \\ 1 \end{bmatrix} \qquad \varepsilon_t = \begin{bmatrix} \xi_{3t} \\ 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \\ 0 \end{bmatrix}$$

$$A = \begin{bmatrix} v_1 & v_2 & \cdots & v_n & v_0 \\ 1 & 0 & & 0 & 0 \\ 0 & 1 & & 0 & 0 \\ \vdots & \vdots & & & \\ 0 & 0 & \cdots 1 & 0 & 0 \\ 0 & 0 & \cdots 0 & 0 & 1 \end{bmatrix} .$$

We can write

$$x_{t+1} = Ax_t + \varepsilon_{t+1}$$

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

$$\vdots$$

$$x_{t+j} = A^j x_t + \varepsilon_{t+j} + A\varepsilon_{t+j-1} + \ldots A^{j-1}\varepsilon_{t+1}.$$

Since $\hat{E}_t \varepsilon_{t+k} = 0$ for $k \geq 1$, we have

$$\hat{E}_t x_{t+j} = A^j x_t.$$

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 Λ is the diagonal matrix whose elements are the eigenvalues of A.[9/] Then we have

$$\hat{E}_t x_{t+j} = P\Lambda^j P^{-1} x_t.$$

Finally, let c be the 1x(n+1) row vector (1,0,0,...,0) so that $w_t = cx_t$. We thus have that

(11) $$\hat{E}_t w_{t+j} = cP\Lambda^j P^{-1} x_t.$$

Substituting from (8) and (11) into (7a) gives[10/]

$$n_{1t} = \delta_1 n_{1t-1} - \frac{\delta_1 h_1}{d} cP \sum_{i=0}^{\infty} (\frac{1}{\delta_2}\Lambda)^i P^{-1} x_t$$

$$+ \frac{\delta_1 h_1}{d}(\frac{f_0}{1 - \frac{1}{\delta_2}}) + \frac{\delta_1 h_1}{d}(\frac{1}{1 - \frac{\rho_1}{\delta_2}}) a_{1t}.$$

Let λ_i be the iith element of Λ. Since $\delta_2 = \frac{1}{\delta_1 b}$, we have that $|\frac{\lambda_i}{\delta_2}| = |\lambda_i \delta_1 b| < 1$ by virtue of the assumption that w_t is of exponential order less than $\frac{1}{b}$, i.e., that $|\lambda_i \cdot b| < 1$. Then the infinite sum above converges and we can write

(12) $$n_{1t} = \delta_1 n_{1t-1} - \frac{\delta_1 h_1}{d} cP [\frac{1}{1 - \frac{\lambda_i}{\delta_2}}]_{ii} P^{-1} x_t$$

$$+ \frac{\delta_1 h_1}{d}(\frac{f_0}{1 - \frac{1}{\delta_2}}) + \frac{\delta_1 h_1}{d}(\frac{1}{1 - \frac{\rho_1}{\delta_2}}) a_{1t}$$

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

Let us write (12) as[11/]

(13) $$n_{1t} = \delta_1 n_{1t-1} + \alpha_1 w_t + \alpha_2 w_{t-1} + \ldots + \alpha_n w_{t-n+1} + \alpha_0$$

$$+ \frac{\delta_1 h_1}{d}(\frac{f_0}{1-\delta_1 b}) + a'_{1t}$$

where

$$(14) \qquad (\alpha_1, \alpha_2, \ldots, \alpha_n, \alpha_0) = -\frac{\delta_1 h_1}{d} cP\left[\frac{1}{1-\lambda_i \delta_1 b}\right]_{ii} P^{-1}$$

$$a'_{1t} = \frac{\delta_1 h_1}{d}\left(\frac{1}{1-\rho_1 \delta_1 b}\right) a_{1t}.$$

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

$$(15) \qquad n_{2t} = \mu_1 n_{2t-1} + \beta_1 w_t + \beta_2 w_{t-1} + \ldots + \beta_n w_{t-n+1} + \beta_0$$

$$+ \frac{\mu_1 h_2}{e}\left(\frac{f_0}{1-\mu_1 b}\right) + a'_{2t}$$

where

$$(16) \qquad (\beta_1, \beta_2, \ldots, \beta_n, \beta_0) = -p\frac{\mu_1 h_2}{e} cP\left[\frac{1}{1-\lambda_i \mu_1 b}\right] P^{-1}$$

$$a'_{2t} = \frac{\mu_1 h_2}{e}\left(\frac{1}{1-\rho_2 \mu_1 b}\right) a_{2t}.$$

Equations (14) and (16) succinctly summarize how the distributed lag coefficients, the α's and β's, reflect the combination of forecasting (through the parameters of P and Λ) and optimization (through the parameters d, δ, and μ) elements. Clearly, the decision rules (13) and (15) are not invariant with respect to changes in the stochastic process for real wages (10), 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

(17) $$n_{1t} = (\delta_1+\rho_1)n_{1t-1} - \rho_1\delta_1 n_{1t-2} + (\alpha_2+\alpha_1 v_2-\alpha_2\rho_1)w_{t-1}$$

$$+ (\alpha_3+\alpha_1 v_2-\alpha_2\rho_1)w_{t-2} +\ldots+ (\alpha_n+\alpha_1 v_{n-1}-\alpha_{n-1}\rho_1)w_{t-n+1}$$

$$+ (\alpha_1 v_n-\alpha_n\rho_1)w_{t-n} + [\alpha_1\xi_{3t}+(a'_{1t}-\rho_1 a'_{1t-1})].$$

From our earlier assumptions, $\hat{E}_{t-1}[\alpha_1\xi_{3t}+(a'_{1t}-\rho_1 a'_{1t-1})] = 0$, so that (17) is the (vector) autoregression for n_{1t}. In particular, we have

(18) $$\hat{E}_{t-1}n_{1t} = (\delta_1+\rho_1)n_{1t-1} - \rho_1\delta_1 n_{1t-2} + (\alpha_2+\alpha_1 v_2-\alpha_2\rho_1)w_{t-1}$$

$$+ (\alpha_3+\alpha_1 v_2-\alpha_2\rho_1)w_{t-2} +\ldots+ (\alpha_n+\alpha_1 v_{n-1}-\alpha_{n-1}\rho_1)w_{t-n+1}$$

$$+ (\alpha_1 v_n-\alpha_n\rho_1)w_{t-n}.$$

Similarly, we have for n_{2t}

(19) $$n_{2t} = (\mu_1+\rho_2)n_{2t-1} - \rho_2\mu_1 n_{2t-2} + (\beta_2+\beta_1 v_2-\beta_2\rho_2)w_{t-1}$$

$$+ (\beta_3+\beta_1 v_2-\beta_2\rho_2)w_{t-2} +\ldots+ (\beta_n+\beta_1 v_{n-1}-\beta_{n-1}\rho_2)w_{t-n+1}$$

$$+ (\beta_1 v_n-\beta_n\rho_2)w_{t-n} + [\beta_1\xi_{3t}+(a'_{2t}-\rho_2 a'_{2t-1})].$$

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

(a) $$n_{1t} = (\delta_1+\rho_1)n_{1t-1} - \rho_1\delta_1 n_{1t-2} + (\alpha_2+\alpha_1 v_2-\alpha_2\rho_1)w_{t-1}$$

$$+ (\alpha_3+\alpha_1 v_2-\alpha_2\rho_1)w_{t-1} +\ldots+ (\alpha_n+\alpha_1 v_{n-1}-\alpha_{n-1}\rho_1)w_{t-n+1}$$

$$+ (\alpha_1 v_n-\alpha_n\rho_1)w_{t-n} + u_{1t}$$

(20) (b) $$n_{2t} = (\mu_1+\rho_2)n_{2t-1} - \rho_2\mu_1 n_{2t-2} + (\beta_2+\beta_1 v_2-\beta_2\rho_2)w_{t-1}$$

$$+ (\beta_3+\beta_1 v_2-\beta_2\rho_2)w_{t-2} +\ldots+ (\beta_n+\beta_1 v_{n-1}-\beta_{n-1}\rho_2)w_{t-n+1}$$

$$+ (\beta_1 v_n-\beta_n\rho_2)w_{t-n} + u_{2t}$$

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

where

$$u_t \equiv \begin{bmatrix} u_{1t} \\ u_{2t} \\ u_{3t} \end{bmatrix} \equiv \begin{bmatrix} \alpha_1 \xi_{3t} + (a'_{1t} - \rho_1 a'_{1t-1}) \\ \beta_1 \xi_{3t} + (a'_{2t} - \rho_2 a'_{2t-1}) \\ \xi_{3t} \end{bmatrix} = \begin{bmatrix} n_{1t} - \hat{E}_{t-1} n_{1t} \\ n_{2t} - \hat{E}_{t-1} n_{2t} \\ w_t - \hat{E}_{t-1} w_t \end{bmatrix} .$$

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., $w_{t-1}, \ldots, w_{t-n}$, each 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, ρ_1, ρ_2, $v_1, \ldots, v_n$, so that there are (n+5) parameters to be estimated. As it turns out, the model is overidentified for $n \geq 1$.

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

$$(21) \quad \begin{cases} \phi_1 = -\left(\frac{f_1 h_1}{d} + (1+b)\right) \\ \phi_2 = -\left(\frac{f_1 h_2}{e} + (1+b)\right) \\ \left(1 + \frac{\phi_1}{b} z + \frac{1}{b} z^2\right) = (1-\delta_1 z)(1-\delta_2 z) \\ \left(1 + \frac{\phi_2}{b} z + \frac{1}{b} z^2\right) = (1-\mu_1 z)(1-\mu_2 z) \\ (\alpha_1, \alpha_2, \ldots, \alpha_n, \alpha_0) = -\frac{\delta_1 h_1}{d} cP\left[\frac{1}{1-\lambda_i \delta_1 b}\right]_{ii} P^{-1} \\ (\beta_1, \beta_2, \ldots, \beta_n, \beta_0) = -p\frac{\mu_1 h_2}{e} cP\left[\frac{1}{1-\lambda_i \mu_1 b}\right]_{ii} P^{-1} \\ A = P \Lambda P^{-1}. \end{cases}$$

Estimates of the free parameters $\theta = (f_1, d, e, \rho_1, \rho_2, v_1, \ldots, v_n)$ are obtained by using the method of maximum likelihood to estimate the vector autoregression (20), subject to (21).[12/] Let $\hat{u}_t = (\hat{u}_{1t}, \hat{u}_{2t}, \hat{u}_{3t})'$ be the sample residual vector associated with the parameter values θ. Under the assumption that u_t is a trivariate normal vector with $Eu_t u_t' = V$, the likelihood function of a sample of observations on the residuals extending over t=1, ..., T is

$$(22) \qquad L(\theta) = (2\pi)^{-\frac{1}{2}3T} |V|^{-\frac{T}{2}} \exp(-\frac{1}{2} \sum_{t=1}^{T} \hat{u}_t' V^{-1} \hat{u}_t).$$

As shown by Wilson [1973] and Bard [1974], maximum likelihood estimates of θ with V unknown can be obtained by minimizing $|\hat{V}|$ with respect to θ, 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], (p. 62-66)).[13/] The value of the likelihood function turns out to be

$$\log L(\theta) = -(1/2)mT \log (2\pi) - (1/2)T \{\log|\hat{V}|+m\}$$

where m is the number of variates, equal to three in the present model.

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

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

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

I also used a likelihood ratio statistic to test the constrained vector autoregression (20) against a second and even less constrained alternative, namely, an unconstrained trivariate vector autoregression with n lagged values of n_1, n_2, and w on the right side of each equation. Let $\tilde{V}_u$ be the estimated sample covariance matrix of the residuals in the unrestricted vector autoregression. Then the appropriate likelihood ratio statistic is given by $T\{\log_e|\hat{V}_r|-\log_e|\tilde{V}_u|\}$. Since the unconstrained parameterization now has 9n free parameters, the likelihood ratio is asymptotically distributed as chi-square with {9n-(n+5)} degrees of freedom.

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 paralleling 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 $Ew_t a_{1s} = Ew_t a_{2s} = 0$ 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 [1978a] 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_{js}'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 $\hat{E}[w_t | w_{t-1}, w_{t-2}, \ldots, n_{1t-1}, n_{1t-2}, \ldots, n_{2t-1}, n_{2t-2}, \ldots] = \hat{E}[w_t | w_{t-1}, w_{t-2}, \ldots]$. 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 [1977a] 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 those correlations. On the other hand, the alternative "projection estimator" attributes all of those correlations 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 correlations 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 extending from 1947I through 1972IV, of which n observations at the beginning of the sample are lost when the order of the wage autoregression is set at n. 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 $\bar{h}_t$ 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 $\bar{h}_t n_{1t}$. 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.[14/] 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.

For reasons developed in Sargent [1978b], I would argue that seasonally unadjusted data are the ones that ought to be used. Briefly, this view follows from the assumption that agents are themselves observing and responding to the seasonally unadjusted variates, so that the cross-equation restrictions delivered by the model pertain to the seasonally unadjusted data. Seasonal adjustment of the data could cause rejection of the cross-equation rational expectations restrictions when they are in fact true. However, arguments have been made against this position in advocacy of seasonally adjusted data in exactly the present context (see Sims [1976]). For this reason, I report some results for both seasonally adjusted and unadjusted data.

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.[15/] 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 <u>not</u> 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 <u>1967</u> dollars, while employment is measured in millions of men.

Table 5 reports estimates of the model for n=4 for the seasonally adjusted data. The free parameters were f_1, d, e, ρ_1, ρ_2, v_1, v_2, v_3, and v_4 with b being fixed at .95, h_1 at 37, h_2 at 17, and the premium p at 1.5. Since n=4, for the more constrained of our two alternative hypotheses, the likelihood ratio statistic is asymptotically distributed as chi-square with q=(3n+4)-(n+5)=7 degrees of freedom. The likelihood ratio is 9.53 which has a "marginal confidence level" of .783. The marginal confidence level is defined as follows. Let X be a chi-square random variable with q degrees of freedom. Let x be the value of the likelihood ratio statistic. Then the marginal confidence level is defined as Prob{X<x} under the null hypothesis. High values of the confidence level lead to rejecting the hypothesis. The likelihood ratio statistic in this case indicates that the hypothesis can't be rejected at marginal significance levels below .20. However, versus the less constrained alternative hypothesis, the marginal confidence level is .9864, which indicates that the data do contain substantial evidence against the hypothesis. Notice the different lag shapes and the magnitudes of the distributed lag coefficients of straight-time employment and overtime employment in the real wage, the αs and βs, respectively. Overtime employment is estimated to be more responsive to the real wage. Further, the straight-time adjustment cost parameter d is estimated to be much larger than the overtime adjustment cost parameter e. That is why n_{1t} depends more strongly on n_{1t-1} than n_{2t} does on n_{2t-1}, i.e., why δ_1 is estimated to exceed μ_1.

Since the likelihood ratio test assumes that the u's are serially uncorrelated, Table 5 also reports three statistics KS(n_1), KS(n_2), and KS(w) which are Kolmogorov-Smirnov statistics from the cumulated periodograms for u_1, u_2, and u_3, that is, for the estimated innovations for n_1, n_2, and w, respectively, implied by the vector autoregression constrained by the model. The Kolmogorov-Smirnov statistic recorded is the maximum absolute deviation of the cumulated periodogram of the disturbance from its theoretical value under the assumption that the disturbances are serially uncorrelated. Durbin [1969] reports tables for the distribution of this statistic, though they are not applicable where lagged dependent variables are included as regressors, as in the present case. It is nevertheless of some comfort that the Kolmogorov-Smirnov statistics in Table 5 and in subsequent tables do not signal dangerous levels of serial correlation. Notice that the Kolmogorov statistics are greater for the n_1 and n_2 innovations than for the w innovation. This is symptomatic of the fact that the model fits an n^{th}-order Markov process in w but only permits two lagged own-values to enter the autoregressions for n_1 and n_2, thereby leaving it more likely that the model will neglect some higher-order serial correlation for n_1 and n_2. This pattern for the Kolmogorov-Smirnov statistics repeats itself in the subsequent tables.

Table 5 also reports the estimated covariance matrix of the innovations $V = Eu_t u_t'$. Recall that

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

where $\xi_{1t} = a_{1t}' - \rho_1 a_{1t-1}'$, $\xi_{2t} = a_{2t}' - \rho_2 a_{2t-1}'$, and where

$$B = \begin{bmatrix} 1 & 0 & \alpha_1 \\ 0 & 1 & \beta_1 \\ 0 & 0 & 1 \end{bmatrix}, \qquad \xi_t = \begin{bmatrix} \xi_{1t} \\ \xi_{2t} \\ \xi_{3t} \end{bmatrix}.$$

Then, since $\xi_t = B^{-1}u_t$, the covariance matrix of ξ_t can be estimated from $E\xi_t\xi_t' = B^{-1}VB^{-1'}$, an estimate of which is also reported in Table 5. The correlation between the innovations to a_{1t}' and to a_{2t}', i.e., ξ_{1t} and ξ_{2t}, is estimated to be .748. The correlation between the innovations to a_{1t}' and w_t, i.e., ξ_{1t} and ξ_{3t} is .3061, while that between ξ_{2t} and ξ_{3t} is .1700. I had expected ξ_{1t} and ξ_{2t} to be even more highly correlated than they are.

As it happens, the estimates reported in Table 5 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 6. In view of the form of 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_{1t-2}, n_{2t-1}, n_{2t-2} are, respectively, $(\delta_1+\rho_1)$, $-\delta_1\rho_1$, $(\mu_1+\rho_2)$, and $-\mu_1\rho_2$. <u>If it were not</u> for the constraints across μ_1 and the β's and across δ_1 and the α's and the appearance of ρ_1 and ρ_2 elsewhere on the right side of (20), the parameters δ_1, ρ_1, μ_1, and ρ_2 would not be identified, since it would be impossible to distinguish the effects of δ_1 from ρ_1 and the effects of μ_1 from ρ_2. The presence of lagged w'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 5 and 6 shows that Table 5 is a high (ρ_1,ρ_2) - low (δ_1,μ_1) solution, while Table 6 reports the high (δ_1,μ_1) - low (ρ_1,ρ_2) solution. Notice that for the Table 6 estimates, $\rho_1 + \delta_1 = 1.528$ and $\rho_1\delta_1 = .555$, while for the Table 5 estimates, $\rho_1 + \delta_1 = 1.526$ while $\rho_1\delta_1 = .552$.

Figures 1 and 2 depict two views of the likelihood surface as a function of δ_1 and ρ_1. The likelihood surface has a ridge and is characterized by two local maxima. Figure 3 depicts iso-likelihood contours in the (δ_1, ρ_1) of the identification of (δ_1, ρ_1) and of (u_1, ρ_2).

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.

Table 7 reports the estimates for the seasonally unadjusted data with n=4. The estimates indicate d >> e and are qualitatively similar to those described above. For testing the model versus the more constrained of the two alternative hypotheses, the marginal confidence level is .53. Versus the less constrained alternative, the marginal confidence level is .68. These results indicate that the sample does not contain strong evidence against the hypothesis.

Table 8 reports estimates of the model for the seasonally unadjusted data with n=8. The likelihood ratio statistic for testing against the more constrained alternative hypothesis is now distributed asymptotically as chi-square with fifteen degrees of freedom under the null hypothesis that the model is correct. Once again, both likelihood ratios indicate that the sample doesn't contain much evidence against the model. For the seasonally unadjusted data with n=8, Table 9 reports the maximum likelihood estimates of the vector autoregression (20), both unconstrained and constrained by the restrictions of the model (21). The constrained and unconstrained estimates are close except in one respect: the model-constrained vector autoregressions for n_1 and n_2 have coefficients on lagged w's that are generally much smaller in absolute value than their unconstrained counterparts. This pattern is also reflected in Tables 10 through 14. Table 10 shows the vector moving average representation implied by the model-constrained estimates while Table 11 shows a decomposition of variance of the 35-quarter ahead forecast error variances. Tables 12 and 13 show the corresponding moving average representation and decomposition of variance for the unconstrained estimates that are reported in Table 9. Comparison of Tables 10 and 12, on the one hand, and Tables 11 and 13, on the other hand, indicates that while the constrained model captures the same response of n_1 and n_2 to their own innovations that is depicted in the unconstrained estimates, the constrained model substantially underestimates the responses of n_1 and n_2 to innovations in w. The moving average representation implied by the model-constrained estimates have one standard deviation wage innovations giving rise to much smaller movements in n_1 and n_2 than are those associated with one standard deviation own-innovations in n_1 and n_2. Contrast this with the relatively sizable responses of n_1 and n_2 to real wage innovations in the unconstrained estimates. The decompositions of variance in Tables 11 and 13 indicate the extent to which the constrained model attributes less of a role to real wage innovations in driving n_1 and n_2.

Notice how both Tables 10 and 12 show n_2 responding more quickly to an own-innovation than does n_1.

The estimates in Tables 10-13 came from the data that are residuals from regressions on constant, trend, trend-squared, and three seasonal dummies. Table 14 is the counterpart of Table 13 where trend-squared has been omitted. The effect of dropping trend-squared is to increase somewhat the percentage of the variance of the 35-quarter ahead prediction error in n_1 or n_2 that is explained by innovations in the real wage. The results in Table 14 are presented to form a bridge to the estimates of Neftci and those summarized in the introduction, which included trend but not trend-squared terms.

The vector autoregresions summarized in Tables 9-14 all impose the extensive zero restrictions incorporated in (20), e.g., lagged n_2's don't appear in the autoregression for n_1. Tables 15 and 16 report summary statistics for fourth-order vector autoregressions with <u>no</u> such zero restrictions built in, i.e., four lags of each variable appear in the autoregression for each of n_1, n_2, and w. A constant, trend, and three seasonal dummies are also included in the regressions. Table 15 reports marginal significance levels appropriate for testing the null hypothesis that n_1 or n_2 or w fails to Granger-cause each of the other variables. These F-statistics are consistent with Neftci's results and indicate stronger evidence for Granger causality flowing from w to n_1 and n_2 than from n_1 or n_2 to w. However, the statistics also indicate Granger-causality from n_1 to n_2 and from n_2 to n_1, patterns which are ruled out by the model (20) and (21). The data indicate dynamic interactions between n_1 and n_2 that the model in its present form cannot account for. The decompositions of variance of 35-quarter ahead forecast errors in Table 16 once again reinforce Neftci's findings in confirming that substantial percentages of the variance in employment forecasting errors are attributable to innovations in the real wage.

In summary, while the model usually passes the likelihood ratio tests I have calculated, it does seem to do violence to two aspects of the data. First, the model generates estimates that seem to <u>understate</u> the magnitude of the inverse influence exerted by the real wage on employment. Second, <u>a priori</u> the model neglects dynamic interactions between

n_1 and n_2 that seem to be there. On the first point, the maximum likelihood estimates of the parameters d and e, which also influence the response to w of n_1 and n_2, respectively, seem mainly to have been chosen to permit the model to capture the response of n_1 and n_2 to their own innovations. As a by-product, this involved understating the responses of n_1 and n_2 to w, which seems less costly in terms of the likelihood function than misstating the response to own innovations. Perhaps a richer specification of the Markov processes for a_{1t} and a_{2t}, say permitting them to be second-order processes would permit enough flexibility to remedy this feature. Permitting the Markov processes for a_{1t} and a_{2t} to depend on lagged cross terms a_2 and a_1, respectively, would provide one way to remedy the second deficiency of the model, for it would potentially permit dynamic interactions between n_1 and n_2 of the kind revealed by Table 15. Another way to account for those dynamic interactions would be to let costs of adjustment for n_1 depend on the level of n_2, and vice versa. This could be done while remaining within the linear-quadratic framework of this paper. However, extensions in each of these directions, while feasible, are costly both in the sense that they reduce the degree of overidentification of the model and in the sense that they make maximum likelihood estimation more expensive.

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.

Footnotes

*Work on this paper was supported by the Federal Reserve Bank of Minneapolis, which is not responsible for the conclusions. Robert Litterman very ably performed the rather involved calculations reported in this paper. Helpful comments from a referee are gratefully acknowledged.

1/ For data on the left-side variable extending from 1951I-1972IV, which more closely matches Neftci's period than mine, the marginal significance level for testing the null hypothesis that real wages don't Granger-cause employment is .0745, and for the null hypothesis that employment doesn't Granger-cause the real wage is .5012. These autoregressions included constant, trend, and three seasonal dummies.

2/ Applications of related methods are contained in Sargent [1977a, 1978]. John Taylor [1978] uses a minimum distance estimator to estimate a macroeconomic model subject to rational expectations restrictions.

3/ 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.

4/ 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 [1977b]. It would be straightforward to carry along n firms, each facing the <u>same</u> wage process and operating under the <u>same</u> functional form for its objective function (1), yet each having <u>different</u> 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.

5/ See Sargent [1977b], Chapters IX and XIV.

6/ See Sargent [1977b].

7/ See Sargent [1977b]. 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).

8/ In the statistical literature the linear least squares projection operator $\hat{E}$ is often referred to as the "wide sense expectation" operator.

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

10/ Here I am using that $\left(\sum_{i=0}^{\infty} (\frac{1}{\delta_2})^i \rho_1^i \right) a_{1t} = \frac{1}{1 - \frac{\rho_1}{\delta_2}} a_{1t}$ since $|\rho_1| < \frac{1}{b}$ and $|\mu_2| > \frac{1}{b}$, so that the infinite sum converges.

11/ Engineers directly obtain solutions of the form (13) by solving matrix Ricatti equations, e.g., see Kwakernaak and Sivan [1972]. In their jargon, our system is <u>not</u> "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 $(\frac{1}{b})$.

12/ 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, p, h_1, and h_2 will be fixed <u>a priori</u>.

13/ The likelihood function was maximized by using a derivative-free hill climbing method with a Davidon-Fletcher-Powell algorithm for updating the Hessian. The complicated nature of the restrictions (21) led me to opt for a derivative-free method over an algorithm that required even analytical first deviatives. My attempts numerically to calculate asymptotic standard errors from the inverse of the information matrix were unsuccessful as one or two diagonal elements turned out to be negative.

14/ 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. I guessed at these values for h_1 and h_2. The guess for h_1 measured in hours per week seemed reasonable after having inspected the time series for average weekly hours. For purposes of constructing the data on n_{2t}, the choice of both h_1 and h_2 matters. For the purpose of estimating the demand functions, given the data on n_1 and n_2, only the ratio of h_1 to h_2 matters as proportional changes in d and e can cancel the effects of proportionate increases in h_1 and h_2.

15/ 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

Anderson, T. W. and John B. Taylor, "Conditions for Strong Consistency of Least Squares Estimates in Linear Models," August 1976. Technical Report No. 213, Institute for Mathematical Studies in the Social Sciences, Stanford University.

Bard, Yonathan. Nonlinear Parameter Estimation, New York, Academic Press: 1974.

Barro, R. J., and H. I. Grossman. "A General Disequilibrium Model of Income and Employment," American Economic Review, 61 (March 1971): 82-93.

Bodkin, R. G. "Real Wages and Cyclical Variations in Employment: A Reconsideration," Canadian Journal of Economics, 2 (August 1969): 353-374.

Dunlop, J. T. "The Movement of Real and Money Wage Rates," Economic Journal, 48 (September 1938): 413-434.

Durbin, J. "Tests for Serial Correlation in Residual Analysis Based on the Periodogram of Least-Squares Residuals," Biometrika, 56 (March 1969): 1-15.

Granger, C. W. J. "Investigating Causal Relations by Econometric Models and Cross-Spectral Methods," Econometrica, 37 (July 1969): 424-438.

Graves, R., and L. Telser. Functional Analysis in Mathematical Economics, Chicago: University of Chicago Press, 1971.

Gould, J. P. "Adjustment Costs in the Theory of Investment of the Firm," Review of Economic Studies, 35 (January 1969): 47-56.

Holt, C., F. Modigliani, T. F. Muth, and H. A. Simon. Planning Production, Inventories and Work Force, Englewood Cliffs, New Jersey, Prentice Hill: 1960.

Keynes, J. M. "Relative Movements of Real Wages and Output," Economic Journal, 59 (March 1939): 34-51.

Kwakernaak, H., and R. Sivan. Linear Optimal Control Systems, New York, Wiley: 1972.

Lucas, Robert E. "Econometric Policy Evaluation: A Critique." In The Phillips Curve and the Labor Market edited by Karl Brunner and Allan Meltzer, Vol. 1 of the Carnegie-Rochester Conferences on Public Policy, a supplementary series to the Journal of Monetary Economics, Amsterdam: North Holland, 1976.

_______________. "Adjustment Costs and the Theory of Supply," Journal of Political Economy, 75 (August 1967, Part I): 321-334.

_______________. "Capacity, Overtime, and Empirical Production Functions," American Economic Review Papers and Proceedings, 60 (May 1970): 23-27.

Ljung, L. "Consistency of the Least Squares Identification Method," IEEE Transactions on Automatic Control (October 1976): 779-781.

Phelps, Edmund S. "A Note on Short-Run Employment and Real Wage Rate Under Competitive Commodity Markets," International Economic Review, 10 (June 1969): 220-232.

Neftci, Salih N. "A Time Series Analysis of the Real Wages--Employment Relationship," Journal of Political Economy, forthcoming 1978.

Sargent, Thomas J., and Neil Wallace. "The Elasticity of Substitution and Cyclical Behavior of Productivity, Wages and Labor's Share," American Economic Review Papers and Proceedings, 64 (May 1974): 257-263.

Sargent, Thomas J. "Notes on Macroeconomic Theory," University of Minnesota, manuscript, 1977b, Chapters IX and XIV, forthcoming, Academic Press: 1978.

_______________. "The Demand for Money During Hyperinflations Under Rational Expectations: I," International Economic Review, 18 (February 1977a), 59-82.

_______________. "Rational Expectations, Econometric Exogeneity, and Consumption," Journal of Political Economy, 1978a.

_______________. "'Comment' on Seasonal Adjustment and Multiple Time Series Analysis by Kenneth Wallis." Presented at NBER/Bureau of Census, Conference on Seasonal Adjustment of Economic Time Series held in Washington, D.C., September 9-10, 1976. Proceedings of Conference to be published by Bureau of Census, 1978b.

Sims, Christopher A. "Money, Income, and Causality," American Economic Review, 62 (September 1972): 540-552.

_______________. "Response to Sargent's Comment on Wallis' Paper," manuscript, University of Minnesota, November 1976.

Solow, R. M., and J. E. Stiglitz. "Output, Employment, and Wages in the Short Run," Quarterly Journal of Economics, 82 (November 1968): 537-560.

Tarshis, L. "Changes in Real and Money Wage Rates," Economic Journal, 49 (March 1939): 150-154.

Taylor, John B. "Estimation and Control of a Macroeconomic Model with Rational Expectations," Columbia University, manuscript, February 1978.

Treadway, A. B. "On Rational Entrepreneurial Behavior and the Demand for Investment," Review of Economic Studies, 36 (April 1969): 227-240.

Wilson, G. T. "The Estimation of Parameters in Multivariate Time Series Models," Journal of the Royal Statistical Societies, B, 1973(1), 76-85.

Appendix on Vector Autoregressions and Moving Averages

Let x_t be an (nxl) vector jointly covariance stationary, linearly indeterministic stochastic process. The m^{th} order vector autoregression for this process is

(a) $$x_t = \alpha + \sum_{j=1}^{m} A_j x_{t-j} + \varepsilon_t^m$$

where ε_t^m is an (nxl) vector of least squares disturbances. Here α is an (nxl) vector and the A_j's are nxn matrices that under mild regularity conditions are uniquely determined by the population orthogonality conditions $E\varepsilon_t^m = 0$ and $E\varepsilon_t^m x'_{t-j} = 0_{nxn}$, j=1, 2, ..., m. The ε_t^m process is termed the process of innovations, the parts of x_t that can't be predicted linearly from m lagged x_t's; ε_t^m is the process of one-step ahead prediction errors. If $m=\infty$, the orthogonality conditions imply that $E\varepsilon_t\varepsilon'_{t-s} = 0$ for $s\neq 0$, having the practical implication that if m is taken to be big enough, as we shall assume, the ε_t^m vector is serially uncorrelated. If we solve the vector difference equation (a) for x_t backwards in terms of the ε process and ignore transient terms, we get the vector moving average representation

(b) $$x_t = \alpha' + \sum_{j=0}^{\infty} C_j \varepsilon_{t-j}^m$$

where α' is an (nxl) vector of constants, where C_j is an (nxn) matrix and $C_0 \equiv I$. The matrix Fourier transforms of the A_j's and C_j's are related by $(I - A_1 e^{-iw} - \ldots - A_m e^{-iwm})^{-1} = \sum_{j=0}^{\infty} C_j e^{-iwj}$. The (nxl) vector process ε_t^m is composed of disturbances that are mutually orthogonal at all nonzero lags and leads (by the orthogonality conditions), but $E\varepsilon_t^m \varepsilon_t^{m'} = \Sigma$ is not in general diagonal. To illustrate how to construct a moving average representation with a disturbance process that is orthogonal contemporaneously as well as at all lags, let n=2 and consider the transformation

$$\begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \rho & 1 \end{pmatrix} \begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix} = Fu_t$$

where $\rho = E\varepsilon_{1t}\varepsilon_{2t}/E\varepsilon_{1t}^2$. Here we are choosing $u_{1t}=\varepsilon_{1t}$ and are decomposing ε_{2t} by the least squares projection equation $\varepsilon_{2t}=\rho\varepsilon_{1t}+u_{2t}$ where the least squares orthogonality condition $Eu_{2t}\varepsilon_{1t}=0$ implies that $\rho=E\varepsilon_{1t}\varepsilon_{2t}/E\varepsilon_{1t}^2$. Here u_{2t} is the part of ε_{2t} that is orthogonal to ε_{1t}. By construction, u_{1t} and u_{2t} are orthogonal. Therefore, a new moving average representation in terms of mutually orthogonal disturbances at all lags is given by

$$\begin{aligned} x_t &= \alpha' + \sum_{j=0}^{\infty} C_j F u_{t-j} \\ &= \alpha' + \sum_{j=0}^{\infty} D_j u_{t-j} \end{aligned}$$

where $D_j=C_jF$. Of course, there is more than such one choice of u_t processes that does the job. For example, in the n=2 example, we could have selected $u_{1t}=\varepsilon_{2t}$ and then chosen u_{2t} as the part of ε_{1t} that is orthogonal to ε_{2t}. In the text, for the n=2 case, I have calculated moving average representations for both of the ways of choosing u_t discussed above. More generally any choice of $u_t=F^{-1}\varepsilon_t$ that makes $Eu_tu_t'=F^{-1}\sum F'^{-1}$ a diagonal matrix can be used to deliver a moving average representation in terms of a u process that is orthogonal contemporaneously as well as at all leads and lags.

In the n=2 case, the first mentioned way of defining u_t is equivalent with changing the form the vector autoregression (a) by adding current x_{1t} to the right side of the autoregression for x_{2t}, and then solving the vector difference equation for a moving average representation in terms of the vector of residuals from this pair of autoregressions. The second mentioned way of defining u_t amounts to changing the form of the vector autoregression (a) by adding current x_{2t} to the right side of the autoregression for x_{1t} (leaving current x_{1t} excluded from the autoregression for x_{2t}) and calculating the moving average in terms of the residuals from these equations.

The k-step ahead error in forecasting x_t linearly from its own past is given by

$$x_t - \hat{E}_{t-k}x_t = C_0\varepsilon_t + \dots + C_{k-1}\varepsilon_{t-k+1}$$

$$= D_0u_t + \dots + D_{k-1}u_{t-k+1}$$

where $\hat{E}_{t-k}x_t$ is the linear least squares forecast of x_t given x_{t-k}, x_{t-k-1}, ... From the extensive orthogonality conditions built in we have that the covariance matrix of k-step ahead prediction errors is

$$E(x_t-\hat{E}_{t-k}x_t)(x_t-\hat{E}_{t-k}x_t)' = D_0Eu_tu_t'D_0' + \dots + D_{k-1}Eu_tu_t'D_{k-1}'.$$

By calculating the diagonal terms in this formula, we achieve a decomposition of the variance k-step ahead prediction error into the parts attributable to variance in the n components of u_t. For every choice of u_t process, there is such a decomposition of variance.

Under certain regularity conditions, least squares estimates of the vector autoregression (a) are known to be statistically consistent (Anderson and Taylor [1976] and Ljung [1976]). For a more extensive discussion of vector stochastic processes and some macroeconomic applications, see Sargent [1977b].

Table 1

Vector Autoregressions for Seasonally Unadjusted Data
(1948I-1972IV)*

	Dependent var: n_{1t}		Dependent var: w_t	
	Coefficient	Standard Error	Coefficient	Standard Error
Constant	7.7038	2.7640	.17353	.1124
Trend	.0506	.0168	.00103	.0006
Fourth Quarter Dummy	.5780	.4250	.02149	.0172
First Quarter Dummy	-.9797	.3163	.03616	.0128
Second Quarter Dummy	1.9887	.3883	.00646	.0158
n_{1t-1}	1.5946	.1075	-.00343	.0043
n_{1t-2}	-.9403	.2006	.00402	.0081
n_{1t-3}	.4128	.2001	-.00315	.0081
n_{1t-4}	-.1604	.1049	.00163	.0042
w_{t-1}	-1.5407	2.5467	.97586	.1036
w_{t-2}	2.0531	3.5659	-.02126	.1451
w_{t-3}	-4.5508	3.5039	.09912	.1426
w_{t-4}	1.4698	2.5500	-.13212	.1037
$\bar{R}^2$	.9969		.99790	
d.w.	1.9835		2.04370	
s.e.	.3677		.01500	
Marginal Significance level on lagged n's	.0000		.00000	
Marginal Significance level on lagged w's	.0910		.86900	

*Observation period on left-side variables.

Table 2

Vector Moving Average Representation of Real Wage and Aggregate Employment (1948I-1972IV)*

Lag	(1)	(2)	(3)	(4)
0	.3697	0	0	.0150
1	.5897	-.00126	-.0231	.0146
2	.5946	-.00177	-.0287	.0140
3	.5464	-.00253	-.0840	.0149
4	.5025	-.00329	-.1553	.0139
5	.4444	-.00359	-.2103	.0130
6	.3741	-.00370	-.2580	.0123
7	.3080	-.00371	-.2983	.0115
8	.2520	-.00359	-.3262	.0108
9	.2048	-.00339	-.3426	.0102
10	.1661	-.00316	-.3493	.0096
11	.1357	-.00291	-.3480	.0091
12	.1125	-.00266	-.3406	.0086
13	.0950	-.00243	-.3285	.0081
14	.0820	-.00221	-.3135	.0076
15	.0724	-.00202	-.2966	.0071
16	.0652	-.00184	-.2789	.0067
17	.0597	-.00169	-.2611	.0063
18	.0554	-.00155	-.2436	.0059
19	.0519	-.00143	-.2269	.0055
20	.0488	-.00132	-.2110	.0051
21	.0460	-.00123	-.1962	.0048
22	.0434	-.00114	-.1824	.0045
23	.0410	-.00106	-.1696	.0042

Column (1): Response of employment to one standard deviation innovation in employment.

Column (2): Response of real wage to one standard deviation innovation in employment.

Column (3): Response of employment to one standard deviation innovation in real wage.

Column (4): Response of real wage to one standard deviation innovation in real wage.

Correlation of innovations in employment and real wage is .2442.

*Observation period for left-hand side variables. For method of construction of vector moving average, see appendix.

Table 3

Decompositions of Variance of Forecast Errors*

Employment

K	Variance of K-Step Ahead Forecast Error	Percent of variance in K-step ahead forecast error explained by orthogonalized innovation in: Employment	Real Wage
K = 1.	.1367	94.03	5.96
K = 2.	.4783	95.24	4.75
K = 3.	.8244	95.59	4.40
K = 4.	1.1076	96.50	3.49
K = 5.	1.3462	97.04	2.95
K = 6.	1.5423	96.74	3.25
K = 7.	1.7017	95.41	4.58
K = 8.	1.8407	93.06	6.93
K = 9.	1.9705	89.96	10.03
K = 10.	2.0956	86.47	13.52
K = 11.	2.2169	82.91	17.08
K = 12.	2.3334	79.51	20.48
K = 20.	2.9620	64.14	35.85
K = 35.	3.2381	59.18	40.81

Real Wage

Real Wages	Variance of K-step Ahead Forecast Error	Percent of variance in K-step ahead forecast error explained by orthogonalized innovation in: Employment	Real Wage
K = 1.	.00022	0	100.00
K = 2.	.00043	.34	99.65
K = 3.	.00062	.71	99.28
K = 4.	.00083	1.25	98.74
K = 5.	.00101	2.02	97.97
K = 6.	.00117	2.78	97.21
K = 7.	.00132	3.45	96.54
K = 8.	.00145	4.04	95.95
K = 9.	.00156	4.53	95.46
K = 10.	.00166	4.91	95.08
K = 11.	.00175	5.20	94.79
K = 12.	.00183	5.41	94.58
K = 20.	.00220	5.87	94.12
K = 35.	.00238	5.91	94.08

Standard error of orthogonalized innovation in real wage = .3586.
Standard error of orthogonalized innovation in employment = .01505.

---

*The orthogonalized innovation in employment here equals the innovation in employment, while the orthogonalized innovation in the real wage equals that part of the innovation in the real wage that is orthogonal to the innovation in employment.

Table 4

Decompositions of Variance of Forecast Errors*

Employment

K	Variance of K-Step Ahead Forecast Error	Percent of variance in K-step ahead forecast error explained by orthogonalized innovation in: Employment	Real Wage
K = 1.	.136	100.0	0
K = 2.	.478	99.8	.10
K = 3.	.824	99.8	.15
K = 4.	1.107	99.2	.71
K = 5.	1.346	97.7	2.27
K = 6.	1.542	95.3	4.68
K = 7.	1.701	92.0	7.92
K = 8.	1.840	88.1	11.87
K = 9.	1.970	83.8	16.16
K = 10.	2.095	79.5	20.47
K = 11.	2.216	75.4	24.52
K = 12.	2.333	71.8	28.18
K = 20.	2.962	56.6	43.38
K = 35.	3.238	51.7	48.21

Real Wage

Real Wages	Variance of K-step Ahead Forecast Error	Percent of variance in K-step ahead forecast error explained by orthogonalized innovation in: Employment	Real Wage
K = 1.	.00022	5.96	94.03
K = 2.	.00043	4.34	95.65
K = 3.	.00062	3.47	96.52
K = 4.	.00083	2.74	97.25
K = 5.	.00101	2.24	97.75
K = 6.	.00117	1.95	98.04
K = 7.	.00132	1.78	98.21
K = 8.	.00145	1.67	98.32
K = 9.	.00156	1.61	98.38
K = 10.	.00166	1.56	98.43
K = 11.	.00175	1.52	98.47
K = 12.	.00183	1.48	98.51
K = 20.	.00220	1.26	98.73
K = 35.	.00238	1.17	98.82

Standard error of orthogonalized innovation in real wage = .3697.
Standard error of orthogonalized innovation in employment = .0146.

---

*The orthogonalized innovation in the real wage here just equals the innovation in the real wage, while the orthogonalized innovation in employment equals that part of the employment innovation that is orthogonal to the innovation in the real wage.

Table 5

First Solution of Likelihood Equations
Seasonally Adjusted Data (n=4)
1948I-1972IV*

$f_1 = 19.80$	$\rho_1 = .9372$	$v_1 = .9542$
$d = 2377.90$	$\rho_2 = .7800$	$v_2 = .0052$
$e = 104.02$	$\mu_1 = .2002$	$v_3 = .0743$
$\delta_1 = .5886$		$v_4 = -.1867$
$\alpha_1 = -.0185$	$\beta_1 = -.0600$	$KS(n_1) = .0706$
$\alpha_2 = .0009$	$\beta_2 = -.0001$	$KS(n_2) = .0744$
$\alpha_3 = .0017$	$\beta_3 = -.0004$	$KS(w) = .0346$
$\alpha_4 = .0044$	$\beta_4 = .0026$	

	.9220E-01	.2000E+00	.1298E-02
V =		.7747E+00	.2077E-02
			.1949E-03

	.9225E-01	.2002E+00	.1301E-02
$B^{-1}VB^{-1'}$ =		.7749E+00	.2089E-02
			.1949E-03

$|\hat{V}_r| = .5497E\text{-}05$, $|\hat{V}_u| = .4998E\text{-}05$

$T\{\log|\hat{V}_r| - \log|\hat{V}_u|\} = 9.5271$

Marginal confidence level = .7830

$|\tilde{V}_u| = .34743E\text{-}05$

$T\{\log|\hat{V}_r| - \log|\tilde{V}_u|\} = 45.881$

Marginal confidence level = .9868

---

*Period of observation on the dependent variables.

Table 6

Seasonally Adjusted Data--Second Solution of Likelihood Equations (n=4) 1948I-1972IV*

f_1 = .5386	v_1 = .9548	$KS(n_1)^+$ = .0755
d = 2367.87	v_2 = .0039	$KS(n_2)$ = .0765
e = 122.737	v_3 = .0721	KS(w) = .0309
ρ_1 = .5962	v_4 = -.1823	
ρ_2 = .2080		
δ_1 = .9317	μ_1 = .7778	
α_1 = -.0135	β_1 = -.3428	
α_2 = .0141	β_2 = .0821	
α_3 = .0160	β_3 = .1124	
α_4 = .0190	β_4 = .1769	

$$V = \begin{pmatrix} .9176E-01 & .1984E+00 & .1296E-02 \\ & .7718E+00 & .2071E-02 \\ & & .1949E-03 \end{pmatrix}$$

$$B^{-1}VB^{-1'} = \begin{pmatrix} .9180E-01 & .1988E+00 & .1299E-02 \\ & .7732E+00 & .2138E-02 \\ & & .1949E-03 \end{pmatrix}$$

$|\hat{V}_r|$ = .55074E-05, $|\hat{V}_u|$ = .4998E-05

$T\{\log|\hat{V}_r| - \log|\hat{V}_u|\}$ = 9.7113

Marginal confidence level = .7945

---

*Period of observation on the dependent variables.

+KS(n_1), KS(n_2), KS(w) denote Kolmogorov-Smirnov statistics on cumulated periodograms of innovations of n_1, n_2, and w respectively.

Table 7

Seasonally Unadjusted Data (n=4)
1948I-1972IV*

f_1 = .4718	v_1 = .9159	$KS(n_1)$ = .0679
d = .3266.00	v_2 = .0189	$KS(n_2)$ = .0736
e = 78.60	v_3 = .0959	KS(w) = .0252
ρ_1 = .3972	v_4 = .1938	
ρ_2 = .1024		
δ_1 = .9487	μ_1 = .7426	
α_1 = -.0078	β_1 = -.5085	
α_2 = .0104	β_2 = .0865	
α_3 = .0117	β_3 = .1322	
α_4 = .0137	β_4 = .2361	

	.1404E+00	.2668E+00	.1147E-02
V =		.8174E+00	.9501E-03
			.1945E-03

	.1405E+00	.2674E+00	.1148E-02
$B^{-1}VB^{-1'}$ =		.8184E+00	.1049E-02
			.1945E-03

$|\hat{V}_r|$ = .7864E-05, $|\hat{V}_u|$ = .7360E-05 $|\tilde{V}_u|$ = .58289E-05

$T\{\log|\hat{V}_r|-\log|\hat{V}_u|\}$ = 6.6248 $T\{\log|\hat{V}_r|-\log|\tilde{V}_u|\}$ = 29.94658

Marginal confidence level = .5310 Marginal confidence level = .68344

*Period of observation on the dependent variables.

Table 8

Seasonally Unadjusted Data (n=8)
1949I-1972IV*

f_1 = .3612	v_1 = .8719	v_5 = .0322
d = .3266.29	v_2 = .0982	v_6 = .0795
e = 75.6750	v_3 = .1183	v_7 = -.1688
ρ_1 = .4094	v_4 = -.2537	v_8 = .0397
ρ_2 = .0571		
δ_1 = .9569	μ_1 = .7687	$KS(n_1)$ = .0756
α_1 = -.3790	β_1 = -.7970	$KS(n_2)$ = .0706
		KS(w) = .0308
α_2 = -.0745	β_2 = -.0417	
α_3 = -.0448	β_3 = .0211	
α_4 = -.0045	β_4 = .1232	
α_5 = -.1010	β_5 = -.0335	
α_6 = -.0989	β_6 = -.0203	
α_7 = -.0787	β_7 = .0356	
α_8 = -.1505	β_8 = -.0857	

	.1355E+00	.2721E+00	.9675E-03
V =		.8420E+00	.1147E-02
			.1791E-03

	.1362E+00	.2734E+00	.1035E-02
$B^{-1}VB^{-1'}$ =		.8439E+00	.1289E-02
			.1791E-03

$|\hat{V}_r|$ = .6802E-05, $|\hat{V}_u|$ = .6163-05

$T\{\log|\hat{V}_r|-\log|\hat{V}_u|\}$ = 9.4610

Marginal confidence level = .1478

$|\tilde{V}_u|$ = .339897E-05

$T\{\log|\hat{V}_r|-\log|\tilde{V}_u|\}$ = 66.599337

Marginal confidence level = .7680

---

*Period of observation on the dependent variables.

Table 9

Vector Autoregressions (n=8)
Seasonally Unadjusted Data
(1949I-1972IV)

(20a)	Unconstrained	Constrained by (21)
n_{1t-1}	1.4040	1.3663
n_{1t-2}	-.4305	-.3918
w_{t-1}	-.7105	-.2498
w_{t-2}	-2.8277	-.0515
w_{t-3}	-2.3616	-.0309
w_{t-4}	1.0005	-.0031
w_{t-5}	6.8486	-.0698
w_{t-6}	-.1849	-.0683
w_{t-7}	-6.2083	-.0543
w_{t-8}	.1310	.0466
(20b)		
n_{2t-1}	.8361	.8258
n_{2t-2}	-.0710	-.0439
w_{t-1}	-.5042	-.6911
w_{t-2}	-10.2706	-.0548
w_{t-3}	.2643	.0277
w_{t-4}	-6.8648	.1616
w_{t-5}	17.6580	-.0440
w_{t-6}	-7.3990	-.0266
w_{t-7}	6.7369	.0468
w_{t-8}	-12.0341	-.0268
(20c)		
w_{t-1}	.8557	.8719
w_{t-2}	.1021	.0982
w_{t-3}	.0699	.1183
w_{t-4}	-.2183	-.2536
w_{t-5}	.7217	-.0322
w_{t-6}	.9416	.0795
w_{t-7}	-.2382	-.1688
w_{t-8}	.0574	-.0397

Table 10

Moving Average Representation Implied By Model For Seasonally Unadjusted (Table 8) Estimates

Response to a One-Standard Deviation Innovation in n_1

Lag	n_1	w	n_2
0	.4324	0	0
1.	.5909	0	0
2.	.6379	0	0
3.	.6401	0	0
4.	.6247	0	0
5.	.6028	0	0
6.	.5788	0	0
7.	.5547	0	0
8.	.5312	0	0
9.	.5085	0	0
10.	.4366	0	0
11.	.4657	0	0
12.	.4456	0	0
13.	.4264	0	0
14.	.4081	0	0
15.	.3905	0	0
16.	.3737	0	0
17.	.3576	0	0
18.	.3422	0	0
19.	.3275	0	0
20.	.3134	0	0
21.	.2999	0	0
22.	.2869	0	0
23.	.2746	0	0
24.	.2628	0	0
25.	.2514	0	0
26.	.2406	0	0
27.	.2302	0	0
28.	.2203	0	0
29.	.2108	0	0
30.	.2018	0	0
31.	.1931	0	0

Table 10 (cont.)

Response to a One-Standard Deviation Innovation in w

Lag	n_1	w	n_2
0	0	.0144	0
1.	-.0036	.0125	-.0099
2.	-.0088	.0123	-.0177
3.	-.0148	.0137	-.0230
4.	-.0212	.0110	-.0257
5.	-.0281	.0096	-.0268
6.	-.0354	.0095	-.0269
7.	-.0432	.0060	-.0258
8.	-.0493	.0044	-.0236
9.	-.0540	.0030	-.0205
10.	-.0572	.0007	-.0169
11.	-.0588	-.0003	-.0132
12.	-.0591	-.0012	-.0095
13.	-.0582	-.0023	-.0060
14.	-.0561	-.0026	-.0030
15.	-.0533	-.0028	-.0004
16.	-.0499	-.0029	.0016
17.	-.0461	-.0026	.0031
18.	-.0422	-.0023	.0041
19.	-.0384	-.0020	.0046
20.	-.0348	-.0015	.0047
21.	-.0315	-.0010	.0045
22.	-.0285	-.0006	.0040
23.	-.0260	-.0002	.0034
24.	-.0239	.0000	.0027
25.	-.0222	.0002	.0019
26.	-.0208	.0004	.0012
27.	-.0198	.0005	.0006
28.	-.0190	.0006	.0000
29.	-.0184	.0006	-.0003
30.	-.0179	.0006	-.0007
31.	-.0176	.0005	-.0009

Table 10 (cont.)

Response to a One-Standard Deviation Innovation in n_2

Lag	n_1	w	n_2
0	0	0	1.080
1.	0	0	.892
2.	0	0	.689
3.	0	0	.530
4.	0	0	.407
5.	0	0	.313
6.	0	0	.240
7.	0	0	.185
8.	0	0	.142
9.	0	0	.109
10.	0	0	.084
11.	0	0	.064
12.	0	0	.049
13.	0	0	.038
14.	0	0	.029
15.	0	0	.022
16.	0	0	.017
17.	0	0	.013
18.	0	0	.010
19.	0	0	.007
20.	0	0	.006
21.	0	0	.004
22.	0	0	.003
23.	0	0	.002
24.	0	0	.002
25.	0	0	.001
26.	0	0	.001
27.	0	0	.000
28.	0	0	.000
29.	0	0	.000
30.	0	0	.000
31.	0	0	.000

Correlation Matrix of Innovations

	n_1	w	n_2
n_1	1.00	.197	.808
w		1.000	.135
n_2			1.000

Table 11

Variance Decompositions for Forecast Errors
Implied By Model (Table 8 and 9 Estimates)
Seasonally Unadjusted Data
Percentage of 35-Step Ahead Forecast Error
Variance in x Accounted for by "Orthogonalized
Innovations" in n_1, w, n_2:

	n_1	w	n_2
$x = n_1$	98.3	1.71	--
$x = w$	0	100.	0
$x = n_2$	63.79	1.38	34.84

Orthogonalization order*: w, n_1, n_2.

*See note to Table 13.

Table 12

Moving Average Representation Implied by Maximum Likelihood Estimates of Vector Autoregression, Unconstrained Seasonally Unadjusted Data (1948I-1972IV)

Response to a One-Standard Deviation Innovation in n_1

Lag	n_1	w	n_2
0	.4005	0	0
1.	.5623	0	0
2.	.6170	0	0
3.	.6241	0	0
4.	.6106	0	0
5.	.5886	0	0
6.	.5635	0	0
7.	.5377	0	0
8.	.5123	0	0
9.	.4877	0	0
10.	.4642	0	0
11.	.4417	0	0
12.	.4203	0	0
13.	.3999	0	0
14.	.3805	0	0
15.	.3621	0	0
16.	.3445	0	0
17.	.3278	0	0
18.	.3119	0	0
19.	.2967	0	0
20.	.2823	0	0
21.	.2686	0	0
22.	.2556	0	0
23.	.2432	0	0
24.	.2314	0	0
25.	.2202	0	0
26.	.2095	0	0
27.	.1993	0	0
28.	.1896	0	0
29.	.1804	0	0
30.	.1717	0	0
31.	.1633	0	0

Table 12 (cont.)

Response to a One Standard Deviation
Innovation in w

Lag	n_1	w	n_2
0	0	.0144	0
1.	-.0102	.0123	-.0072
2.	-.0640	.0120	-.1607
3.	-.1631	.0125	-.2287
4.	-.2592	.0097	-.3764
5.	-.2533	.0087	-.2306
6.	-.2135	.0090	-.2083
7.	-.2416	.0051	-.0945
8.	-.2613	.0037	-.1603
9.	-.2869	.0021	-.2076
10.	-.3095	-.0001	-.2269
11.	-.3044	-.0008	-.2143
12.	-.3028	-.0017	-.2149
13.	-.3077	-.0029	-.1868
14.	-.2987	-.0029	-.1863
15.	-.2910	-.0031	-.1644
16.	-.2786	-.0031	-.1331
17.	-.2589	-.0026	-.1044
18.	-.2428	-.0023	-.0753
19.	-.2256	-.0019	-.0451
20.	-.2060	-.0013	-.0249
21.	-.1897	-.0008	-.0022
22.	-.1730	-.0004	.0154
23.	-.1573	.0000	.0267
24.	-.1449	.0002	.0357
25.	-.1333	.0005	.0400
26.	-.1234	.0007	.0397
27.	-.1158	.0007	.0386
28.	-.1090	.0008	.0343
29.	-.1036	.0007	.0284
30.	-.0996	.0006	.0224
31.	-.0960	.0005	.0156

Table 12 (cont.)

Response to a One Standard Deviation
Innovation in n_2

Lag	n_1	w	n_2
0	0	0	1.0595
1.	0	0	.8859
2.	0	0	.6654
3.	0	0	.4934
4.	0	0	.3653
5.	0	0	.2704
6.	0	0	.2001
7.	0	0	.1481
8.	0	0	.1096
9.	0	0	.0811
10.	0	0	.0600
11.	0	0	.0444
12.	0	0	.0329
13.	0	0	.0243
14.	0	0	.0180
15.	0	0	.0133
16.	0	0	.0098
17.	0	0	.0073
18.	0	0	.0054
19.	0	0	.0040
20.	0	0	.0029
21.	0	0	.0021
22.	0	0	.0016
23.	0	0	.0012
24.	0	0	.0008
25.	0	0	.0006
26.	0	0	.0004
27.	0	0	.0003
28.	0	0	.0002
29.	0	0	.0001
30.	0	0	.0001
31.	0	0	.0001

Correlation Matrix of Innovations

	n_1	w	n_2
n_1	1.00	.2066	.7971
w		1.0000	.1350
n_2			1.0000

Table 13

Decomposition of Variance of Forecast Error for Unconstrained Estimates, Seasonally Unadjusted Data (1948I-1972IV)

Percentage of 35-Step Ahead Forecast Error Variance in x Accounted for by "Orthogonalized Innovations" in n_1, w, n_2

Orthogonalization order*: n_1, w, n_2

	n_1	w	n_2
$x = n_1$	74.71	25.29	0
$x = w$	4.27	95.73	0
$x = n_2$	49.14	20.12	30.73

Orthogonalization order: n_1, n_2, w

	n_1	w	n_2
$x = n_1$	74.71	25.23	.06
$x = w$	4.27	95.49	.24
$x = n_2$	49.14	19.03	31.83

Orthogonalization order: w, n_1, n_2

	n_1	w	n_2
$x = n_1$	87.76	12.24	0
$x = w$	0	100	0
$x = n_2$	52.22	17.05	30.73

---

*Orthogonalization order refers to the procedure described in the appendix of defining an orthogonal u process from $u_t = F\varepsilon_t$. If the orthogonalization order is n_1, w, n_2, then the "orthogonalized n_1 innovation" is simply the n_1 innovation; the "orthogonalized w innovation" is the part of the w innovation orthogonal to the n_1 innovation; the "orthogonalized n_2 innovation" is the part of the n_2 innovation that is orthogonal to both the n_1 and w innovations.

Table 14+

Decomposition of Variance of Forecast Errors, Seasonally Unadjusted Data (1948I-1972IV)

Percentage of 35-Step Ahead Forecast Error Variance in x Accounted for by "Orthogonalized Innovations" in n_1, w, n_2

Orthogonalization order*: n_1, w, n_2

	n_1	w	n_2
$x = n_1$	50.74	49.26	0
$x = w$	3.58	96.42	0
$x = n_2$	45.68	24.71	29.62

Orthogonalization order: n_1, n_2, w

	n_1	w	n_2
$x = n_1$	50.74	49.23	.03
$x = w$	3.58	96.36	.06
$x = n_2$	45.68	24.31	30.01

Orthogonalization order: w, n_1, n_2

	n_1	w	n_2
$x = n_1$	64.66	35.34	0
$x = w$	0	100	0
$x = n_2$	46.80	23.59	29.62

---

*See note to Table 13.

+Data are residuals from regressions on constant, trend, and three seasonal dummies, with no trend squared terms, in contradistinction to the Table 13 results.

Table 15

Summary Statistics for Fourth-Order Vector Autoregressions for (n_1, n_2, w)

Seasonally Unadjusted data+
(1948I-1972IV)

Marginal Significance Levels*
Pertinent for Testing Null Hypothesis
That Lagged n_1 or n_2 or w's Have Zero
Coefficients in Autoregression for x

	n_1	n_2	w
$x = n_1$	.0000	.0000	.2000
$x = n_2$	.0395	.0000	.0446
$x = w$	.6857	.6128	.0000

---

*Where f is the calculated value of the pertinent F-statistic, the marginal significance level is defined as prob{F>f} under the null hypothesis.

+Regressions included a constant, trend, and three seasonal dummies.

Table 16

Decomposition of Variance of Forecast Error Implied by Vector Autoregression for (n_1, n_2, w) Seasonally Unadjusted Data

(1948I-1972IV)

Percentage of 35-Quarter Ahead Forecast Error Variance in x Accounted for by "Orthogonalized Innovation" in n_1, w, n_2

Orthogonalization order*: n_1, w, n_2

	n_1	w	n_2
$x = n_1$	21.82	48.74	29.44
$x = w$	.76	98.39	.85
$x = n_2$	23.64	16.25	60.25

Orthogonalization order: w, n_1, n_2

	n_1	w	n_2
$x = n_1$	26.90	43.66	29.44
$x = w$	2.11	97.03	.85
$x = n_2$	20.82	18.93	60.24

---

*See note to Table 13.