

Estimation in A Disequilibrium Model and
the Value of Information

Stephen M. Goldfeld
Richard E. Quandt

Econometric Research Program
Research Memorandum No. 169
July 1974

Estimation in a Disequilibrium Model and
the Value of Information^{*}

by

Stephen M. Goldfeld

Richard E. Quandt

1. Introduction

Preliminary Notions. A number of recent papers has studied the problem of estimating demand and supply functions for markets that are in disequilibrium. By disequilibrium we mean a state in which the quantities demanded (D) and supplied (S) are not equal to one another and in which the observed quantity Q is given by the short side of the market, i.e. $Q = \min(D,S)$. Such markets present a number of novel problems of specification as well as of estimation. These issues were first raised by Fair and Jaffee in their treatment of the market for housing starts [2] and were subsequently considered, among others, by Fair and Kelejian [3] , Amemiya [1] , Hartley [6] and Maddala and Nelson [7] . Some theoretical (as contrasted with econometric) issues are discussed by Grossman [5] .

Despite this spate of recent articles, the study of disequilibrium models is still in its infancy, especially with respect to actual computational

* We gratefully acknowledge comments from the members of the NBER-NSF Workshop on Segmented and Switching Regressions held June 3-4, 1974 at the University of Wisconsin and in particular those of G.S. Maddala. We are indebted to Rehka Nadkarni for computer programming. Financial support was received from NSF Grant GS-43747X.

experience. As far as we know, such experience has been exclusively with the data used by Fair and Jaffee. Furthermore, there are some formulations which, due to the presence of complex nonlinearities, have not even been estimated. These reasons alone point up the need for further study of disequilibrium models. An additional reason for such study is to spell out more precisely the relationship between the appropriate estimating technique and the amount of information assumed to be available to the investigator. Since this relationship may vary from underlying model to model, it is also desirable to investigate a model with somewhat different features than the housing market.

In particular, the present paper has several specific objectives.

- (1) In the remainder of this section we present a brief review of the basic demand-supply disequilibrium models and also introduce a somewhat richer model stemming from Suits' work [9] on the watermelon market.
- (2) In Section 2 we develop maximum likelihood methods for estimating the coefficients of this model under a variety of assumptions concerning how much information is available and analyze some theoretical properties of the resulting likelihood functions.
- (3) In Section 3 we report the results of reestimating the Suits model with the new techniques and also extensively examine the computational aspects and estimation properties of the various methods with the aid of Monte Carlo experiments.
- (4) In the same section we analyze the value of information in the context of the Suits model. In particular, we compare in detail two of the three estimators, introduced in Section 2, which make use of differing amounts of information or data.

Supply and Demand Disequilibrium. The simplest form of the Fair-Jaffee model is as follows:

$$D = D(x^D, p) + u_1 \quad (1-1)$$

$$S = S(x^S, p) + u_2 \quad (1-2)$$

$$Q = \min(D, S) \quad (1-3)$$

where for simplicity the price p is taken to be exogenous (although this is by no means necessary) and where x^D and x^S represent vectors of other variables entering the demand and supply functions. The observed variables are p and Q , with D and S being latent or unobserved. Diagrammatically this situation is as in Figure 1 with the observations clustered around the line segments AB and BC . More complicated models of this type in which p is endogenous are also discussed in [2] and [7]. Since in the present model Q is the only observable random variable among Q , D and S , the problem is to derive the pdf of Q from the joint pdf of D and S . Denoting by $f(Q|D<S)$ the conditional pdf of Q given $D < S$, we have from the specification of the problem

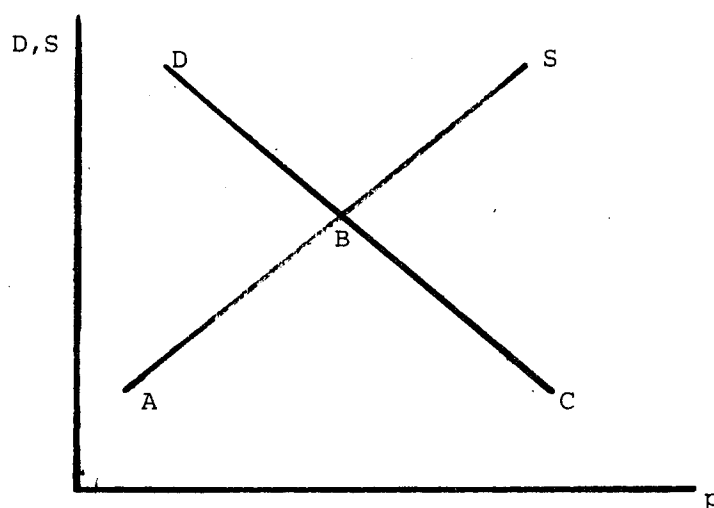
$$h(Q) = f(Q|D<S) \Pr\{D<S\} + f(Q|D \geq S) (1 - \Pr\{D<S\}) \quad (1-4)$$

But

$$f(Q|D<S) = \int_Q^\infty g(Q, S|D<S) ds = \frac{1}{\Pr\{D<S\}} \int_Q^\infty g(Q, S) ds \quad (1-5)$$

where $g(Q, S)$ and $g(Q, S|D<S)$ respectively denote the joint and joint conditional pdf of D and S with Q replacing D , and a similar expression

Figure 1



holds for the second part of (1-6) . Hence (1-4) becomes

$$h(Q) = \int_Q^{\infty} g(Q, S) dS + \int_Q^{\infty} g(D, Q) dD \quad (1-6)$$

where in the two terms either $Q = D$ and we integrate out S over $Q \leq S$ or $Q = S$ and we integrate out D over $Q \leq D$.

Indexing observations by i ,

$$L = \prod_i h(Q_i) \quad (1-7)$$

A slightly more complicated model arises if p is made endogenous by

¹Equation (1-4) has the general form of the sum of two densities, each weighted by a certain probability and thus appears to be superficially similar to the λ -weighted densities considered in [8] . However, in the present case, the probabilities are not constant from observation to observation and furthermore they cancel out altogether when the pdf is written in the form of joint rather than conditional densities in (1-6).

the addition to (1-1) , (1-2) , (1-3) of, say, a dynamic adjustment equation

$$\Delta P = \gamma(D-S) + \beta X^P + u_3 \quad (1-8)$$

where X^P is a vector of some additional regressors. We now require the joint pdf of the observable random variables Q and p which is derived from the joint pdf $g(D,S,p)$ analogously to the previous case and is

$$h(Q,p) = \int_Q^\infty g(Q,S,p) dS + \int_Q^\infty g(D,Q,p) dD \quad (1-9)$$

In both of the previous formulations it was implicitly assumed that D and S are unobservable. In the contrary case, and employing the first and simpler model for purposes of illustration, we let I_1 denote the set of indices for which $D < S$ and I_2 the set for which $D \geq S$. The density function for Q then is

$$h(Q) = \begin{cases} f(Q | D < S) \Pr\{D < S\} & \text{if } D < S \\ f(Q | D \geq S) (1 - \Pr\{D < S\}) & \text{otherwise} \end{cases} \quad (1-10)$$

The corresponding likelihood function then is

$$L = \prod_{I_1} f(Q | D < S) \Pr\{D < S\} \prod_{I_2} f(Q | D \geq S) (1 - \Pr\{D < S\}) =$$

$$= \prod_{I_1} \int_Q^{\infty} g(Q,S) dS \prod_{I_2} \int_Q^{\infty} g(D,Q) dD \quad (1-11)$$

which follows from applying (1-5) .

Computational experience with models of this type and particularly with the likelihood function arising from (1-9) has been limited; whatever there exists so far is based on the Fair-Jaffee data. In view of these reasons it appears fortunate that we encountered Suits' early article. This turned out to be even more propitious in the light of the fact that the Suits model contains some new complexities as well as a more natural way in which questions about the value of information or data can be introduced.

The Suits Model. A particularly interesting disequilibrium model which presents some new complexities for the derivation of the likelihood function is Suits' model of the watermelon market. Define the following variables, measured in terms of natural logarithms:

q_t = crop of watermelons

p_t = price of watermelons

x_t = amount of watermelons harvested

C_t = price of cotton

T_t = price of commercial truck

J_t = dummy variable representing government cotton policy

K_t = dummy variable representing World War II

w_t = southern farm wage rate

Y_t = disposable income

N_t = population

F_t = freight cost .

The equations of the model are as follows

$$q_t = a_1 p_{t-1} + a_2 C_{t-1} + a_3 J_t + a_4 T_{t-1} + a_5 K_t + a_6 + u_{1t} \quad (1-12)$$

$$x_t = a_7 p_t - a_7 w_t + a_8 q_t + a_9 + u_{2t} \quad (1-13)$$

$$p_t = a_{10} Y_t + a_{11} x_t - (a_{10} + a_{11}) N_t + a_{12} F_t + a_{13} + u_{3t} \quad (1-14)$$

if $x_t < q_t$ and

$$q_t = a_1 p_{t-1} + a_2 C_{t-1} + a_3 J_t + a_4 T_{t-1} + a_5 K_t + a_6 + u_{1t} \quad (1-15)$$

$$p_t = a_{10} Y_t + a_{11} q_t - (a_{10} + a_{11}) N_t + a_{12} F_t + a_{13} + u_{3t} \quad (1-16)$$

if $x_t = q_t$.

Equation (1-12) (or(1-15)) describes the determination of the crop and states that the intended crop is a function of the lagged own price and the lagged prices of competing crops and some other variables. Equation (1-14) is a standard demand equation in per capita terms and states that the current price is a function of the amount brought to the market, income, etc. Equation (1-13) is the "harvest equation" and states that the amount actually harvested is a function of the current price and wage and of the size of the crop itself. This equation thus states that under certain circumstances (e.g., low current price and high

current wage) it may not be worthwhile to harvest the entire crop; of course, if the equation were to require more to be harvested than had actually grown, the equation would not be required to hold and the harvest would equal the crop, i.e. $x_t = q_t$.

It is not our intention to analyze in detail whether (1-12) through (1-16) are a suitable specification of the market for watermelons and these questions are discussed in [9] . It is clear that alternative specifications are possible and that one might wish to introduce, (a) a more explicit optimization model for the determination of the harvested quantity and (b) an adaptive expectations model for the price in the q-equation. However, we shall concentrate on the model essentially as stated for two principal reasons: (1) it presents an interesting extension of disequilibrium models and (2) by alternately specifying q_t to be observed or unobserved it permits us to estimate econometrically the value of information. More specifically, a model such as (1-12) through (1-16) may be thought to hold under one of two polar extreme assumptions: (1) that we are given observations on all the variables, including q_t (and this is the model adopted by Suits himself) and (2) that we have no direct observations between the submodel consisting of (1-12) to (1-14) and the submodel consisting of (1-15) , (1-16) . This is analogous to having observations on D and S in the simple models discussed earlier which led to (1-11) as the likelihood function and to lacking such information which leads to a quite different likelihood function. We now turn to the derivation of these functions.

2. Derivation of the Likelihood Functions

In order to simplify the derivations it is convenient to rewrite the equations of the Suits model in somewhat condensed form by letting the scalars, z_{1t} , z_{2t} , z_{3t} represent various (sets of) predetermined variables. The basic equations of the model are

$$q_t = b_1 z_{1t} + b_2 + u_{1t} \quad (2-1)$$

$$x_t = b_3 p_t + b_4 q_t + b_5 z_{2t} + b_6 + u_{2t} \quad (2-2)$$

$$p_t = b_7 z_{3t} + b_8 x_t + b_9 + u_{3t} \quad (2-3)$$

In replacing the predetermined variables in (2-1) by z_{1t} we are consciously neglecting the fact that p_{t-1} is endogenous. The pdf we derive for the endogenous variables for the t th observation will be conditional on p_{t-1} and the resulting likelihood function will be correct if we assume that p_0 is nonstochastic.

According to the original specification the values of the endogenous variables are generated by (2-1), (2-2) and (2-3) if $x_t < q_t$; if $x_t = q_t$ then (2-2) is ineffective. The nature of this problem is illustrated in Figure 1.

Quantities are measured on the vertical axis and price on the horizontal one. According to normal expectations equation (2-2) is positively sloped and (2-3) negatively. Equation (2-2) depends on q_t and hence the line representing it would shift as q_t changes; we neglect this dependence in the diagram without losing the essential features of the problem. The quantity of crop q_t depends on exogenous variables only and represents farmers' past intentions; hence if its magnitude is given by q_1 , the solution

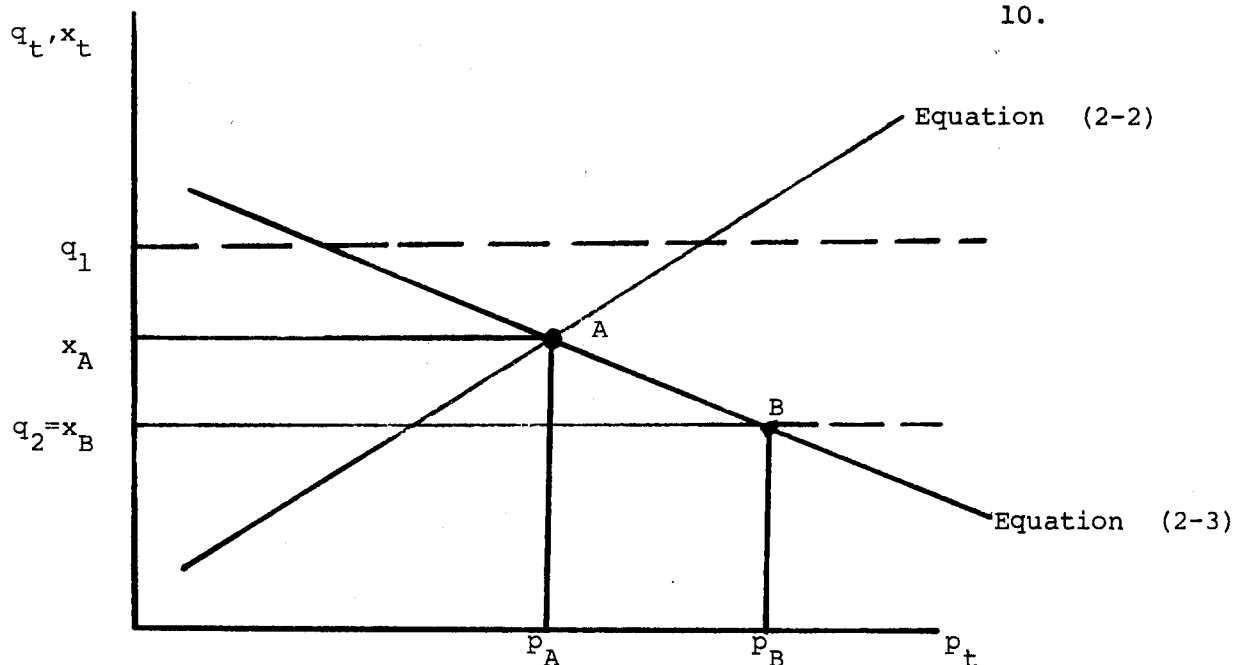


Figure 1

to (2-2) and (2-3) yields a desired harvest or x -value given by $x_A < q_1$ and thus the price-harvest quantity pair generated by nature is (p_A, x_A) corresponding to the intersection A . If, on the other hand, the quantity of crop q_2 is less than x_A , then the amount of the harvest is $x_B = q_2$ and the generated pair is (p_B, x_B) corresponding to point B . An entire sample will contain a mixture of points corresponding to A and B respectively.

We now turn to deriving the likelihood function under the two alternative specifications that q_t is or is not observable. In any event we shall assume that $u_{it} \sim N(0, \sigma_i^2)$ ($i = 1, 2, 3$), $E(u_{it} u_{js}) = 0$ for all $i \neq j$ and $t \neq s$.

Specification 1: q_t unobserved. It is obvious from Figure 1 that the quantity that must be substituted in Equation (2-3) to find the observed price is the lesser of q_t and x_t , i.e., the lesser of the crop and the intended harvest. Defining y_t to be the observed quantity coming to the market, the model can be rewritten as follows:

$$q_t = b_1 z_{1t} + b_2 + u_{1t} \quad (2-4)$$

$$x_t = b_3 p_t + b_4 q_t + b_5 z_{2t} + b_6 + u_{2t} \quad (2-5)$$

$$p_t = b_7 z_{3t} + b_8 y_t + b_9 + u_{3t} \quad (2-6)$$

$$y_t = \min(q_t, x_t) \quad (2-7)$$

where the quantity variable introduced into the price equation (2-6) is y_t , which is given by (2-7)². Under the present specification only p_t and y_t are observed and our next step will be to deduce the joint pdf of p_t and y_t from that of u_{1t} , u_{2t} , u_{3t} . Since only one quantity variable is observable in the present case, the specification given by (2-4) to (2-7) can be transformed into two sets of equations corresponding to the regime in which $x_t < q_t$ and the regime $x_t = q_t$. For this purpose we introduce v_{1t} , v_{2t} to represent the quantity in each regime and w_{1t} , w_{2t} the corresponding price. Then, since q_t is unobserved, we have

$$v_{1t} = b_3 w_{1t} + b_4 (b_1 z_{1t} + b_2) + b_5 z_{2t} + b_6 + (u_{2t} + b_4 u_{1t}) \quad (2-8)$$

$$w_{1t} = b_7 z_{3t} + b_8 v_{1t} + b_9 + u_{3t} \quad (2-9)$$

if $v_{1t} < v_{2t}$ and

$$v_{2t} = b_1 z_{1t} + b_2 + u_{1t} \quad (2-10)$$

$$w_{2t} = b_7 z_{3t} + b_8 v_{2t} + b_9 + u_{3t} \quad (2-11)$$

²The reader will note the similarity to Equations (1-1) to (1-3) and (1-8), with the exception however, that in the present example the observed quantity does enter the price equation, thus departing from the recursive nature of the previous model; i.e. the usual Fair-Jaffee model.

otherwise. Equation (2-8) is obtained by substituting (2-4) into (2-5) and replacing x_t by v_{1t} , p_t by w_{1t} ; (2-9) is the same as (2-6) with the same substitution of symbols. We in fact observe (y_t, p_t) either as (v_{1t}, w_{1t}) or as (v_{2t}, w_{2t}) depending on whether $v_{1t} < v_{2t}$ or not. The random variables $v_{1t}, w_{1t}, v_{2t}, w_{2t}$ have a singular normal distribution and it can be verified that $w_{1t} - w_{2t} = b_8(v_{1t} - v_{2t})$; however, the set of random variables, v_{1t}, w_{1t}, v_{2t} has a nonsingular normal distribution and so does the set v_{1t}, w_{2t}, v_{2t} ; both induced by the distributions of u_{1t}, u_{2t}, u_{3t} .

We derive each distribution in turn. Equations (2-8), (2-9) and (2-10) are a system of simultaneous equations with a normally distributed error term vector $\eta_t' = (\eta_{1t}, \eta_{2t}, \eta_{3t}) = (u_{2t} + b_4 u_{1t}, u_{3t}, u_{1t})$ with $E(\eta_t) = 0$ and

$$\Sigma_1 = E(\eta_t \eta_t') = \begin{bmatrix} \sigma_2^2 + b_4^2 \sigma_1^2 & 0 & b_4 \sigma_1^2 \\ 0 & \sigma_3^2 & 0 \\ b_4 \sigma_1^2 & 0 & \sigma_1^2 \end{bmatrix} \quad (2-12)$$

Since $\det(\Sigma_1) = \sigma_1^2 \sigma_2^2 \sigma_3^2$, the joint pdf of v_{1t}, w_{1t}, v_{2t} then is

$$f(v_{1t}, w_{1t}, v_{2t}) = \frac{|1 - b_3 b_8|}{(2\pi)^{3/2} \sigma_1 \sigma_2 \sigma_3} \exp \left\{ -\frac{1}{2} \left[\eta_t' \Sigma_1^{-1} \eta_t \right] \right\} \quad (2-13)$$

where we substitute for the η vector from (2-8), (2-9) and (2-10) and where $1 - b_3 b_8$ is the Jacobian of the transformation.

The second distribution is obtained by replacing w_{1t} in (2-8) by $w_{2t} + b_8(v_{1t} - v_{2t})$ and rearranging to yield the three equations

$$v_{1t} = \frac{1}{1-b_3 b_8} \left[b_3 w_{2t} - b_3 b_8 v_{2t} + b_4 b_1 z_{1t} + b_5 z_{2t} + b_4 b_2 + b_6 \right] + \frac{u_{2t} + b_4 u_{1t}}{1-b_3 b_8} \quad (2-14)$$

$$w_{2t} = b_7 z_{3t} + b_8 v_{2t} + b_9 + u_{3t} \quad (2-15)$$

$$v_{2t} = b_1 z_{1t} + b_2 + u_{1t} \quad (2-16)$$

with error term vector $\xi_t = \left(\frac{u_{2t} + b_4 u_{1t}}{1-b_3 b_8}, u_{3t}, u_{1t} \right)$ and $E(\xi_t) = 0$,

$$\Sigma_2 = E(\xi_t \xi_t') = \begin{bmatrix} \frac{\sigma_2^2 + b_4^2 \sigma_1^2}{(1-b_3 b_8)^2} & 0 & \frac{b_4 \sigma_1^2}{1-b_3 b_8} \\ 0 & \sigma_3^2 & 0 \\ \frac{b_4 \sigma_1^2}{1-b_3 b_8} & 0 & \sigma_1^2 \end{bmatrix} \quad (2-17)$$

and $\det(\Sigma_2) = \sigma_1^2 \sigma_2^2 \sigma_3^2 / (1-b_3 b_8)^2$. Hence, $g(v_{1t}, w_{2t}, v_{2t})$ is

$$g(v_{1t}, w_{2t}, v_{2t}) = \frac{|1-b_3 b_8|}{(2\pi)^{3/2} \sigma_1 \sigma_2 \sigma_3} \exp \left\{ -\frac{1}{2} \left[\xi_t' \Sigma_2^{-1} \xi_t \right] \right\} \quad (2-18)$$

where we replace the ξ -vector from (2-14), (2-15), (2-16) and where the Jacobian of the transformation is unity but the term $|1-b_3b_8|$ reappears nevertheless as a result of its appearance in $\det(\sum_2)$.

The joint density of y_t, p_t is clearly given by

$$h(y_t, p_t) = h_1(y_t, p_t | v_{1t} < v_{2t}) \Pr\{v_{1t} < v_{2t}\} + h_2(y_t, p_t | v_{1t} \geq v_{2t}) (1 - \Pr\{v_{1t} < v_{2t}\}) \quad (2-19)$$

$$\text{but} \quad h_1(y_t, p_t | v_{1t} < v_{2t}) = \int_{y_t}^{\infty} f(y_t, p_t, v_{2t} | v_{1t} < v_{2t}) dv_{2t} =$$

$$\frac{1}{\Pr\{v_{1t} < v_{2t}\}} \int_{y_t}^{\infty} f(y_t, p_t, v_{2t}) dv_{2t}$$

and similarly for $h_2(y_t, p_t | v_{1t} \geq v_{2t})$. Hence (2-19) becomes³

$$h(y_t, p_t) = \int_{y_t}^{\infty} f(y_t, p_t, v_{2t}) dv_{2t} + \int_{y_t}^{\infty} g(v_{1t}, p_t, y_t) dv_{1t} \quad (2-20)$$

We now obtain the first term of (2-20) by integrating (2-13) and the second term by integrating (2-18).

Consider the exponent part of (2-13) and substitute for the η 's from (2-8) - (2-10). We obtain for $\eta_t \sum_{l=1}^{-1} \eta_t$

$$\begin{aligned} \eta_t \sum_{l=1}^{-1} \eta_t &= \frac{(v_{1t}^{-b_3} w_{3t}^{-b_4} z_{4t}^{-b_5} z_{2t}^{-b_4} b_2^{-b_6})^2}{\sigma_2^2} + \frac{(w_{1t}^{-b_7} z_{3t}^{-b_8} v_{1t}^{-b_9})^2}{\sigma_3^2} + \\ &+ \frac{(v_{2t}^{-b_1} z_{1t}^{-b_2})^2 (\sigma_2^2 + b_4^2 \sigma_1^2)}{\sigma_1^2 \sigma_2^2} - \frac{2b_4 (v_{1t}^{-b_3} w_{3t}^{-b_4} z_{4t}^{-b_5} z_{2t}^{-b_4} b_2^{-b_6}) (v_{2t}^{-b_1} z_{1t}^{-b_2})}{\sigma_2^2} \end{aligned}$$

³Equation (2-20) is analogous to (1-9) but with the additional complication that unlike (1-9), the two integrals in (2-20) involve different functions.

Completing the square on v_{2t} , substituting the observed y_t, p_t for v_{1t}, w_{1t} , and rearranging we obtain for the first term of (2-20)

$$\int_{y_t}^{\infty} f(y_t, p_t, v_{2t}) dv_{2t} = \frac{|1-b_3b_8|}{2\pi\sigma_3(\sigma_2^2+b_4^2\sigma_1^2)^{1/2}} \exp\left\{-\frac{1}{2}\left[\frac{(y_t-b_3p_t-b_4b_1z_{1t}-b_5z_{2t}-b_4b_2-b_6)^2}{\sigma_2^2+b_4^2\sigma_1^2} + \frac{(p_t-b_7z_{3t}-b_8y_t-b_9)^2}{\sigma_3^2}\right]\right\} \int_{y_t}^{\infty} \frac{(\sigma_2^2+b_4^2\sigma_1^2)^{1/2}}{\sqrt{2\pi}\sigma_1\sigma_2} \exp\left\{-\frac{1}{2}\left[\frac{\sigma_2^2+b_4^2\sigma_1^2}{\sigma_1^2\sigma_2^2}\left[(v_{2t}-b_1z_{1t}-b_2)\right.\right.\right. \\ \left.\left.\left.-\frac{b_4\sigma_1^2(y_{1t}-b_3p_t-b_4b_1z_{1t}-b_5z_{2t}-b_4b_2-b_6)}{\sigma_2^2+b_4^2\sigma_1^2}\right]^2\right]\right\} dv_{2t} \quad (2-21)$$

Denoting the term before the integral sign in (2-21) by $\psi_1(y_t, p_t)$ and the cumulative integral of $N(0,1)$ from $-\infty$ to ℓ_t by $\Phi(\ell_t)$, (2-21) can be written as

$$\int_y^{\infty} f(y_t, p_t, v_{2t}) dv_{2t} = \psi_1(y_t, p_t) (1-\Phi(\ell_t)) \quad (2-22)$$

$$\text{where } \ell_t = \frac{y_t(\sigma_2^2+b_4^2\sigma_1^2)-\sigma_2^2(b_1z_{1t}+b_2)-b_4\sigma_1^2(y_t-b_3p_t-b_5z_{2t}-b_6)}{(\sigma_2^2+b_4^2\sigma_1^2)^{1/2}\sigma_1\sigma_2} \quad (2-23)$$

The second term of (2-20) is obtained in analogous fashion. Substituting (2-14) to (2-16) in the exponent part of (2-18) we obtain

$$\begin{aligned}
\xi_t' \xi_t^{-1} \xi_t &= \left[v_{1t} - \frac{1}{1-b_3b_8} (b_3w_{2t} - b_3b_8v_{2t} + b_4b_1z_{1t} + b_5z_{2t} + b_4b_2 + b_6) \right]^2 \frac{(1-b_3b_8)^2}{\sigma_2^2} \\
&+ \left[w_{2t} - b_7z_{3t} - b_8v_{2t} - b_9 \right]^2 / \sigma_3^2 + \left[v_{2t} - b_1z_{1t} - b_2 \right]^2 \frac{(\sigma_2^2 + b_4^2\sigma_1^2)}{\sigma_1^2\sigma_2^2} \\
&- 2 \left[v_{1t} - \frac{1}{1-b_3b_8} (b_3w_{2t} - b_3b_8v_{2t} + b_4b_1z_{1t} + b_5z_{2t} + b_4b_2 + b_6) \right] \left[v_{2t} - b_1z_{1t} - b_2 \right] \cdot \\
&\cdot b_4(1-b_3b_8) / \sigma_2^2
\end{aligned}$$

Completing the square on v_{1t} we obtain for the second part of (2-20)

$$\begin{aligned}
\int_{Y_t}^{\infty} g(v_{1t}, p_t, y_t) dv_{1t} &= \frac{|1-b_3b_8|}{(2\pi)^{3/2} \sigma_1 \sigma_2 \sigma_3} \exp \left\{ -\frac{1}{2} \left[\frac{(y_t - b_1z_{1t} - b_2)^2}{\sigma_1^2} + \frac{(p_t - b_7z_{3t} - b_8y_t - b_9)^2}{\sigma_3^2} \right] \right\} \\
&\cdot \int_{Y_t}^{\infty} \exp \left\{ -\frac{1}{2} \left[(1-b_3b_8) \left(v_{1t} - \frac{1}{1-b_3b_8} (b_3p_t - b_3b_8y_t + b_4b_1z_{1t} + \right. \right. \right. \\
&\left. \left. \left. + b_5z_{2t} + b_4b_2 + b_6) - b_4(y_t - b_1z_{1t} - b_2) \right) \right]^2 / \sigma_2^2 \right\} dv_{1t} \\
&= \frac{1}{2\pi\sigma_1\sigma_3} \exp \left\{ -\frac{1}{2} \left[\frac{(y_t - b_1z_{1t} - b_2)^2}{\sigma_1^2} + \frac{(p_t - b_7z_{3t} - b_8y_t - b_9)^2}{\sigma_3^2} \right] \right\} \cdot \\
&\int_{Y_t}^{\infty} \frac{1}{\sqrt{2\pi}\sigma_2} \exp \left\{ -\frac{1}{2} \left[(1-b_3b_8)v_{1t} - b_3p_t + (b_3b_8 - b_4)y_t - b_5z_{2t} - b_6 \right]^2 / \sigma_2^2 \right\} dv_{1t} \\
&= \psi_2(y_t, p_t) (1 - \Phi(\ell_t')) \tag{2-24}
\end{aligned}$$

where $\psi_2(y_t, p_t)$ is the term before the integral sign and

$$\ell'_t = \frac{1}{\sigma_2} \left[y_t (1-b_4) - b_3 p_t - b_5 z_{2t} - b_6 \right] \quad (2-25)$$

Hence the likelihood function has the form

$$L = \prod_{t=1}^n \left[\psi_1(y_t, p_t) (1-\Phi(\ell_t)) + \psi_2(y_t, p_t) (1-\Phi(\ell'_t)) \right] \quad (2-26)$$

Specification 1-A: q_t unobserved but sample separation known.

Somewhat more but still incomplete information is available if we assume that q_t is itself not generally observed, but that we have prior knowledge as to when x_t is less than and when it is equal to q_t .⁴ Let I_1 denote the set of indices for which $x_t < q_t$ and I_2 the set for which $x_t = q_t$. It follows immediately that the likelihood function is multiplicative as (1-11) but consists of the same parts as (2-26), hence

$$L = \prod_{t \in I_1} \psi_1(y_t, p_t) (1-\Phi(\ell_t)) \prod_{t \in I_2} \psi_2(y_t, p_t) (1-\Phi(\ell'_t)) \quad (2-27)$$

Specification 2: q_t observed. Once it is assumed (as in 1-A) that the sample separation is known, it may be more plausible to assume that it is actually observed. If this is the case we can decide by inspection whether

⁴Of course, this means that q_t is known when we have the information that $x_t = q_t$.

$x_t < q_t$ or $x_t = q_t$. This is analogous in the case of the demand-supply model of (1-1) and (1-2) of observing D and S. In that case the appropriate likelihood function is of the form (1-11).

Let I_1 and I_2 be the index sets as before. If q_t, x_t, p_t are all observed and $x_t < q_t$, the pdf of the observable variables is simply the joint pdf of q_t, x_t, p_t given by

$$\begin{aligned} \psi_3(q_t, x_t, p_t) = & \frac{|1-b_3b_8|}{(2\pi)^{3/2}\sigma_1\sigma_2\sigma_3} \exp \left\{ -\frac{1}{2} \left[\frac{(q_t - b_1z_{1t} - b_2)^2}{\sigma_1^2} + \right. \right. \\ & \left. \left. + \frac{(x_t - b_3p_t - b_4q_t - b_5z_{2t} - b_6)^2}{\sigma_2^2} + \frac{(p_t - b_7z_{3t} - b_8x_t - b_9)^2}{\sigma_3^2} \right] \right\} \end{aligned} \quad (2-28)$$

When $q_t = x_t$, the pdf of the observable variables x_t, p_t is simply $\psi_2(x_t, p_t)(1-\phi(\lambda_t'))$. Hence the likelihood function in the present case is

$$L = \prod_{t \in I_1} \psi_3(q_t, x_t, p_t) \prod_{t \in I_2} \psi_2(x_t, p_t)(1-\phi(\lambda_t')) \quad (2-29)$$

The Probability $\Pr\{v_{1t} < v_{2t}\}$. Although the probability $\Pr\{v_{1t} < v_{2t}\}$ cancels out in all likelihood functions, it is nevertheless of considerable interest as a quantity to be computed from estimates of the structural parameters since it allows us to estimate the probability that x_t was less than q_t for any t . This is, of course, of particular importance for

Specification 1.

In order to obtain $\Pr\{v_{1t} < v_{2t}\}$ obtain the reduced form for v_{1t} from (2-8) and (2-9). This yields

$$v_{1t} = \frac{1}{1-b_3b_8} \left[b_4b_1z_{1t} + b_5z_{2t} + b_3b_7z_{3t} + b_3b_9 + b_4b_2 + b_6 + b_4u_{1t} + u_{2t} + b_3u_{3t} \right]$$

and using this together with (2-10) in the probability statement yields

$$\Pr\{v_{1t} < v_{2t}\} = 1 - \Phi(\lambda_t'') \quad (2-30)$$

where

$$\lambda_t'' = \frac{b_1(b_4 - 1 + b_3b_8)z_{1t} + b_5z_{2t} + b_3b_7z_{3t} + b_3b_9 + (b_4 - 1)b_2 + b_2b_3b_6b_8}{(\sigma_1^2(1 - b_3b_8 - b_4)^2 + \sigma_2^2 + b_3^2\sigma_3^2)^{1/2}}$$

where we assumed that $1 - b_3b_8 > 0$ (which it normally will be).

Maximum Likelihood Estimation. Having obtained the likelihood functions under three separate specifications, the natural procedure for estimating the parameters is to maximize the corresponding likelihood function. Although the various likelihood functions are composed of similar or identical parts, their behavior is strikingly different. Some aspects of these differences are contained in the following two propositions.

Proposition 1. Maximum likelihood estimates for Specification 2 produce OLS estimates for the coefficients b_1, b_2, σ_1 .

Proof: We can factor the term $\exp \left\{ -\frac{1}{2} \frac{(a_t - b_1z_{1t} - b_2)^2}{\sigma_1^2} \right\} / \sigma_1 \sqrt{2\pi}$ from each

ψ_3 and each ψ_2 in (2-29). Equation (2-29) can then be written as

$$L = \left(\frac{1}{\sqrt{2\pi} \sigma_1} \right)^n \exp \left\{ - \frac{1}{2\sigma_1^2} \sum_{t=1}^n (q_t - b_1 z_{1t} - b_2)^2 \right\} K \quad (2-31)$$

where K does not depend on b_1 , b_2 and σ_1 . QED.

Proposition 2. The likelihood function for Specification 1 is unbounded and hence maximizing it produces inconsistent estimates.

Proof: We shall explore the behavior of the likelihood function at a point in parameter space such that $b_4 = 0$. Then (2-20) can be written as

$$\begin{aligned} h(y_t, p_t) &= \frac{1}{2\pi} \frac{|1 - b_3 b_8|}{\sigma_2 \sigma_3} \exp \left\{ - \frac{1}{2} \left[\frac{(y_t - b_3 p_t - b_5 z_{2t} - b_6)^2}{\sigma_2^2} + \frac{(p_t - b_7 z_{3t} - b_8 y_t - b_9)^2}{\sigma_3^2} \right] \right\} \times \\ &(1 - \Phi(l_t)) + \frac{1}{2\pi \sigma_1 \sigma_3} \exp \left\{ - \frac{1}{2} \left[\frac{(y_t - b_1 z_{1t} - b_2)^2}{\sigma_1^2} + \frac{(p_t - b_7 z_{3t} - b_8 y_t - b_9)^2}{\sigma_3^2} \right] \right\} \times \\ &(1 - \Phi(l_t')) \end{aligned} \quad (2-32)$$

where now

$$l_t = \frac{y_t - b_1 z_{1t} - b_2}{\sigma_1} \quad (2-33)$$

$$l_t' = \frac{y_t - b_3 p_t - b_5 z_{2t} - b_6}{\sigma_2} \quad (2-34)$$

The likelihood function (2-26) has the same form as before, with appropriate changes made in ψ_1, ψ_2, ℓ and ℓ' . Let S denote the set of n points in 3-space with coordinates (y_t, p_t, z_{2t}) , i.e., $S = \{(y_1, p_1, z_{21}), \dots, (y_n, p_n, z_{2n})\}$. Let \bar{S} be the convex hull of S . \bar{S} is bounded and there exists a supporting hyperplane defined by coefficients $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ such that $\alpha_1 \neq 0$ and

- (i) $\alpha_1 y_k + \alpha_2 p_k + \alpha_3 z_{2k} + \alpha_4 = 0$ for some k
- (ii) $\alpha_1 y_t + \alpha_2 p_t + \alpha_3 z_{2t} + \alpha_4 \leq 0$ for all t
- (iii) $\alpha_1 y_j + \alpha_2 p_j + \alpha_3 z_{2j} + \alpha_4 = 0$ for $j \neq k$ occurs with probability zero.

We then have $b_3 = -\alpha_2/\alpha_1$, $b_5 = -\alpha_3/\alpha_1$, $b_6 = -\alpha_4/\alpha_1$. In the present case all z_{2t} are positive; one example of a supporting hyperplane is given by $b_3 = 0$, $b_5 = \max \left(\frac{y_t}{z_{2t}} \right)$, $b_6 = 0$.

Now choose $\sigma_1 \neq 0$ and consider the behavior of the likelihood function as $\sigma_2 \rightarrow 0$. Since $\sigma_1 \neq 0$, ℓ_t is finite for all t and $1 - \Phi(\ell_t) \neq 0, 1$ for all t . Since σ_2 does not enter ψ_2 , $\psi_{2t} \neq 0$ for all t . As $\sigma_2 \rightarrow 0$ we observe that

- (i) $\psi_1(y_k, p_k) \rightarrow \infty$ since b_3, b_5, b_6 are so chosen that $(y_k - b_3 p_k - b_5 z_{2k} - b_6) = 0$;
- (ii) $1 - \Phi(\ell_k) = 1/2$
- (iii) $\ell_t \rightarrow -\infty$ and hence $1 - \Phi(\ell_t) \rightarrow 1$ for all $t \neq k$.

It follows that in

$$L = \prod_{t=1}^n \left\{ \psi_1(y_t, p_t) (1 - \Phi(\ell_t)) + \psi_2(y_t, p_t) (1 - \Phi(\ell'_t)) \right\}$$

the terms $\psi_2(y_t, p_t) (1 - \Phi(\ell'_t)) > 0$ for all t and $\psi_1(y_k, p_k) (1 - \Phi(\ell_k)) \rightarrow \infty$;
hence $L \rightarrow \infty$. QED.

The unboundedness of the likelihood function means that any attempt to find a global maximum will produce the inconsistent estimates $b_4 = 0$, σ_1 arbitrarily close to 0 . This difficulty is similar to the one that occurs in the mixture distribution models and can be avoided by bounding σ_2 away from zero. A convenient method of doing this in both cases is to require a priori knowledge of the ratio of the variances σ_1^2 , σ_2^2 and to set $\sigma_1^2 = k\sigma_2^2$. The computations reported in Section 3 will employ this restriction and will also examine the sensitivity of the results to alternative assumptions about k . It should be mentioned in conclusion that Proposition 2 is of a negative nature in that we have not proved the consistency of the ML estimates that employ the above restriction.

3. Empirical Results

The Suits Model. The model was originally estimated as follows. Equation (1-12) was estimated by ordinary least squares from data for the years 1919-1951. Equations (1-13) and (1-14) were estimated by limited information maximum likelihood ⁵ for the years 1930-1951 (since x_t was believed

⁵This estimation method consequently ignores the difficult points raised in [2] , [3] , [6] and [7] .

unavailable prior to 1930) with the harvest equation being estimated with only those data points for which $x_t < q_t$. These original results are reproduced in the first column of Table 1.

We extended the data series backwards by obtaining the previously missing observations for x_t . There is some question about comparability of the two sets of data, particularly with respect to the variable F . We first reestimated the equations using OLS for all three equations, but again estimating the harvest equation only from points for which $x_t < q_t$. These results are in the second column of Table 1. The constant terms and the coefficients of the dummy variables are very different; this however is simply due to scaling. The other parameter estimates are quite reasonably similar to those of Suits except for a_{12} which now has the wrong sign and appears to be statistically significant. We have no explanation for this phenomenon except the observation that our extension of the data series may have produced numbers not fully comparable with the earlier ones. We also find a higher income elasticity of crop: $\partial y_t / \partial Y_t = a_{10} / a_{11}$ and is 1.37 for Suits and 1.75 for the OLS estimates.

We next reestimated the model employing Specifications 1, 1-A and 2. The resulting estimates are in columns 3, 4 and 5 of Table 1. The results for Specification 1-A and 2 are quite similar to one another and are in general quite good. In particular, for Specification 2 the results for the first equation are identical to the OLS estimates, as was shown in Proposition 1. The reader may note that the standard errors for the first equation are slightly different; this is due to the fact that in the

Table 1

Results for the Suits Model*

<u>Coefficient</u>	<u>Original Suits Results</u>	<u>Recomputation Using OLS</u>	<u>Specification 1</u>	<u>Specification 1-A</u>	<u>Specification 2</u>
a ₁	.587(.156)	.584(.156)	1.622(1.378)	.613(.169)	.584(.141)
a ₂	-.320(.095)	-.354(.084)	-.744(.640)	-.350(.090)	-.354(.076)
a ₃	34.41(27.41)	.067(.051)	.104(.163)	.094(.055)	.067(.046)
a ₄	-.141(.238)	-.102(.205)	-.664(.736)	-.116(.215)	-.102(.185)
a ₅	-155.97(45.17)	-.361(.094)	-1.150(.955)	-.396(.105)	-.361(.085)
a ₆	768.735	4.509(.626)	.883(4.740)	4.355(.673)	4.509(.566)
a ₇	.237(.110)	.224(.095)	1.001(.450)	.279(.123)	.291(.112)
a ₈	1.205(.114)	1.135(.107)	.583(.563)	.822(.125)	.974(.083)
a ₉	-118.041	-1.178(.783)	1.293(3.895)	.831(.837)	-.166(.612)
a ₁₀	1.530(.088)	1.284(.291)	1.193(.278)	1.195(.284)	1.192(.285)
a ₁₁	-1.110(.246)	-.732(.501)	-1.500(-)	-1.481(.539)	-1.504(.528)
a ₁₂	-.682(.183)	2.152(.617)	2.435(.573)	2.429(.606)	2.437(.605)
a ₁₃	-140.163	9.651(.969)	10.742(.640)	10.715(.998)	10.748(.987)

* Numbers in parentheses are (ordinary or asymptotic)⁶ standard errors.

⁶The standard errors are derived from the negative inverse of the matrix of second partial derivatives of the loglikelihood function and differentiation was accomplished by numerical differencing.

ML method n rather than $n-k$ is employed as a divisor for computing the residual variance. With this adjustment the standard errors agree exactly.

In comparing Specification 1-A and 2 we note that the estimated standard errors are smaller for the second specification for every coefficient except one where there is a slight difference in the third place; a result completely in accord with expectations since Specification 2 uses more information. The coefficients agree with one another very well between the two specifications, except for a_9 which is a constant term and which is not significantly different from zero in either. Specification 2 suggests that the harvest is somewhat more responsive to price and less responsive to the size of the crop than either Suits or our OLS estimates have found. We now find a much lower income elasticity (.79) than before and also a substantially lower price elasticity ($1/a_{11}$), namely $-.66$ rather than $-.90$ for Suits and -1.37 for our OLS estimates. These changes appear reasonable although a_{12} has a reverse sign for these maximum likelihood estimates as well.

We have had considerable difficulty computing estimates by Specification 1 even though the variances were constrained to their Specification 2 ratio. The likelihood function appeared to be very flat over extensive ranges of the parameter space. We did compute suboptimal estimates by fixing the value of a_{11} at -1.5 , approximately the estimated value by Specification 2. Although the signs of the remaining coefficients are the same as in Specification 2 (again with the exception of the constant term a_9), their values are not plausible and none of the coefficients in the crop and harvest equations is significant. The poor performance of Specification 1

may well result from the simultaneous presence of specification error, inadequate data, and excessive denial of information to the model. In the light of these results, it appeared desirable to conduct some Monte Carlo experiments in order to shed some additional light on the computability of estimates, their sampling properties and on the effect of using differing amounts of information in a controlled situation in which the truth is known to the investigator.

Monte Carlo Experiments. The condensed form of the Suits model as given by (2-4) to (2-7) was employed for an empirical examination of estimation under our two specifications of the likelihood function. The exogenous variables z_1 , z_2 , z_3 were chosen to be the same as the variables F , W , and X in the actual Suits model. We experimented with two sample sizes $N = 30$ and 60 . Since the actual data comprise 33 observations, we omitted the last three for the smaller sample size and for the larger sample size we used all 33 and then repeated the first 27. In the basic experiments the true values of the parameters were $b_1 = 1.0$, $b_2 = 0.3$, $b_3 = 2.0$, $b_4 = 1.0$, $b_5 = -1.5$, $b_6 = -4.5$, $b_7 = 1.1$, $b_8 = -1.3$, $b_9 = 5.9$, $\sigma_1^2 = 0.0132$, $\sigma_2^2 = 0.0025$, $\sigma_3^2 = 0.1277$ where the latter three figures were those computed for the actual Suits model by OLS. The parameter values were chosen so as to reproduce approximately the levels of the dependent variables observed in the actual data. Observations on the endogenous variables were generated from Equations (2-4) to (2-7) using error terms distributed normally and independently with the indicated variances. Each experiment consisted of 50 replications of the process of generating a set of data and estimating

the parameters by the two basic specifications; since Specification 1-A is intermediate between these, we performed no large-scale experiments with it. In Specification 1 we have to constrain σ_2^2 and we assumed that $\sigma_2^2 = \frac{1}{k}\sigma_1^2$, where we chose for k in the main set of experiments the true value of σ_1^2/σ_2^2 .

In the principal set of experiments we varied the sample size N , the coefficient b_4 and the variance σ_1^2 . Since Specification 1 uses less information than Specification 2, in that it employs no explicit q -data, we wished to vary b_4 , the coefficient of q in Equation (2-5) and the variance of the error term in Equation (2-4), in order to examine the effect of these variations on the value of additional information. Of course, changing the parameters of the model may well change the fraction of sample points over the replication for which $x_t < q_t$ and the fraction for which $x_t = q_t$. We wished to keep this fraction as nearly constant between experiments as possible since these fractions determine in effect how many observations "belong to each regime" and variation in this fraction has well-known effects on root mean-square errors of estimates [8]. In order to keep these fractions nearly constant, whenever we varied b_4 for an experiment we also undertook a "compensating" variation in b_6 . The basic characteristic of the experiment are described in Table 2. It will be noted that in the six basic experiments the mean fraction of observations over each set of 50 replications is quite stable, ranging from .65 to .69; slightly larger changes are observable for the minimum and maximum fractions over each set of 50 replications.

Table 2

Characteristics of Experiments

Case	N	b_4	b_6	σ_1^2	Mean fraction of observations with $x_t < q_t$	Maximum fraction of observations with $x_t < q_t$	Minimum fraction of observations with $x_t < q_t$
1	60	1.0	-4.5	0.0132	.67	.77	.50
2	30	1.0	-4.5	0.0132	.67	.80	.53
3	60	1.0	-4.5	0.0033	.66	.73	.52
4	60	0.2	0.9	0.0132	.66	.78	.53
5	60	1.7	-9.3	0.0132	.69	.82	.57
6	60	1.0	-4.5	0.0528	.65	.75	.52

Basic Results of Experiments. The overall results are very good and numerical results are displayed in Tables 3 through 7. Not one single computational failure occurred in any of the replications of the experiments. The mean biases of the estimates for Cases 1 and 2 are displayed in Table 3 and for $N=60$ they are very small relative to the true values of the coefficients for both specifications, with the exception of one constant term. The median value of the ratio of bias to true value is .02 for Specification 1 and .01 for Specification 2. In general, looking at either mean biases or at root mean square errors, displayed in Table 4, we find that both specifications estimate the coefficients with a high degree of precision with the largest errors occurring invariably for the constant terms in the equations.

There are several sets of comparisons that are of particular interest for shedding light on the value of information. These are (1) a comparison between Specifications 1 and 2 for any of the Cases; (2) a comparison between

Cases 1 and 2 for examining the effects of variations in N ; (3) a comparison among Cases 1, 4 and 5 for variations in b_4 and (4) a comparison among Cases 1, 3 and 6 for variations in σ_1^2 .

(1) Comparison between Specifications 1 and 2. In Table 5 we display the ratios of the RMSE's for Specification 2 to the corresponding RMSE's for Specification 1. A ratio less than 1.0 indicates that using q -data explicitly has positive value. Of the 66 ratios in Table 5 only 4 are larger than unity; the largest of those is 1.013. The median of all the ratios is .60; disregarding the ratios for the price equation, (2-6) , which account for all the ratios in excess of .905 the median is .58. On the average in this sense having the q -data thus produces a 40 percent improvement. It is not particularly surprising that the improvement is minor for equation (2-6) which does not contain q ; for (2-6) the median of the corresponding ratios is .986. From the estimates for each replication we can compute for each sample point the estimated probability that that sample point belongs to the regime $x_t < q_t$ (or $v_{1t} < v_{2t}$) . These probabilities are computed quite accurately and the mean bias in the probability estimates does not exceed one percent in any Case. We can then compute the RMSE over the replications of these probabilities for each sample point. The average of these N RMSE's is displayed in Table 6. As is to be expected, the RMSE's for Specification 2 are uniformly lower, their ratios ranging from .746 to .911, thus further confirming the value of q -data.

(2) Variations in N . The mean biases decline uniformly for Specification

1 as we go from $N=30$ to $N=60$; they decline for a majority of the coefficients for Specification 2. Even more convincing is the behavior of the root mean square errors (RMSE) which are uniformly smaller for $N=60$ than for $N=30$ for both specifications. The median of the ratios of the RMSE's is 1.39, a not unreasonable estimate of $\sqrt{2}$, i.e. square root of the ratio of the sample sizes. In Table 5 a decline in a ratio from one case to another indicates an increase in the value of information, i.e., it indicates that explicit use of the q -data has caused the RMSE to decline relative to Specification 1 which does not use q -data. Variations in N are ambiguous in this regard, as larger N increases the relative value of q for estimating equation (2-5), reduces the relative value of q for estimating equation (2-4) and has no appreciable effect on equation (2-6).

Table 6 shows the same marked decline in the RMSE's for Case 1 as does Table 4; the ratios average to 1.47, again reasonably close to $\sqrt{2}$.

(3) Variations in b_4 . If b_4 were equal to zero and since $x_t \leq q_t$ in any event, the absence of information on q_t would be expected to have an effect on the efficiency with which (2-5) and (2-6) can be estimated only to the extent that we are unable to classify the sample points between regimes.⁷ If $b_4 \neq 0$ we have the additional complication of the unobservable q_t entering (2-5); the larger the absolute value of b_4 , the more valuable would appear knowledge of the q -data for estimating (2-5) and obviously also for estimating (2-4). The relevant cases are ordered 4, 1 and 5 in order

⁷If no changes of regimes were present, the model would be formally analogous to Goldberger's Case 4 [4].

of increasing b_4 . The RMSE's for Specification 1 for equations (2-4) and (2-5) uniformly increase with increasing b_4 . Equation (2-6) shows relatively little change in the RMSE, though whatever change there is, is in the opposite direction. For equations (2-4) and (2-5) the increase in the RMSE's from Case 4 to Case 5 ranges from a factor of 1.08 for equation (2-4) to factors up to 6.20 for equation (2-5). The Table 5 ratios show unambiguous results only for equation (2-5); for that equation it is clearly true that the superiority of Specification 2 estimates increases with the value of b_4 . The same unambiguous picture emerges from the Table 6 RMSE's.

(4) Variations in σ_1^2 . The relevant Cases are ordered 3, 1 and 6 with increasing values of σ_1^2 . As should be expected, the RMSE's for equation (2-4) by Specification 2 change in proportion to the proportionate change in the square root of the residual variance. For Specification 1 the same RMSE's increase somewhat less than proportionately. The greater is σ_1^2 , the more valuable it ought to be for the estimation of (2-5) to know the value of q as is indicated by the uniform and marked decline of the Table 5 ratios for the coefficients b_4 , b_5 , b_6 . Again, no particular effects are observable for equation (2-6) and the Table 6 RMSE's further support our conclusion of the relative deterioration of Specification 2 estimates as σ_1^2 increases.

Additional Results. Several other aspects of the experiments help to characterize the success of the two estimating methods. These aspects are (1) the ratio of the average asymptotic standard deviation to the RMSE; (2) the sensitivity of the experiments to the ratio σ_1^2/σ_2^2 assumed to be true for estimating purposes, and (3) the distribution of the estimates.

(1) Ratio of the average asymptotic standard deviation to the RMSE. If

the estimated parameters are jointly sufficient statistics then the diagonal elements of the negative inverse of the matrix of second partial derivatives of the loglikelihood function are consistent estimates of the asymptotic variances. Hence under these circumstances the average over the replication of their square roots divided by the corresponding RMSE converges to 1 as $N \rightarrow \infty$. These ratios are displayed in Table 7. Although there is no rigorous test of when the ratios are "close enough" to unity, the results appear satisfactory on the whole. The ratios pertaining to equation (2-6) again tend to show no systematic behavior although in comparing Case 1 with Case 2 the ratios behave perversely. Otherwise increasing sample size causes the ratios to become very much closer to unity. The ratios for Specification 1 tend to deteriorate as we move from Case 4 to 1 to 5 and also as we move from 3 to 1 to 6, suggesting in conformity with previously observed behavior that as the characteristics of the experiment change in the direction of an increased role of q_t in the model, the quality of the estimates by the method that does not make use of q deteriorates.

(2) Variations in the assumed σ_1^2/σ_2^2 ratio. In the six principal experiments we had employed for estimating purposes the constraint $k\sigma_2^2 = \sigma_1^2$ where the value of k was set at the true value of σ_1^2/σ_2^2 . Since this is prior information that needs to be supplied by the investigator we examined the sensitivity of the results to inaccurate assumptions about k . In fact, we repeated Case 1 (for Specification 1) assuming alternately a value of k equal to twice and to one half the true value. There was a slight increase on the average in the RMSE's, but no increase exceeded 6 percent and some actually

declined. The mean biases changed relatively little when k was one half its true value. Although they increased quite substantially in some instances for k double its true value (up to a factor of 10), they still resulted in an acceptable median ratio of mean bias to true value of coefficient of .04. Thus the estimation process seems relatively insensitive to the actual k -value assumed.

(3) Distribution of the Estimates. Since maximum likelihood estimates are, under general conditions, asymptotically normally distributed, we investigated whether in our finite sample the normal distribution provides an adequate description of the estimates. For each experiment and each estimated coefficient we tested the hypothesis that the quantity $(\text{est. coeff} - \text{true coeff})/\text{est. RMSE}$ is distributed as $N(0,1)$. The sampling distributions were compared to the theoretical one by the Kolmogorov-Smirnov test.

Out of the 138 estimated coefficients (6 Cases \times (11 + 12)) only 13 exceeded the critical value of the Kolmogorov-Smirnov statistic at the .10 level. The median value of this statistic was .0989 for Specification 1 coefficient estimates and .1088 for Specification 2. We conclude in general that we cannot reject the hypothesis that the estimates are normally distributed. It is also noteworthy that massive improvement occurs in the fit of $N(0,1)$ as we go from Case 2 ($N = 30$) to Case 1 ($N=60$) and that nine of the 13 rejections of the null hypothesis pertain to estimates of residual variances.

4. Conclusions

We have succeeded in formulating the stochastic specification of a disequilibrium model in several alternative ways and in computing maximum likelihood estimates for the alternatives, both in a realistic economic example and in some Monte Carlo experiments. These computations establish the computability of models of this type and shed considerable light on both the finite sampling properties of the estimates and on the value of information as measured by the impact upon these properties of using more or less information or data. Several open questions remain, of course, such as (1) are the maximum likelihood estimates proposed here consistent; (2) what are appropriate ways of testing hypotheses about the existence of disequilibrium; (3) are unbounded likelihood functions likely to occur frequently in practice in models of this type; and (4) what properties do estimates have from likelihood functions that are theoretically unbounded but for which an interior maximum also exists.

Table 3

Mean Biases for Cases 1 and 2

<u>Coefficient</u>	<u>Specification 1</u>		<u>Specification 2</u>	
	<u>Case 1</u>	<u>Case 2</u>	<u>Case 1</u>	<u>Case 2</u>
b_1	.0029	.0568	-.0233	-.0188
b_2	-.0141	-.3491	.1542	.1282
b_3	.0461	.0601	.0010	-.0116
b_4	-.0051	-.0578	.0055	-.0015
b_5	-.0420	-.0465	.0004	.0098
b_6	-.0311	.2720	-.0446	.0314
b_7	.0121	.0431	.0091	.0382
b_8	-.0288	-.0824	-.0197	-.0753
b_9	.1043	.2322	.0663	.2203
σ_3^2	-.0039	-.0064	-.0049	-.0073
σ_1^2	-.0005	-.0010	.0000	-.0006

Table 4

Root Mean Square Errors
Specification 1

<u>Coefficient</u>	<u>Case 1</u>	<u>Case 2</u>	<u>Case 3</u>	<u>Case 4</u>	<u>Case 5</u>	<u>Case 6</u>
b ₁	.1318	.2775	.0752	.1310	.1411	.2002
b ₂	.8409	1.7179	.4806	.8354	.9107	1.2884
b ₃	.2317	.3326	.1452	.1120	.3554	.4675
b ₄	.1596	.2205	.1028	.0672	.3246	.2239
b ₅	.1664	.2563	.0966	.0732	.2607	.3555
b ₆	1.2624	1.7687	.7784	.4142	2.5668	2.0549
b ₇	.2140	.2952	.2151	.2167	.2001	.2090
b ₈	.3168	.4172	.3401	.3283	.2752	.2462
b ₉	1.8878	2.5232	1.9602	1.9394	1.7309	1.6530
σ_3^2	.0323	.0420	.0335	.0330	.0312	.0281
σ_1^2	.0026	.0059	.0006	.0026	.0031	.0136

Specification 2

<u>Coefficient</u>	<u>Case 1</u>	<u>Case 2</u>	<u>Case 3</u>	<u>Case 4</u>	<u>Case 5</u>	<u>Case 6</u>
b ₁	.0867	.1202	.0434	.0867	.0874	.1734
b ₂	.5628	.7820	.2814	.5628	.5674	1.1256
b ₃	.0920	.1471	.0907	.0897	.0890	.0849
b ₄	.0419	.0677	.0484	.0413	.0558	.0335
b ₅	.0582	.0977	.0585	.0575	.0585	.0545
b ₆	.4464	.7077	.4502	.3011	.5813	.4110
b ₇	.2114	.2851	.2080	.2149	.1948	.2061
b ₈	.3086	.4031	.3254	.3233	.2675	.2357
b ₉	1.8676	2.5091	1.9700	1.9410	1.7445	1.5862
σ_3^2	.0309	.0425	.0319	.0327	.0300	.0268
σ_1^2	.0023	.0032	.0006	.0023	.0024	.0091
σ_2^2	.0006	.0010	.0007	.0007	.0006	.0007

Table 5

Ratios of RMSE's: (Specification 2 ÷ Specification 1)

<u>Coefficient</u>	<u>Case 1</u>	<u>Case 2</u>	<u>Case 3</u>	<u>Case 4</u>	<u>Case 5</u>	<u>Case 6</u>
b_1	.658	.433	.577	.662	.619	.866
b_2	.669	.455	.586	.674	.623	.874
b_3	.391	.442	.625	.801	.250	.182
b_4	.263	.307	.471	.615	.172	.150
b_5	.350	.381	.606	.786	.224	.153
b_6	.354	.400	.578	.727	.226	.200
b_7	.988	.966	.967	.992	.974	.986
b_8	.974	.966	.957	.985	.972	.957
b_9	.989	.994	1.005	1.001	1.008	.959
σ_3^2	.959	1.013	.952	.990	.961	.952
σ_1^2	.874	.543	.905	.877	.751	.668

Table 6

Average RMSE for $P\{v_1 < v_2\}$

	<u>Case 1</u>	<u>Case 2</u>	<u>Case 3</u>	<u>Case 4</u>	<u>Case 5</u>	<u>Case 6</u>
Specification 1	.0695	.1127	.0665	.0680	.0747	.0747
Specification 2	.0628	.0840	.0631	.0620	.0622	.0610

Table 7

Ratio of the Average over Replications of the Asymptotic
Standard Deviation to the Root Mean Square Error

<u>Coefficient</u>	<u>Specification 1</u>					
	<u>Case 1</u>	<u>Case 2</u>	<u>Case 3</u>	<u>Case 4</u>	<u>Case 5</u>	<u>Case 6</u>
b ₁	.9810	.7868	1.0023	1.0204	.9207	.9986
b ₂	.9838	.8078	1.0049	1.0236	.9162	.9965
b ₃	.8635	.8146	.9191	.8268	.8948	.7005
b ₄	.8909	.8820	1.0419	1.0058	.8134	.6641
b ₅	.8926	.7811	1.0183	.9214	.9073	.6864
b ₆	.9175	.8876	1.0560	1.0338	.8839	.7257
b ₇	.7698	.8014	.7732	.7566	.8279	.7691
b ₈	.7881	.8706	.7699	.8042	.8584	.8809
b ₉	.8177	.8768	.8037	.8439	.8347	.8746

<u>Coefficient</u>	<u>Specification 2</u>					
	<u>Case 1</u>	<u>Case 2</u>	<u>Case 3</u>	<u>Case 4</u>	<u>Case 5</u>	<u>Case 6</u>
b ₁	1.0006	1.0033	1.0006	1.0006	.9915	1.0005
b ₂	.9982	.9982	.9982	.9982	.9989	.9982
b ₃	.9043	.7633	.9162	.8883	.8989	.9399
b ₄	.9942	.8110	.9389	1.0712	.9135	.9952
b ₅	1.0460	.8438	1.0356	1.0072	1.0045	1.0743
b ₆	.8569	.7161	.8696	.9773	.8452	.8607
b ₇	.7729	.8207	.7938	.7603	.8364	.7680
b ₈	.7887	.8763	.7843	.8032	.8420	.8774
b ₉	.8144	.8687	.7913	.8356	.8089	.8874

REFERENCES

- [1]. Amemiya, T., "A Note on the Fair-Jaffee Model," Technical Report No. 112, Institute for Mathematical Studies in the Social Sciences, Stanford University, October 1973.
- [2]. Fair, R.C. and D.M. Jaffee, "Methods of Estimation for Markets in Disequilibrium," Econometrica, 40(1972), 497-514.
- [3]. Fair, R.C. and H.H. Kelejian, "Methods for Estimation for Markets in Disequilibrium: A Further Study," Econometric Research Program, Research Memo, No. 135, Princeton University, January 1972.
- [4]. Goldberger, A.S., "Maximum Likelihood Estimation of Regressions Containing Unobservable Independent Variables," International Economic Review, 13(1972), 1-15.
- [5]. Grossman, Herschel I., "The Nature of Quantities in Market Disequilibrium," American Economic Review, LXIV(1974), 509-514.
- [6]. Hartley, M.J., "On the Estimation of Markets in Disequilibrium," Discussion Paper No. 275, Economic Research Group, State University of New York at Buffalo, February 1974.
- [7]. Maddala, G.S. and F.D. Nelson, "Maximum Likelihood Methods of Markets in Disequilibrium," Econometrica, forthcoming.
- [8]. Quandt, R.E., "A New Approach to Estimating Switching Regressions," Journal of the American Statistical Association, 67(1972), 306-310.
- [9]. Suits, D., "An Econometric Model of the Watermelon Market," Journal of Farm Economics, 37(1955), 237-251.