METHODS OF ESTIMATION FOR MARKETS IN DISEQUILIBRIUM: A FURTHER STUDY

BY RAY C. FAIR AND HARRY H. KELEJIAN

This paper is concerned with the problem of estimating demand and supply schedules in disequilibrium markets. The results of Fair and Jaffee are expanded in three ways. (1) Their directional method 1 is modified to yield consistent estimates. (2) A maximum likelihood alternative to their quantitative method is proposed. (3) The price equation is generalized to be a multivariate, stochastic function, and a method is proposed for estimating demand and supply schedules in this case.

1. INTRODUCTION

IN A RECENT paper Fair and Jaffee [4] considered the problem of estimating demand and supply schedules in disequilibrium markets. They suggested four possible methods of estimation: a general maximum likelihood method for finding the optimal separation of the sample period into demand and supply regimes; two “directional” methods, which relied on price-setting information to separate the sample period; and a “quantitative” method, which relied on price-setting information to adjust the observed quantity for the effects of rationing. The Fair-Jaffee study is subject to several limitations. First, Fair and Jaffee found that the general maximum likelihood method was not computationally feasible. Second, their directional method 1, although yielding a correct sample separation under the assumptions of the model, does not yield consistent estimates. Finally, their quantitative method is based on a rather strict assumption about price-setting behavior, namely that price changes are strictly proportional to excess demand.

The purpose of this paper is to expand upon the results of Fair and Jaffee in three ways. First, their directional method 1 will be modified to yield consistent estimates, and then this modified technique will be used to estimate a particular model so that these estimates can be compared to the directional method 1 estimates. Second, a maximum likelihood alternative to the quantitative method will be proposed under the same strict assumption that price changes are proportional to excess demand. Third, and most important, the strict assumption about price-setting behavior will be relaxed and a method will be proposed for estimating supply and demand schedules under the much weaker assumption that the price equation is a multivariate and stochastic relationship.

1 The authors would like to thank Dwight M. Jaffee for helpful comments. He is not, of course, responsible for any shortcomings of this paper.

177
2. DIRECTIONAL METHODS OF ESTIMATION

The Model

The model under consideration in this section is

\begin{align*}
(1a) \quad D_t &= X_{1t}\beta_1 + P_{t-1}\beta_2 + u_{1t}, \\
(1b) \quad S_t &= X_{2t}\beta_3 + P_{t-1}\beta_4 + u_{2t}, \\
(1c) \quad \Delta P_t &= P_t - P_{t-1} = f(D_t - S_t),
\end{align*}

and

\begin{align*}
(1d) \quad Q_t &= \min\{D_t, S_t\} \quad (t = 1, 2, \ldots, T),
\end{align*}

where, at time \( t \), \( D_t \) and \( S_t \) are the quantities demanded and supplied respectively, \( Q_t \) is the actual quantity observed, \( P_t \) is the price, \( X_{1t} \) and \( X_{2t} \) are vectors of predetermined variables, and \( u_{1t} \) and \( u_{2t} \) are the disturbance terms. The vectors \( \beta_1 \) and \( \beta_3 \) are vectors of parameters, conformably defined. The stochastic assumptions are

\begin{align*}
(2) \quad E[u_{1t} | X_t] &= E[u_{2t} | X_t] = 0, \\
E[u_{1t}^2 | X_t] &= \sigma_1^2, \quad E[u_{2t}^2 | X_t] = \sigma_2^2,
\end{align*}

where \( X_t = (X_{1t}, X_{2t}, P_{t-1}); u_{1t} \) and \( u_{2t} \) are assumed to be continuous.

The problem of estimation concerning the parameters of (1a) and (1b) is that the price equation, (1c), implies that prices do not adjust in every period in such a manner as to equate \( D_t \) and \( S_t \). Therefore, unless some adjustments are made, all of the observations on \( Q_t \) cannot be used in the estimation of equations (1a) and (1b).

Directional Method I

Fair and Jaffee's directional method I is based on the assumption that \( f(D_t - S_t) \geq 0 \) when \( D_t - S_t \geq 0 \). Under this assumption, if \( \Delta P_t > 0 \), then \( Q_t = S_t \); if \( \Delta P_t < 0 \), then \( Q_t = D_t \); and if \( \Delta P_t = 0 \), then \( Q_t = D_t = S_t \). Directional method I takes those sample points for which \( \Delta P_t \geq 0 \) and estimates the supply equation, and takes those sample points for which \( \Delta P_t \leq 0 \) and estimates the demand equation.

Although directional method I yields a correct sample separation under the above assumptions, the coefficient estimates are not consistent. For instance, according to the method

\begin{align*}
(3) \quad Q_t &= D_t = X_{1t}\beta_1 + P_{t-1}\beta_2 + u_{1t}, \quad \text{when } \Delta P_t \leq 0, \\
\end{align*}

and

\begin{align*}
(4) \quad Q_t &= S_t = X_{2t}\beta_3 + P_{t-1}\beta_4 + u_{2t}, \quad \text{when } \Delta P_t \geq 0.
\end{align*}

2 In this section the price terms are assumed to enter the demand and supply equations with a lag rather than contemporaneously. In Sections 3 and 4 the price terms are allowed to enter the demand and supply equations contemporaneously.
Now, the ordinary least squares parameter estimates of (3) and (4) are inconsistent because the means of the disturbance terms are no longer independent of $X_{1t}$, $X_{2t}$, and $P_{t-1}$ over the relevant sample points. To see this, consider $E[u_{1t}|X_t, \Delta P_t \leq 0]$. In light of (1c), (1d), and the assumption that $f(D_t - S_t) \geq 0$ when $D_t - S_t \geq 0$:

\begin{equation}
E[u_{1t}|X_t, \Delta P_t \leq 0] = E[u_{1t}|X_t, D_t \leq S_t].
\end{equation}

Let $D_t^* = X_{1t}\beta_1 + P_{t-1}\beta_2$, and $S_t^* = X_{2t}\beta_3 + P_{t-1}\beta_4$. Then from (1a), (5) may be written as

\begin{equation}
E[u_{1t}|u_{1t}, \Delta P_t < 0, X_t] = \frac{\int_{-\infty}^{\infty} \int_{-\infty}^{S_t - D_t^*} u_{1t} \frac{\partial}{\partial \phi} \frac{g_1(u_{1t}, \phi|X_t)}{g_1(\phi|X_t)} \, d\phi \, du_{1t}}{\int_{-\infty}^{\infty} \int_{-\infty}^{S_t - D_t^*} \frac{\partial}{\partial \phi} \frac{g_1(u_{1t}, \phi|X_t)}{g_1(\phi|X_t)} \, d\phi}.
\end{equation}

Clearly, the expectation in (7) will not, in general, be independent of $X_t$ unless $u_{1t}$ and $\phi_t$ are independent, in which case $g_2$ factors.

**A Consistent Method of Estimation**

The demand and supply equations, (3) and (4), can be consistently estimated by a maximum likelihood technique that is conditional on the segmentation of the sample. Let $g_3(u_{1t}|\Delta P_t \leq 0, X_t)$ be the conditional density of $u_{1t}$ given $X_t$ and $\Delta P_t \leq 0$, and let $g_4(u_{2t}|\Delta P_t > 0, X_t)$ be the conditional density of $u_{2t}$ given $X_t$ and $\Delta P_t > 0$. In view of the above assumptions, these conditional densities can be written as $g_3(u_{1t}|\phi_t = S_t^* - D_t^*, X_t)$ and $g_4(u_{2t}|\phi_t > S_t^* - D_t^*, X_t)$ respectively. Now, the maximum likelihood estimators of the parameters $\beta_1, \beta_2, \beta_3, \beta_4, \sigma_1^2, \sigma_2^2$, and $\sigma_{12}$, can be obtained by maximizing the likelihood function

\begin{equation}
\mathcal{L} = \prod_{\Delta P_t < 0} g_3(u_{1t}|\phi_t \leq S_t^* - D_t^*, X_t) \quad \prod_{\Delta P_t > 0} g_4(u_{2t}|\phi_t > S_t^* - D_t^*, X_t),
\end{equation}

where $u_{1t} = Q_t - D_t^*, u_{2t} = Q_t - S_t^*$, and the products are taken, respectively, over the periods for which $\Delta P_t < 0$ and $\Delta P_t > 0$. Equation (8) is defined in terms of strict inequalities because the disturbance terms are assumed to be continuous and thus the probability that $\Delta P_t = 0$ is zero.

**Empirical Results**

The likelihood technique can be implemented as follows: First, note that

\begin{equation}
g_3(u_{1t}|\phi_t \leq S_t^* - D_t^*, X_t) = \frac{\int_{-\infty}^{S_t^* - D_t^*} \frac{\partial}{\partial \phi} \frac{g_2(u_{1t}, \phi_t|X_t)}{g_1(\phi_t|X_t)} \, d\phi_t}{\int_{-\infty}^{\infty} \frac{\partial}{\partial \phi} \frac{g_2(u_{1t}, \phi_t|X_t)}{g_1(\phi_t|X_t)} \, d\phi_t}.
\end{equation}

\(^3\) See Mood and Graybill [12, Chs. 1–5] for a discussion of the concepts in the development up through (7).
where, as above, $g_2(u_{1t}, \phi_t|X_t)$ is the joint density of $u_{1t}$ and $\phi_t$, given $X_t$. Likewise,

$$g_4(u_{2t}, \phi_t|X_t) = \frac{\int_{S_{t}^{*} - D_{t}^{*}} g_3(u_{2t}, \phi_t|X_t) d\phi_t}{\int_{S_{t}^{*} - D_{t}^{*}} g_1(\phi_t|X_t) d\phi_t},$$

where $g_2(u_{1t}, \phi_t|X_t)$ is the joint density of $u_{2t}$ and $\phi_t$, given $X_t$. If the joint density of $u_{1t}$ and $u_{2t}$, conditional on $X_t$, is specified to be normal, then $g_1(\phi_t|X_t)$ will be normal and $g_2(u_{1t}, \phi_t|X_t)$ and $g_3(u_{2t}, \phi_t|X_t)$ will each be jointly normal. Therefore, since it is quite easy numerically to evaluate normal integrals, an attempt can be made to maximize the likelihood function in (8) using a nonlinear maximization program. Note that the parameters $\beta_1$, $\beta_2$, $\beta_3$, and $\beta_4$ enter both the limits of the integral and the integrand. The parameters $\sigma_1^2$, $\sigma_2^2$, and $\sigma_{12}$ enter only into the integrand.

In order to see whether it is feasible to maximize (8), the housing starts model that Fair and Jaffee used as an example in their study was also used as an example in this study. The model consists of one demand equation and one supply equation:

$$HS_t^D = \alpha_0 + \alpha_1 t + \alpha_2 \sum_{i=1}^{t-1} HS_i + \alpha_3 RM_{t-2} + u_{1t},$$

and

$$HS_t^S = \psi_0 + \psi_1 t + \psi_2 DF6_{t-1} + \psi_3 DHF3_{t-2} + \psi_4 RM_{t-1} + u_{2t},$$

where $HS_t^D$ and $HS_t^S$ denote the demand for and supply of housing starts respectively, $RM_{t-1}$ and $RM_{t-2}$ denote the mortgage rate lagged one and two months respectively, $DF6_{t-1}$ denotes the six-month moving average of the flow of deposits into savings and loan associations (SLAs) and mutual savings banks lagged one month, and $DHF3_{t-2}$ denotes the three-month moving average of the flow of borrowings by SLAs from the Federal Home Loan Bank lagged two months.a Fair and Jaffee assumed that the error terms, $u_{1t}$ and $u_{2t}$, were first-order serially correlated, but for present purposes serial correlation problems will be ignored. Serial correlation questions will be considered at the end of this section. Fair and Jaffee also used seasonally unadjusted data and seasonal dummy variables, but for purposes here seasonally adjusted data were used.

The results of estimating equations (11) and (12) by directional method 1 and by the consistent likelihood technique are presented in Table I. The price variable in the model is the mortgage rate. Two nonlinear maximization techniques were tried in the maximization of the likelihood function: the quadratic hill-climbing technique of Goldfed, Quandt, and Trotter [9] and the technique of Powell [13]. The quadratic hill-climbing technique requires first and second derivatives, and for this purpose numerical first and second derivatives were used. The normal integrals were evaluated using the ERF function in the FORTRAN library. It turned out that the likelihood function was not very well behaved. The function

\[ ^a \text{See Fair [3, Ch. 8] for a more complete description of this model.} \]
was very flat with respect to the parameter $\sigma_{12}$, for example, and the function appeared to have many local maxima. The quadratic hill-climbing technique and Powell's technique worked about equally well in their ability to find local maxima. The maximum likelihood estimates presented in Table I correspond to the largest value of the likelihood function found after considerable experimentation, but there is no guarantee that this is the global maximum.

**TABLE I**

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Directional Method I</th>
<th>Maximum Likelihood Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha_0$</td>
<td>223.7</td>
<td>223.4</td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>2.428</td>
<td>2.429</td>
</tr>
<tr>
<td>$\alpha_2$</td>
<td>-.0188</td>
<td>-.0119</td>
</tr>
<tr>
<td>$\alpha_3$</td>
<td>-.2032</td>
<td>-.2013</td>
</tr>
<tr>
<td>$\psi_0$</td>
<td>15.53</td>
<td>15.49</td>
</tr>
<tr>
<td>$\psi_1$</td>
<td>-.195</td>
<td>-.209</td>
</tr>
<tr>
<td>$\psi_2$</td>
<td>.0515</td>
<td>.0521</td>
</tr>
<tr>
<td>$\psi_3$</td>
<td>.0469</td>
<td>.0519</td>
</tr>
<tr>
<td>$\psi_4$</td>
<td>.1017</td>
<td>.1019</td>
</tr>
<tr>
<td>$\sigma^2_1$</td>
<td>151.04</td>
<td>222.90</td>
</tr>
<tr>
<td>$\sigma^2_2$</td>
<td>76.39</td>
<td>69.06</td>
</tr>
<tr>
<td>$\sigma_{12}$</td>
<td>—</td>
<td>59.92</td>
</tr>
</tbody>
</table>

The maximum likelihood estimates in Table I are quite close to the directional method I estimates, which suggests that for this particular example the bias using directional method I does not appear to be very great. Whether this is true in general is, of course, not clear.

**Serial Correlation Questions**

If the error terms in equations (1a) and (1b) are serially correlated, then it turns out that the coefficients of (1a) and (1b) are not identified if it is assumed that $u_{1t}$ and $u_{2t}$ are continuous random variables. Assume, for example, that the error terms are first-order serially correlated:

$$u_{1t} = u_{1t-1} \rho_1 + \varepsilon_{1t},$$  

(13)

and

$$u_{2t} = u_{2t-1} \rho_2 + \varepsilon_{2t},$$  

(14)

where the assumptions in (2) now pertain to $\varepsilon_{1t}$ and $\varepsilon_{2t}$ rather than to $u_{1t}$ and $u_{2t}$. Using (13) and (14), equations (1a) and (1b) can be written:

$$D_t = D_{t-1} \rho_1 + X_{1t} \beta_1 - X_{1t-1} \beta_1 + P_{t-1} \beta_2 - P_{t-2} \beta_2 \rho_1 + \varepsilon_{1t},$$  

(15)
The problem with estimating equations (15) and (16) is that the explanatory variables \( D_{t-1} \) and \( S_{t-1} \) will generally not be observed at the same time. For instance, given the above assumptions, only if \( \Delta P_{t-1} = 0 \) are both \( S_{t-1} \) and \( D_{t-1} \) observed; otherwise only one of them is observed. However, if \( u_{11} \) and \( u_{21} \) are continuous random variables, the probability that \( \Delta P_{t} = 0 \), for any \( t \), is zero. Consequently, as the sample size approaches infinity, that portion of it corresponding to time periods for which \( \Delta P_{t} = 0 \) will remain finite. Now, if it is recalled, from either (7) or (9) and (10), that a consistent estimation technique, for either or both equations, necessarily involves observations on all of the predetermined variables, the result concerning lack of identification follows. The same situation also holds if \( D_{t-1} \) and \( S_{t-1} \) enter directly as explanatory variables in equations (1a) and (1b) rather than entering indirectly by way of the serial correlation assumption.

3. QUANTITATIVE METHODS OF ESTIMATION

The Model

The model under consideration in this section is

\[
(17a) \quad D_t = X_{1t} \beta_1 + P_t \beta_2 + u_{1t},
\]

\[
(17b) \quad S_t = X_{2t} \beta_3 + P_t \beta_4 + u_{2t},
\]

\[
(17c) \quad \Delta P_t = \gamma (D_t - S_t),
\]

and

\[
(17d) \quad Q_t = \min \{ D_t, S_t \} \quad (t = 1, 2, \ldots, T).
\]

The model in this section differs from the model in Section 2 in that the price term is allowed to enter contemporaneously in the demand and supply equations and the change in price is assumed to be directly proportional to the level of excess demand. The model is assumed to be identified. Fair and Jaffee demonstrated that the above model can be estimated by relating \( Q_t \) to \( D_t \) and \( S_t \) by means of equations (17c) and (17d).\(^6\)

\(^5\) For directional method I this problem of identification does not arise since one ignores the problems that arise because of sample segmentation and one chooses as sample points for, say, the demand equation only those points for which both \( D_t \) and \( D_{t-1} \) are observed. This sample segmentation requires throwing away one observation for every switching point. For their empirical work using directional method I, Fair and Jaffee did not actually throw away the requisite number of observations, but assumed that at a switching point both \( D_t \) and \( D_{t-1} \) or \( S_t \) and \( S_{t-1} \) were observed.

\(^6\) For example, if \( \Delta P_t \geq 0 \), then \( S_t = Q_t \) and \( D_t = Q_t + (1/\gamma) \Delta P_t. \)
The model in (17a)-(17d) can also be estimated by the maximum likelihood technique in a manner similar to that done for the model in Section 2. First, the sample can be partitioned as follows:

\[(18)\]  
\[Q_t = X_{it}\beta_1 + P_t\beta_2 + u_{2t},\]  
\[Q_t = X_{it}\beta_1 + P_t\left(\beta_2 - \frac{1}{\gamma}\right) - P_{t-1}\frac{1}{\gamma} + u_{1t},\]  
when \(\Delta P_t \geq 0,\)

and

\[(19)\]  
\[Q_t = X_{it}\beta_1 + P_t\beta_2 + u_{1t},\]  
\[Q_t = X_{it}\beta_1 + P_t\left(\beta_4 - \frac{1}{\gamma}\right) - P_{t-1}\frac{1}{\gamma} + u_{2t},\]  
when \(\Delta P_t \leq 0.\)

Consider the equations in (18) corresponding to the sample segmentation, \(\Delta P_t \geq 0.\) These equations can be considered as a two equation system in the variables \(Q_t\) and \(P_t.\) Therefore, the likelihood function for the equations of (18), given the sample segmentation, is based on the joint conditional density of \(u_{1t}\) and \(u_{2t}\), given \(X_t\) and \(\Delta P_t \geq 0,\) where, as in Section 2, \(X_t = (X_{1t}, X_{2t}, P_{t-1}).\) Let this density be \(g_6(u_{1t}, u_{2t} \mid \Delta P_t \geq 0, X_t).\) Then, in a manner not dissimilar from that of Section 2,

\[(20)\]  
\[g_6(u_{1t}, u_{2t} \mid \Delta P_t \geq 0, X_t) = g_6(u_{1t}, u_{2t} \mid D_t \geq S_1, X_t)\]
\[= g_6(u_{1t}, u_{2t} \mid X_{1t}\beta_1 + P_t\beta_2 + u_{1t} \geq X_{2t}\beta_3 + P_t\beta_4 + u_{2t}, X_t)\]
\[= g_6(u_{1t}, u_{2t} \mid x_1u_{1t} + x_2u_{2t} \geq G(X_t), X_t),\]

where the last step of (20) is obtained by replacing \(P_t\) by its reduced form expression in \(u_{1t}, u_{2t},\) and the predetermined variables, putting all terms not involving \(u_{1t}\) or \(u_{2t}\) on the right hand side, and denoting the resulting expression as \(G(X_t).\) The parameters \(x_1\) and \(x_2\) are functions of the parameters in equations (17a)-(17c).

The likelihood function for the equations of (19), given the sample separation, is based on the joint conditional density of \(u_{1t}\) and \(u_{2t}\), given \(X_t\) and \(\Delta P_t \geq 0.\) Let this density be \(g_7(u_{1t}, u_{2t} \mid \Delta P_t \geq 0, X_t).\) The derivation of \(g_7(u_{1t}, u_{2t} \mid \Delta P_t \leq 0, X_t)\) is almost identical to that for \(g_6(u_{1t}, u_{2t} \mid \Delta P_t \geq 0, X_t).\) Now, the maximum likelihood estimators of the parameters of equations (17a)-(17c) are obtained by maximizing the likelihood function

\[(21)\]  
\[\mathscr{L} = \prod_{\Delta P_t > 0} g_6(u_{1t}, u_{2t} \mid \Delta P_t \geq 0, X_t)J_1 \times \prod_{\Delta P_t < 0} g_7(u_{1t}, u_{2t} \mid \Delta P_t \leq 0, X_t)J_2,\]

where \(u_{1t}\) and \(u_{2t}\) are replaced in both products of (21) by their corresponding expressions in (18) and (19), and \(J_1\) and \(J_2\) are the corresponding Jacobians of transformation from \(u_{1t}\) and \(u_{2t}\) to \(Q_t\) and \(P_t.\)
The likelihood technique can be implemented in a manner similar to that described in Section 2, although the situation is somewhat more complicated in the present case. The joint conditional density $g_5$, for example, can be obtained as

$$
g_5(u_{1t}, u_{2t}|z_1 u_{1t} + x_2 u_{2t} \geq G(X_t), X_t) = \frac{g_2(u_{1t}, u_{2t}|X_t)}{\int_{x_1 u_{1t} + x_2 u_{2t} \geq G(X_t)} g_2(u_{1t}, u_{2t}|X_t) \, du_{1t} \, du_{2t}},$$

where $g_2(u_{1t}, u_{2t}|X_t)$ is the joint density of $u_{1t}$ and $u_{2t}$ conditional on $X_t$ and where the double integral in the denominator represents the probability that $z_1 u_{1t} + x_2 u_{2t} \geq G(X_t)$. A similar expression can be derived for the joint conditional density $g_6$. If the joint density of $u_{1t}$ and $u_{2t}$, conditional on $X_t$, is specified to be normal, then it is possible to evaluate numerically the double integral in (22). It is thus possible to attempt to maximize the likelihood function in (21) using a nonlinear maximization program.

4. A METHOD OF ESTIMATION FOR THE GENERALIZED MODEL

The Generalized Model

The models described in Sections 2 and 3 contain price equations which are nonstochastic functions of only one variable, namely excess demand. In this section the price equation is generalized to be a multivariate, stochastic function. The model is taken to consist of equations (17a), (17b), (17d), and

$$(17c') \Delta P_t = \beta_5 (D_t - S_t) + X_{3t} \beta_6 + u_{3t},$$

where $X_{3t}$ is a vector of predetermined variables and $\beta_6$ is a vector of parameters.

A Method of Estimation

Because the price equation is multivariate and stochastic, the observed quantity, $Q_t$, cannot be strictly identified with either $D_t$ or $S_t$ on the basis of observed price changes. Hence, $Q_t$ must be related to $D_t$ and $S_t$ probabilistically on the basis of observed price changes. Define a selector variable $r_t$, where

$$r_t = 1 \quad \text{if } D_t > S_t, \quad r_t = 0 \quad \text{if } D_t < S_t.$$

Using this variable, equations (17a), (17b), and (17d) can be written as

$$Q_t = r_t S_t + (1 - r_t) D_t.$$

Now, from $(17c')$ $r_t = 1$ if $\Delta P_t > X_{3t} \beta_6 + u_{3t}$, and $r_t = 0$ if $\Delta P_t < X_{3t} \beta_6 + u_{3t}$. In light of these relations, the conditional density of $r_t$ given $\Delta P_t$ and $X_t$, where $X_t$

---

7 Since the disturbance terms of the model are assumed to be continuous variables, the probability is zero that $D_t$ will equal $S_t$. Thus, the problem of defining $r_t$ when $D_t = S_t$ can be ignored.
now includes $X_3$, as well as $X_{1t}, X_{2t}$, and $P_{t-1}$, can be expressed as

$$ (25) \quad \text{prob} \left( r_t = 1 | \Delta P_t, X_t \right) = H(\Delta P_t, X_t), $$

$$ \text{prob} \left( r_t = 0 | \Delta P_t, X_t \right) = 1 - H(\Delta P_t, X_t), $$

where $H(\Delta P_t, X_t) = \text{prob} \left( u_{3t} \leq \Delta P_t - X_3 \beta_3 | \Delta P_t, X_t \right)$.

The probability statement, $\text{prob} \left( u_{3t} \leq \Delta P_t - X_3 \beta_3 | \Delta P_t, X_t \right)$, can be obtained from the conditional density of $u_{3t}$ given $\Delta P_t$ and $X_t$. This density in turn depends upon the joint density of $\Delta P_t$ and $u_{3t}$ and the marginal density of $\Delta P_t$, both conditional on $X_t$. Finally, these densities can be derived if the joint density of the disturbance terms is specified. Thus, if the joint density of the disturbance terms is specified, the functional form of $H(\Delta P_t, X_t)$ is determined.

Assuming that the joint density of the disturbance terms is specified, the estimation of the model can now be considered. The model under consideration is

$$ (26a) \quad Q_t = r_t S_t + (1 - r_t) D_t, $$

$$ (26b) \quad D_t = X_{1t} \beta_1 + P_t \beta_2 + u_{1t}, $$

$$ (26c) \quad S_t = X_{2t} \beta_3 + P_t \beta_4 + u_{2t}, $$

$$ (26d) \quad \Delta P_t = \beta_5 (D_t - S_t) + X_3 \beta_6 + u_{3t}, $$

and

$$ (26e) \quad r_t = H(\Delta P_t, X_t) + u_{4t}, $$

where, in light of (25), $u_{4t}$ is a random variable such that $E[u_{4t} | \Delta P_t, X_t] = 0$. Since observations on $D_t, S_t,$ and $r_t$ are not available, these variables will first be eliminated from the model. For the sake of having a compact notation, let $D^* = X_{1t} \beta_1 + P_t \beta_2, S^* = X_{2t} \beta_3 + P_t \beta_4, \text{ and } H_t = H(\Delta P_t, X_t).$ Now, $D_t, S_t,$ and $r_t$ can be eliminated from (26a)-(26e) to get

$$ (27a) \quad Q_t = H_t S_t^* + D_t^* - H_t D_t^* + \Omega_t + e_{1t}, $$

and

$$ (27b) \quad \Delta P_t = \beta_5 (D_t^* - S_t^*) + X_3 \beta_6 + e_{2t}, $$

where $\Omega_t = H_t (u_{1t} - u_{1t}), e_{1t} = u_{1t} + u_{4t} (S_t - D_t), \text{ and } e_{2t} = u_{3t} + (u_{1t} - u_{2t}).$ It is clear that $E[e_{2t} | X_t] = 0$. Also, since $E[u_{4t} | \Delta P_t, X_t] = 0$ for any values of $X_3$, and $u_{4t}$. If the disturbance terms are assumed to be normally distributed and independent of $X_t$, then the conditional density of $\Delta P_t$ and $u_{3t}$ is also normal and, therefore, is completely specified by two conditional means and variances and by the covariance of $\Delta P_t$ and $u_{3t}$. The mean and variance of $u_{3t}$ are given by the specifications of the model; the mean and variance of $\Delta P_t$, as well as the covariance of $\Delta P_t$ with $u_{3t}$, are easily derived from the reduced form equation for $\Delta P_t$.

As an example, the reduced form equation for $\Delta P_t$ is linear in the elements of $X_t$ and $u_{1t}, u_{2t},$ and $u_{4t}.$ If the disturbance terms are normally distributed and independent of $X_t$, then the conditional density of $\Delta P_t$ and $u_{3t}$ is also normal and, therefore, is completely specified by two conditional means and variances and by the covariance of $\Delta P_t$ and $u_{3t}$. The mean and variance of $u_{3t}$ are given by the specifications of the model; the mean and variance of $\Delta P_t$, as well as the covariance of $\Delta P_t$ with $u_{3t}$, are easily derived from the reduced form equation for $\Delta P_t$.

Note that for certain specifications of the disturbance terms (e.g., normality) this function will involve integrals. More will be said concerning this function below.
use, it follows from the price equation that $E[u_{4i}(D_i - S_i), X_i] = 0$. Therefore, $E[e_{1i}|X_i] = 0$. Thus, $e_{1i}$ and $e_{2i}$ can be considered as disturbance terms.

Unlike the remaining terms of (27a) and (27b), $\Omega_i$ depends directly upon the disturbance terms $u_{1i}$ and $u_{2i}$. Therefore $\Omega_i$ cannot be considered as part of the regression function; $\Omega_i$ also cannot be considered as a disturbance term because it involves the product of $H_i$ and $(u_{2i} - u_{1i})$ and so will not in general have a mean of zero. The procedure here, therefore, will be to abstract the mean of $\Omega_i$, conditional on $\Delta P_i$ and $X_i$, and incorporate it within the regression function so as to end up with a two-equation system based on the two endogenous variables, $Q_i$ and $\Delta P_i$.

Since the expectation of one variable conditional upon a set of others is, in general, a function of the conditioning variables, it follows that

$$E[\Omega_i|\Delta P_i, X_i] = H_i E[u_{2i} - u_{1i}|P_i, X_i] = H_i L(P_i, X_i),$$

where $L(P_i, X_i)$ is a function of the elements of $X_i$ and $P_i$. It is interesting to note that if $u_{1i}, u_{2i}$, and $u_{3i}$ are assumed to be jointly normal, the joint distribution of $(u_{2i} - u_{1i})$ and $P_i$ will be normal and, therefore, the conditional distribution of $u_{2i} - u_{1i}$ given $P_i$ will be normal; hence, the function $L$ in (28) will be linear in $P_i$ and $X_i$. In any event, the parameters of $L$ will be functions of the parameters of the demand, supply, and price equations, as well as of the variances and covariances of $u_{1i}, u_{2i}$, and $u_{3i}$.

In light of equation (28), it follows that $\Omega_i$ can be expressed as

$$\Omega_i = H_i L(P_i, X_i) + \theta_i,$$

where $E[\theta_i|P_i, X_i] = 0$, and so $E[\theta_i|X_i] = 0$. Therefore, equations (27a) and (27b) can be expressed as

(30a) \[ Q_i = H_i S_i^* + D_i^* - H_i D_i^* + H_i L(P_i, X_i) + \psi_i \]

and

(30b) \[ \Delta P_i = \beta_5(D_i^* - S_i^*) + X_{3i}\beta_6 + e_{2i} \]

$$= X_{1i}(\beta_5 \beta_1) + P_i(\beta_5 \beta_2) - X_{2i}(\beta_5 \beta_3) - P_i(\beta_5 \beta_4) + X_{3i}\beta_6 + e_{2i},$$

where $\psi_i = e_{1i} + \theta_i$, and so $E[\psi_i|X_i] = 0$.

Equations (30a) and (30b) form a simultaneous two-equation system for $Q_i$ and $P_i$, which is nonlinear in the parameters and also in one of the endogenous variables, $P_i$. Aside from the maximum likelihood technique, general results concerning the estimation of such systems are not available. The difficulty in applying the maximum likelihood technique to the system (30a)–(30b) is that the joint distribution of $\psi_i$ and $e_{2i}$ will be, for just about any specification of $u_{1i}, u_{2i}$, and $u_{3i}$, quite complicated. However, a consistent estimation technique can be

\[^{10}\text{Note that although } P_i \text{ does not depend directly on } Q_i \text{ in (30b), the system is fully simultaneous because } \psi_i \text{ and } e_{2i} \text{ are correlated, e.g., they both contain } u_{1i}.\]
developed, subject to certain approximations, using the method of moments because $E[y_t|X_t^1] = E[e_t|X_t^1] = 0$.

First, it will be assumed that $u_{1t}$, $u_{2t}$, and $u_{3t}$ are jointly normal so that $L(P_t, X_t)$ is linear in $P_t$ and $X_t$. Second, it will be assumed that the error of approximation in the expansion of $H_t$ (as a function of $P_t$ and $X_t$) in a Taylor series is negligible after a finite number of terms. Now, substituting the expression for $L(P_t, X_t)$ and the polynomial expansion of $H_t$ into (30a) yields an equation of the form

$$(30a') Q_t = Z_{1t}y_1 + Z_{2t}y_2 + \psi_t,$$

where $Z_{1t}$ is a row vector of observations on known polynomial functions of the predetermined variables, $X_t$, and $Z_{2t}$ is a row vector of observations on known polynomial functions of the endogenous variable $P_t$ and predetermined variables $X_t$. The vectors $y_1$ and $y_2$ are vectors of parameters, the elements of which are nonlinear functions of $\beta_1$ through $\beta_6$ and of the variances and covariances of $u_{1t}$, $u_{2t}$, and $u_{3t}$. The order of these vectors depends on the degree of the polynomial expansion.

Equations (30a') and (30b) form a two-equation system that contains nonlinear, but known, endogenous functions and nonlinear restrictions on the parameters—a system, in other words, that is nonlinear in both variables and parameters. The system may be consistently estimated by a nonlinear two-stage least squares procedure. Specifically, since $Z_{2t}$ is the vector of endogenous functions in (30a'), each element of $Z_{2t}$, one of which is $P_t$, can be regressed on the elements of $Z_{1t}$, and on the predetermined variables in (30b) as well as on powers of these variables.13

Let $\bar{Z}_{2t}$ and $\bar{P}_t$ denote the predicted values of the elements of $Z_{2t}$ and $P_t$. Now, as will be shown below, the basic parameters of the system $\beta_1$ through $\beta_6$ and the variances and covariances of $u_{1t}$, $u_{2t}$, and $u_{3t}$, can be estimated by minimizing

$$(31) \quad S = (Q - Z_1y_1 - \hat{Z}_2y_2)(Q - Z_1y_1 - \hat{Z}_2y_2) + \alpha(\Delta P - X_1\beta_2\beta_1 - \hat{P}\beta_5\beta_2 + X_2\beta_5\beta_3 + \hat{P}\beta_5\beta_4 - X_3\beta_5\beta_6)(\Delta P - X_1\beta_2\beta_1 - \hat{P}\beta_5\beta_2 + X_2\beta_5\beta_3 + \hat{P}\beta_5\beta_4 - X_3\beta_5\beta_6).$$

where $Q, Z_1, \hat{Z}_2, \Delta P, X_1, X_2, X_3,$ and $\hat{P}$ are the vectors and matrices of observations on the corresponding elements, and $\alpha$ is any nonnegative number. If $\alpha$ is taken to be zero, then only information regarding equation (30a') is used to obtain the estimates, whereas if $\alpha$ is taken to be positive, then information regarding equation (30b) is also used in obtaining the estimates.

11 Note, if $u_{1t}$, $u_{2t}$, and $u_{3t}$ are jointly normal, $H_t$ will be of the form

$$H_t = \int_{-\infty}^{\mu_t(P_t, X_t)} f(Z) dZ,$$

where $f(Z)$ is the density for a standard normal variable and $L_t(P_t, X_t)$ is linear in $X_t$ and $P_t$; see (25) and Footnote 7. Therefore, the expansion of $H_t$ is straightforward. In a sense, most econometric systems may be considered as depending upon such polynomial approximations; see, for example, Fisher [6, pp. 127–29].

12 For example, one such function might be $P_t^2 X_{1t}$, where $X_{1t}$ is a predetermined variable.

13 See Kelejian [10].
For a given value of \( \alpha \), the minimization of \( S \) in (31) with respect to the parameters \( \beta_1 \) through \( \beta_6 \) and the variances and covariances of \( u_{1t} \), \( u_{2t} \), and \( u_{3t} \) is quite straightforward and should be able to be handled by nonlinear optimization programs like those used in Section 2. If first and second derivatives of \( S \) are required, these can be computed numerically or else one can go to the bother of actually differentiating \( S \) twice with respect to \( \beta_1 \) through \( \beta_6 \). Indeed, the problem of minimizing \( S \) in (31) does not appear to be as difficult as was the problem of maximizing \( \mathcal{L} \) in (8) since the maximization of \( \mathcal{L} \) required the evaluation of normal integrals, where the limits of the integrals were themselves functions of some of the parameter values. The steps involved in computing the estimates of the parameters of equations (30a) and (30b) are tedious because of the need to expand \( H_t \) in a Taylor series and the need to express \( L(P_t, X_t) \) as an explicit function of \( P_t, X_t \), and the parameter values, but aside from the tediousness the computation of the estimates does not appear infeasible or impractical.

The choice of the value of \( \alpha \) in (31) is somewhat arbitrary. A choice of a value of one means that both equations are weighted equally, and this may be as good a choice as any. One might also want to consider a two-step procedure in which initial estimates of the parameter values are obtained by, say, using \( \alpha = 1 \), then estimating the variances of \( \psi_1 \) and \( e_{2t} \), and then reestimating the parameters taking \( \alpha \) to be the ratio of the estimated variance of \( \psi_1 \) to the estimated variance of \( e_{2t} \). It remains to be shown that the minimization of (31) yields consistent parameter estimates. To see this, let \( \hat{\Psi}_t = Z_{2t} - \hat{Z}_{2t} \) and \( \hat{\Psi}_t = P_t - \hat{P}_t \). Also, rewrite equation (30b) as

\[
(32) \quad \Delta P_t = Z_{3t} \hat{\gamma}_3 + P_t \hat{\gamma}_4 + e_{2t},
\]

where \( Z_{3t} = (X_{1t}, X_{2t}, X_{3t}) \) and \( \hat{\gamma}_3 \) and \( \hat{\gamma}_4 \) are the corresponding vectors of parameters, the elements of which are nonlinear functions of \( \beta_1 \) through \( \beta_6 \). Equations (30a') and (32) can be written:

\[
(33) \quad \hat{Q}_t = Z_{1t} \hat{\gamma}_1 + \hat{Z}_{2t} \hat{\gamma}_2 + \psi_t + \hat{\Psi}_t \hat{\gamma}_2
\]

and

\[
(34) \quad \Delta P_t = Z_{3t} \hat{\gamma}_3 + P_t \hat{\gamma}_4 + e_{2t} + \hat{\Psi}_t \hat{\gamma}_4.
\]

Now, let

\[
(35) \quad Y = \begin{bmatrix} \hat{Q} \\ \sqrt{\alpha} \Delta P_t \end{bmatrix}, \quad Z = \begin{bmatrix} Z_{1t} & \hat{Z}_{2t} & 0 & 0 \\ 0 & 0 & \sqrt{\alpha} Z_3 & \sqrt{\alpha} P_t \end{bmatrix}, \quad \gamma = \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \\ \gamma_4 \end{bmatrix},
\]

\[
\varepsilon = \begin{bmatrix} \psi \\ \sqrt{\alpha} e_2 \end{bmatrix}, \quad \hat{\varepsilon} = \begin{bmatrix} \hat{\psi}_1 \\ 0 \end{bmatrix}, \quad \hat{\gamma}_w = \begin{bmatrix} \gamma_2 \\ \gamma_4 \end{bmatrix}.
\]

\[\text{It should be noted, of course, that the variance of } \psi_1 \text{ is not constant over the sample period, and so what one is obtaining using this procedure is an estimate of the average variance of } \psi_1 \text{ over the sample period.}\]
where the dropping of the $t$ subscripts means that the symbols refer to vectors or matrices of observations. Equations (33) and (34) can now be written, after (34) is multiplied across by $\sqrt{a}$, as

\begin{equation}
Y = Z\gamma + \epsilon + \bar{V}\gamma^*.
\end{equation}

Using this notation, $S$ in (31) is

\begin{equation}
S = (Y - Z\gamma)'(Y - Z\gamma)
= Y'Y - 2\gamma'Z'Y + \gamma'Z'Z\gamma.
\end{equation}

Let $\beta$ be the vector of parameters consisting of $\beta_1$ through $\beta_6$ and of the variances and covariances of $u_{1t}$, $u_{2t}$, and $u_{3t}$. Minimizing $S$ with respect to $\beta$ yields

\begin{equation}
\frac{\partial S}{\partial \beta} = -2\gamma'Z'Y + 2\gamma'Z'Z\gamma = 0,
\end{equation}

where $\gamma$ is $\gamma$ evaluated at $\beta$ and $\frac{\partial}{\partial \beta}$ is the matrix of partial derivatives $\partial \gamma/\partial \beta$ evaluated at $\beta$. Linearizing (38) about $\beta$ yields

\begin{equation}
\gamma'Z'[Z\gamma + Z\gamma(\bar{\beta} - \beta) - Y] = 0,
\end{equation}
or, using (36),

\begin{equation}
\bar{\beta} - \beta = (\gamma'Z'Z\gamma)^{-1}\gamma'Z'(\epsilon + \bar{V}\gamma^*).
\end{equation}

Since $\text{plim}_{T \to \infty} T^{-1}Z\epsilon = 0$, $\text{plim}_{T \to \infty} (\bar{\beta} - \beta)$ will be zero if $\text{plim}_{T \to \infty} T^{-1}(\gamma'Z'\bar{V}\gamma^*)$ is zero. Now,

\begin{equation}
Z'\bar{V} = \begin{bmatrix}
Z_1'\bar{V}_1 & 0 \\
Z_2'\bar{V}_1 & 0 \\
0 & \alpha Z_3'\bar{V}_2 \\
0 & \alpha \bar{P}'\bar{V}_2
\end{bmatrix}.
\end{equation}

Since it was assumed that all of the predetermined variables were used in constructing the calculated values, it follows by the least squares property that $Z'\bar{V} = 0$; therefore, $\bar{\beta}$ is consistent.

5. CONCLUSION

The study of the estimation of disequilibrium models has become quite popular recently. In this paper three methods of estimating disequilibrium models have been proposed. The first two methods—maximum likelihood methods—are concerned with the estimation of models in which the price equation is a non-

---

15 The following derivation is similar, although in a different context, to a derivation given by Aigner and Goldberger [1, p. 715, n. 1].

16 In addition to the study of Fair and Jaffe [4], the following studies are concerned in one way or another with the question of estimating disequilibrium models: Goldfeld and Quandt [8], Quandt [13], Goldfeld, Kedleston, and Quandt [7], Brown and Durbin [2], Farley and Hinich [5], and McGee and Carleton [11]. See Quandt [14] for a brief review of these studies.
stochastic function of excess demand. The third method is concerned with the estimation of a more general model in which the price equation is allowed to be a multivariate, stochastic function. The problems involved in estimating disequilibrium models turn out to be fairly complicated, and for this reason one may in practice want to begin with the estimation of simply-specified models before considering more general models. Nevertheless, it is encouraging that the quite general model considered in Section 4 of this paper appears capable of estimation.

Princeton University
and
New York University

Manuscript received February, 1972; revision received January, 1973.

REFERENCES