SMALL SAMPLE PROPERTIES OF DEMAND SYSTEM ESTIMATES

by

Nicholas M. Kiefer
James G. MacKinnon

Econometric Research Program
Princeton University
Research Memorandum No. #179

June 1975

SMALL SAMPLE PROPERTIES OF DEMAND SYSTEM ESTIMATES*

by

Nicholas M. Kiefer
Princeton University

James G. MacKinnon
Queen's University

April 1975

1. Introduction

Although the estimation of systems of demand equations is one of the major topics in applied econometrics, little is known about the small sample properties of such estimates. This study makes use of sampling experiments in an effort to remedy that deficiency.

It is probably fair to say that the empirical results in the estimation of demand systems have not been entirely satisfactory. Parameter and elasticity estimates have often tended to be highly dependent on the model specified,[1] and in the experience of the authors they have not always been plausible; see Kiefer [1975] and MacKinnon [1975]. There would seen to be two possible explanations for this state of affairs. The first is that the demand systems which have been estimated simply do not describe the data adequately, either because the functional forms are too restrictive, or because aggregation across individuals, simultaneity with the supply side of the economy, durable goods and other complications have been ignored. This will be referred to as the structuralist explanation. An alternative explanation, no less plausible a priori, is that the functional forms which have been estimated characterize the true demand relationships well enough, but that, given the reasonably modest number of observations and fairly small relative price variation in the data series that are typically employed, the stochastic component overwhelms the systematic ones and causes the estimates to be imprecise. This will be referred to as the stochastic explanation.

One of the major purposes of this study is to assess the relative importance of the structuralist and stochastic explanations. Suppose that data were generated by some demand system. If estimates of that system using that data tended to be good, while estimates of other systems using the same data tended to be poor, considerable doubt would be cast upon the stochastic explanation and the structuralist explanation would be reinforced. If, on the other hand, estimates of the true system tended to be poor, the structuralist argument would be considerably weakened.

There are two main competing approaches to the specification of demand systems. One approach is to postulate a restrictive form for the utility function and estimate the demand equations which would result if a single individual maximized that function subject to a budget constraint. The systems of this type which are most commonly employed are the linear expenditure system and its generalizations. The alternative approach is to postulate a demand system which can provide a local approximation to an arbitrary system of demand equations that satisfy the usual conditions. The translog and Rotterdam models are commonly used systems of this type.

In our experiments systems of both types have been employed. Specifically, the true model was postulated to be the linear expenditure system in roughly half our experiments, and the translog system in the other half. These systems were chosen because they are widely used and reasonably simple to estimate. Since the linear expenditure and translog systems are very different in functional form and represent radically different approaches, it seems reasonable to believe that any results which hold for both those systems are likely to hold for most of the demand systems which are commonly estimated. In addition, we were able to investigate the effects of misspecification by estimating one system from data generated by the other.

2. The Linear Expenditure System

The linear expenditure system (LES) is the most venerable and widely estimated system of demand equations. It was first used by Stone [1954], and has more recently been estimated by Parks [1969] and Abbott and Ashenfelter [1974], among others; see Brown and Deaton [1972] for more references. In recent years a number of generalizations of the linear expenditure system have been proposed. One of the simplest is what Wales [1971] calls the generalized linear expenditure system (GLES). The share of expenditure devoted to the i^{th} good in the GLES is

$$w_i = \gamma_i P_i/Y + [q_i P_i^{-\sigma} / \sum_{j=1}^{n} q_j P_j^{1-\sigma}] P_i (Y - \sum_{j=1}^{n} P_i \gamma_i)/Y \ , \qquad (2\text{-}1)$$

where Y is total income and P_i is the price of good i . The parameters of the system are the so-called share parameters, q_1 to q_n , the committed quantities, γ_1 to γ_n , and σ , which is the elasticity of substitution between uncommitted expenditures. This way of writing the GLES is due to MacKinnon [1975].

When σ is 1, the GLES reduces to the LES. The LES can be obtained directly by maximizing the Stone-Geary utility function

$$U(x) = \sum_{i=1}^{n} q_i \log(x_i - \gamma_i) \qquad (2\text{-}2)$$

subject to the constraint that total expenditures be equal to Y . Since (2-1) is homogeneous of degree zero in the q_i's , only $n-1$ of them can be estimated. One possibility is to prespecify one of the q_i's , perhaps by setting q_1 equal to one, for example. If q_i and q_1 differ by a large order of magnitude, this can make estimation difficult. An alternative approach is to set q_1 equal to one minus the sum of the other q_i's , so that the q_i's sum to unity. With this normalization, the share of expenditure devoted to the ith good in the LES can be written as

$$w_i = \gamma_i P_i/Y + q_i(Y-\sum_{j=1}^{n} P_i\gamma_i)/Y \ . \qquad (2\text{-}3)$$

The LES requires the estimation of 2n-1 parameters - n committed quantities and n-1 share parameters. The GLES in addition requires the estimation of σ , for a total of 2n parameters.

Provided that prices and income are positive, (2-1) and (2-3) are always defined, but they do not always make economic sense. If q_i is negative for any good, (2-2) implies that the marginal utility of that good is negative, and the utility function which generates the GLES is not defined. If the value of the committed quantities, $\sum_{i=1}^{n} \gamma_i P_i$, exceeds income, (2-2) and the GLES analogues are not defined. Thus in this case it makes no sense to assert that these systems arise because people maximize utility. If σ in the GLES is negative, the ratios of the purchases of any two goods increases as their price ratio increases. Hence the LES and GLES only make sense if σ and every q_i are positive, and none of the γ_i's are too big. The 'true' parameters chosen for our experiments satisfy these conditions, but the estimated parameters need not always do so. If they often failed to satisfy these minimal conditions of economic reasonableness, it would clearly provide strong evidence in favour of the stochastic explanation.

3. The Translog System

The transcendental logarithmic indirect utility function, which is due to Jorgenson, Christensen and Lau [1973], is

$$V = \alpha \log(P/Y) + (\log(P/Y))'(\beta/2)(\log(P/Y)) \qquad (3\text{-}1)$$

constants. Hence the observable implications of $g(V)$ are identical to those of V itself. Therefore, one element of α and one of β must be restricted a priori. The normalizations chosen are

$$\sum_{i=1}^{n} \alpha_i = -1$$

$$\sum_{i=1}^{n} \beta_{ii} = 1 \ .$$

These restrictions allow V to be a decreasing and convex function of prices, as expected, but do not require it. Since β in the utility function (3-1) is the matrix of a quadratic form, it may be taken to be symmetric. Thus the number of elements of β that must be estimated is $n(n+1)/2 - 1$. In addition, $n-1$ elements of α must be estimated, for a total of $n(n+1)/2-2+n$.

4. Design of the Experiments

Every attempt was made to make the experiments we conducted as close as possible to the situation actually faced by researchers attempting to estimate demand systems from annual time series data.

In order to limit the costs of computation, the number of goods was chosen to be three. Three is the smallest number that requires the estimation of more than one equation, and the smallest number that allows for complementarity. For this case the LES has five parameters which must be estimated, the GLES has six, and the translog has seven.

The number of observations was chosen to be forty. This seems to be typical of the data series available to researchers estimating demand systems.[2] We did not experiment with different numbers of observations because to do so would have significantly increased the cost of computation; the behavior of the estimates with changing sample size is of only peripheral interest; and in practice more than forty observations are rarely available.

Every experiment involved fifty replications. This seemed to us a reasonable compromise between cost and sampling error, and there were no interesting results which were ever in doubt because of an inability to make clearcut inferences.

In any sampling experiment, the choice of the independent variables is critical. A natural way to choose them would be to assume that they come from some given probability distribution, such as the normal or the uniform. But this is clearly not how income and price series are generated. In fact, income and all prices tend to rise sharply over time, but at different rates, so that most of the variation in relative prices takes place only gradually. Examination of some actual income and price series (those used by MacKinnon [1975]) suggested that, outside the war years, they can be well characterized by random processes of the following sort:

$$p_t^i = p_{t-1}^i (g^i + e_t^i + a^i v_t) \qquad (4\text{-}1)$$

where g^i is one plus a rate of growth, e_t^i is a normally distributed error term specific to series i, v_t is a normally distributed error term common to all the series, and a^i is a constant. Thus (4-1) implies that each series (income as well as price) tends on the average to grow at a constant rate (higher for income than for prices, and varying among the latter), but this is modified by random shocks. Some of these random shocks affect individual series only, while some affect all the series to some degree (to reflect the role of the business cycle).

Plausible parameters were chosen for (4-1) and three price and one income series were derived in this way (see the Appendix). These were then normalized to equal unity at the twenty-first observation. This normalization is equivalent to choosing a base year for the price indices and redefining the units in which real quantities are measured. It was carried out in order that the translog system may be interpreted as a Taylor series approximation to an arbitrary system at observation 21 (which is the midpoint of the sample when the first observation is dropped to allow for an autoregressive error structure). The same series were used in all the experiments.

The estimating equations from both models are

$$w_t^i = \hat{w}_t^i + u_t^i \quad i = 1,\dots,n \quad t = 1,\dots,T . \tag{4-2}$$

For each t the budget constraint implies that both actual and fitted shares sum to unity, so that any one share can be determined from the other $n-1$. This also implies that the errors in each period sum to zero, so that the variance-covariance matrix of u is singular. Barten [1969] has shown that as a result only $n-1$ equations need be estimated, and that the parameter estimates will be invariant with respect to the equation which is dropped. In our experiments we dropped the third.

We assume that the u_t^i's follow a first-order autoregressive scheme,

$$u_t = \rho u_{t-1} + \varepsilon_t , \tag{4-3}$$

where u_t is an $(n-1)$-vector whose i^{th} element is u_t^i. ε_t is assumed to come from an $(n-1)$-variate normal distribution with mean vector 0 and variance-covariance matrix $\sum$, constant through time. Berndt and Savin [1975] have extended Barten's analysis to the autoregressive case, and they have shown that ρ must be the same for each equation if the estimates are to be invariant with respect to which equation is dropped.

Our experience with demand system estimates suggests that ρ is likely to be fairly close to one in practice. Therefore in most of our experiments we have chosen ρ to be .9 or .98, although for purposes of comparison we have also experimented with $\rho = 0$.

Since the variance-covariance matrix of all n ε^i's must be singular, the off-diagonal elements cannot all be zero, and it would be a remarkable coincidence if the off-diagonal elements of Σ were all zero. Bearing this in mind, Σ was chosen to be

$$\begin{pmatrix} .000036 & -.000025 \\ -.000025 & .000049 \end{pmatrix} . \qquad (4\text{-}4)$$

This implies that the full (singular) variance-covariance matrix is

$$\begin{pmatrix} .000036 & -.000025 & -.000011 \\ -.000025 & .000049 & -.000024 \\ -.000011 & -.000024 & .000035 \end{pmatrix} .$$

Martix (4-4) was used in all our experiments, regardless of ρ. Since the variance of u is $1/(1-\rho^2)$ times the variance of ε, this means that the random component is larger when ρ is larger, so that one would expect the precision of the estimates to decline. It is not clear, however, that any other choice would be preferable. For example, Σ could have been chosen differently for each ρ so as to keep the variance of u constant. But since the error terms can be predicted better when ρ is larger, one would expect the precision of the estimates to increase with ρ if that were done. The fact that Σ is the same in all the experiments should be kept in mind when interpreting the results.

The purpose of this study is not to compare different estimating techniques. Full information maximum likelihood was used exclusively (see MacKinnon [1975] for the methodology). Since it is known that Σ is not diagonal, limited information methods are clearly inappropriate (and since all parameters appear in all equations, it is not clear that they would be as easily obtained as FIML).

Several ad hoc techniques which are commonly used to estimate demand systems are also not particularly easy to use, and derive their only theoretical justification from being asymptotically equivalent to FIML. Analytical small sample properties of simultaneous equation estimators in the presence of nonlinearities are virtually unknown, so there is no estimator which is clearly best. However, FIML takes account of all restrictions simultaneously and estimates all equations simultaneously, which cannot be said for many other techniques. Of course, its large sample properties are unexcelled, and what evidence there is from sampling experiments seems to indicate that FIML is superior to nonlinear two stage least squares and to various polynomial approximation estimators; see Quandt [1975]. For these reasons, it seems to us very unlikely that any alternative estimating technique would give results noticeably better than FIML. Using FIML therefore gives a reasonable picture of the best results one can expect from demand system estimates.

In some of our experiments the 'true' shares were generated by the LES, in others by the translog. We wished to make the data in all the experiments reasonably comparable, so it was not possible to choose the parameters of the LES and the translog independently of each other. What we did was to choose parameters for the LES which generated share series that seemed reasonable. The translog was then fitted to these shares in order to choose parameters for it. Thus with the parameters that were used both models generate share series that look reasonably similar, so that we are comparing the performance of the two models in similar situations.

5. Structure of the Experiments

One of the main purposes of this investigation was to examine the effects of mis-specifying the error structure. As indicated in the previous section, we assumed that the errors were generated by a first-order autoregressive process, (4-3), with a single parameter ρ common to all equations. In order to obtain the best asymptotic properties, ρ should be estimated by maximum likelihood along with the other parameters. It is not clear that this is the best procedure in small samples, however. Most authors who have estimated demand systems have assumed that ρ is zero; more recently some authors, noting the strong evidence of autocorrelation in the data, have assumed that ρ is one.[3] Both these approaches have the advantage of requiring the estimation of one less parameter than the asymptotically correct procedure, and in addition the former (estimating in levels) has the advantage that one more observation is made available, since the first observation does not have to be dropped.

Three plausible estimating techniques are therefore available: setting $\hat{\rho} = 0$ (estimating in levels), estimating $\hat{\rho}$, and setting $\hat{\rho} = 1$ (estimating the first differences). We examined the performance of these techniques when $\rho = .9$ and when $\rho = .98$, and in addition, for purposes of comparison, tried estimating in levels when $\rho = 0$. A schematic representation of the experiments that were performed is shown in Table 1. There are nine combinations of true values of ρ and assumptions about ρ. Each 'box' in Table 1 shows what experiments were performed with the indicated true value and assumption. For example, the box in the upper left-hand corner shows that three experiments were performed with a true ρ of zero and an assumed ρ of zero: the LES and GLES were estimated when the true system was LES, and the translog was estimated when the true system was translog. Not all possible experiments were performed. It seemed to us unlikely that a true ρ of zero would ever be encountered, and even more unlikely that the econometrician would allow for the possibility of its being non-zero if in fact it were zero. Thus two of the boxes in the top row of the table are empty. The case that seemed most plausible to us was $\rho = .9$, and for that reason the bulk of the experiments were conducted for that case.

Table 1

<u>Structure of the Experiments</u>

Assumed ρ: / True ρ	0	Estimated	1
0.0	LES-LES LES-GLES Tlog-Tlog		
0.9	LES-LES LES-GLES LES-Tlog Tlog-Tlog Tlog-LES Tlog-GLES	LES-LES LES-GLES LES-Tlog Tlog-Tlog Tlog-LES Tlog_GLES	LES-LES LES-GLES LES-Tlog Tlog-Tlog Tlog-LES Tlog-GLES
0.98	LES-LES Tlog-Tlog	LES-LES Tlog-Tlog	LES-LES Tlog-Tlog

Of each pair, the system on the left of the hyphen is the one that actually generated the data, and the system on the right is the one that was estimated.

A total of thirty experiments were performed. Since each experiment involved fifty replications, a total of fifteen hundred nonlinear likelihood functions had to be maximized. The optimization algorithm of Davidon, Fletcher and Powell (See Goldfeld and Quandt [1972] and MacKinnon [1975]) was used in almost all cases. In order to minimize the computational burden, the algorithm was always started at the true values of the parameters. The computer program for the algorithm terminates if either the step-size becomes negligible, or the gradient becomes negligible, or the change in the function value becomes negligible. A termination was accepted as a maximum if the matrix of second derivatives of the log-likelihood function was negative definite. No attempt was made to check for multiple maxima, although the experience reported in Chapter suggests that they may occur for the GLES. Any attempt to deal with this problem would have been prohibitively expensive.

It is customary in reporting the results of sampling experiments to focus almost exclusively on parameter estimates. This would be inadequate in the present instance, for at least two reasons. First of all, the translog and LES systems do not involve the same parameters, so that looking at parameter estimates would tell us little about the performance of the translog when the truth is LES, and vice versa. Secondly, it is quite plausible that in nonlinear systems certain parameters may be estimated very imprecisely, while the economically interesting magnitudes may be estimated with much greater precision. As an example, one would expect the parameters of

$$y = a_1x + a_2x^2 + a_3x^3 + a_4x^4 + a_5x^5 + u$$

to be estimated very imprecisely indeed with most time series data, but that need not imply that estimates of dy/dx are equally imprecise. For two reasons, then, one would like to examine the estimates of economically interesting magnitudes, as well as of parameter estimates.

The economically interesting magnitudes in this context are primarily estimates of elasticities and expenditure shares. Demand equations and systems of demand equations are often estimated simply to derive elasticity estimates, and economists may be more apt to believe these estimates than the systems that were used to derive them. It is therefore of considerable interest to see just how good the elasticity estimates are in our experiments. If estimated demand systems are to be used in econometric models, they must not only provide reliable elasticity estimates, but must also provide reliable estimates of expenditure shares. A system which estimates elasticities perfectly would nevertheless be virtually useless for that purpose if its share estimates were poor.

For these reasons, then, a considerable part of the analysis of our results is devoted to looking at the properties of elasticity and share estimates. Unfortunately, these estimates depend on income and prices, as well as on the parameters of the demand systems. We have chosen to evaluate them at observation 21, the midpoint of the sample, which is also the point where income and all prices are unity , and where the translog can be interpreted as an approximation to an arbitrary system. The (uncompensated) price elasticity of good i with respect to price j evaluated at this point will be represented by η_{ij} , the income elasticity of good i by μ_i , and the share spent on good i by w_i .

The experiments listed in Table 1 can be divided into two main groups -- those in which the model that is estimated is the one that actually generated the data, and those in which it is not. The former will be discussed first, beginning with the experiments where there was in truth no autocorrelation. We do not consider this to be a realistic case, but it provides a base with which to compare the results of the other experiments, and an introduction to the way in which our results are tabulated.

6. Results when the True System is Estimated

The Case of $\rho=0$. Table 2 tabulates the true values, means and medians for the LES, GLES and translog of the parameters of the systems and of the own price elasticities, income elasticities and shares at the midpoint of the sample. Examination of these results suggests that all three systems do very well in this case, and that bias is virtually absent. This is confirmed by the figures tabulated in the columns headed "Bias". These represent the percentage of the time that the estimated value fell below the true value. If the estimator is unbiased and symmetrically distributed, one would expect this number to be near fifty. More precisely, one would expect the actual number of times that the estimated value fell below the true value (half the numbers reported in the table, since the number of replications was fifty) to be distributed according to the binomial distribution with $n = 50$ and $p = .5$. A test for bias which is robust to possible non-normalities of the estimates is to test whether p in fact equals .5. This can be done by using the normal approximation to the binomial, with continuity correction. Critical values at the .05 level are 35.14 and 64.86 percent. Values outside this range have been marked with an asterisk. There is only one asterisk in Table 2, which is no more than the number that could be expected by chance. Moreover, there is no discernable pattern to the bias percentage, and most of them are very close to fifty. It therefore seems safe to conclude that, when $\rho = 0$, all three systems yield unbiased estimates.

Table 3 tabulates the midspread, range and root mean square errors of the estimates. These three different measures of how accurate the estimates are contain different though overlapping information. The midspread is the

Table 7

<u>Location of Estimates, Translog</u>

$\rho=.9$

		$\hat{\rho}=0$			$\hat{\rho}=1$			$\hat{\rho}$ estimated		
	True Value	Mean	Median	Bias	Mean	Median	Bias	Mean	Median	Bias
α_1	-.300	-.301	-.299	48	-.334	-.335	58	-.301	-.301	52
α_2	-.200	-.197	-.194	36	-.207	-.227	60	-.199	-.197	38
β_{11}	.074	.078	.074	50	.113	.097	40	.078	.085	44
β_{22}	.450	.456	.446	56	.514	.479	42	.447	.440	62
β_{12}	.207	.207	.200	54	.244	.234	44	.209	.212	44
β_{13}	.386	.354	.362	60	.363	.362	62	.370	.380	54
β_{23}	.259	.239	.246	54	.218	.243	58	.256	.260	50
ρ	.900							.751	.775	90*
η_{11}	-.580	-.616	-.579	50	-.472	-.517	40	-.600	-.604	58
η_{22}	-2.330	-2.44	-2.43	60	-2.84	-2.05	32*	-2.34	-2.29	44
η_{33}	-.831	-.871	-.859	62	-.781	-.796	36	-.849	-.841	52
μ_1	.522	.522	.504	54	.476	.522	50	.513	.513	54
μ_2	2.880	2.99	2.99	44	3.73	2.64	64	2.92	2.87	52
μ_3	.536	.518	.523	54	.420	.491	54	.536	.550	44
w_1	.300	.301	.299	52	.334	.335	42	.301	.301	48
w_2	.200	.197	.194	64	.207	.227	40	.199	.197	62
w_3	.500	.501	.502	42	.459	.461	66*	.500	.502	46

Table 8

Dispersion of Estimates, LES

$\rho=.9$

	$\hat{\rho}=0$				$\hat{\rho}=1$				$\hat{\rho}$ estimated			
	Mid-Spread	Range	RMSE	k-s	Mid-Spread	Range	RMSE	k-s	Mid-Spread	Range	RMSE	k-s
q_2	.082	.272	.068	.982	.586	1.01	.342	.383	.043	.990	.109	.007
q_3	.088	.359	.062	.863	.592	1.54	.378	.612	.056	.616	.073	.091
γ_1	.089	.369	.075	.990	.232	228.	40.9	.000	.080	.328	.069	.931
γ_2	.235	.987	.198	.997	.194	4979.	511.	.000	.161	2.71	.377	.001
γ_3	.127	.988	.158	.106	.508	23499.	3382.	.000	.129	.574	.122	.635
ρ									.139	.529	.180	.718
η_{11}	.239	1.14	.212	.905	.635	16.8	2.30	.000	.214	4.80	.645	.000
η_{22}	.670	3.08	.585	.935	.380	6.17	.760	.006	.487	1.62	.338	.990
η_{33}	.150	.692	.129	.798	.482	5.70	.679	.020	.141	.624	.131	.679
μ_1	.162	.530	.125	.671	.867	15.8	1.99	.000	.162	2.00	.272	.002
μ_2	.455	1.93	.431	.937	1.81	47.7	6.08	.000	.353	2.65	.430	.137
μ_3	.190	.761	.131	.784	.348	9.21	1.05	.000	.123	.768	.114	.784
w_1	.015	.043	.010	.823	.351	1.63	.333	.405	.019	.448	.051	.000
w_2	.019	.057	.013	.455	.611	1.16	.338	.725	.021	1.07	.118	.000
w_3	.011	.038	.009	.737	.536	1.47	.358	.539	.012	.624	.070	.000

Table 9

Dispersion of Estimates, GLES

$\rho=.9$

	$\hat{\rho}=0$				$\hat{\rho}=1$				$\hat{\rho}$ estimated			
	Mid-Spread	Range	RMSE	k-s	Mid-Spread	Range	RMSE	k-s	Mid-Spread	Range	RMSE	k-s
q_2	.082	.335	.066	.747	.608	71.9	10.2	.000	.051	.230	.043	.924
q_3	.094	.266	.065	.929	.472	26.8	3.76	.000	.058	.177	.040	.999
γ_1	.175	23.5	4.75	.000	.306	314.	39.7	.000	.293	39.0	6.87	.000
γ_2	.460	56.2	10.1	.000	.913	317.	46.0	.000	.544	73.0	13.8	.000
γ_3	.307	94.6	14.8	.000	.606	72425.	7319.	.000	.704	58.0	13.1	.000
σ	.960	4.36	.860	.984	1.47	104.	12.4	.000	1.13	2.70	.688	.663
ρ									.156	.623	.231	.837
η_{11}	.293	.859	.203	.822	.691	3.50	.598	.546	.246	.625	.155	1.000
η_{22}	.709	2.15	.531	.894	.321	8.03	1.05	.000	.480	1.40	.338	.975
η_{33}	.164	.470	.113	.938	.621	8.29	1.05	.003	.127	.432	.103	.872
μ_1	.200	.555	.130	.689	.505	5.49	.690	.073	.164	.454	.107	.812
μ_2	.455	1.91	.391	.801	1.58	114.	16.3	.000	.299	2.18	.335	.323
μ_3	.204	.536	.129	.579	.335	13.3	1.53	.000	.113	.338	.081	.785
w_1	.015	.044	.010	.728	.351	46.1	6.47	.000	.017	.181	.024	.021
w_2	.017	.055	.013	.201	.636	71.9	10.2	.000	.018	.149	.021	.218
w_3	.016	.051	.011	.890	.487	26.9	3.76	.000	.013	.058	.012	.835

Table 10

Dispersion of Estimates, Translog

$\rho = .9$

	$\hat{\rho}=0$				$\hat{\rho}=1$				$\hat{\rho}$ estimated			
	Mid-Spread	Range	RMSE	k-s	Mid-Spread	Range	RMSE	k-s	Mid-Spread	Range	RMSE	k-s
α_1	.016	.056	.011	.539	.143	.463	.118	.885	.016	.056	.012	.909
α_2	.019	.063	.014	.280	.154	.717	.149	.914	.018	.110	.016	.274
β_{11}	.076	.376	.084	.427	.208	.763	.177	.981	.079	.207	.054	.710
β_{22}	.088	.497	.104	.292	.218	1.00	.210	.532	.067	.481	.086	.143
β_{12}	.093	.316	.072	.911	.200	.741	.159	.970	.059	.296	.062	.936
β_{13}	.122	.479	.113	.984	.139	.699	.141	.624	.110	.335	.082	.920
β_{23}	.151	.362	.095	.947	.183	.734	.163	.888	.059	.232	.051	.979
ρ									.142	.600	.199	.579
η_{11}	.274	1.25	.280	.564	.261	1.52	.348	.021	.253	.823	.201	.952
η_{22}	.714	3.39	.616	.942	.780	27.3	4.40	.000	.424	3.18	.447	.536
η_{33}	.149	.516	.141	.621	.150	.793	.162	.341	.131	.400	.095	.999
μ_1	.213	.851	.186	.912	.239	.847	.219	.164	.129	.596	.116	.977
μ_2	.719	2.94	.580	.999	1.43	36.3	6.09	.000	.433	2.58	.373	.515
μ_3	.248	.599	.165	.903	.227	2.77	.407	.006	.120	.415	.094	.980
w_1	.016	.056	.011	.539	.143	.463	.118	.885	.016	.056	.012	.908
w_2	.019	.063	.014	.280	.154	.717	.149	.914	.018	.110	.016	.274
w_3	.017	.046	.011	.869	.156	.573	.132	.999	.013	.070	.012	.716

preset ρ accordingly, while the practicing econometrician has no such information to help him. However, it seems clear that prior information about ρ may be useful, and that estimating ρ by maximum likelihood is not necessarily the best course when the sample size is small.

Tables 11 and 12 examine the reliability of estimated asymptotic standard errors, tabulating the percent of times that one and two standard errors covered the true values. It is clear that when there is autocorrelation, estimated asymptotic standard errors tend to be too low. This is most pronounced when the systems are estimated in levels. For the translog, estimating in first differences seems to yield standard errors that are quite reliable and even perhaps slightly too large, but this is not the case for the LES or GLES. For all three systems, the standard error associated with the estimate of ρ seems to be particularly unreliable.

The LES is a special case of the GLES. If one estimates the GLES, one may wish to test the hypothesis that the true system is LES by testing whether the true value of σ is one. One result of the underestimation of standard errors is that such a test would often reject the true hypothesis that the system is LES; even if ρ is estimated, that mistake would be made about 34 percent of the time. An alternative way to perform the same test is to look at minus twice the difference of the log likelihood functions from the GLES and LES, which should be distributed as $\chi^2(1)$ if the true system is LES. Even when ρ was estimated, this likelihood ratio test led to rejection of the true hypothesis 28 percent of the time. Thus tests based on standard errors and likelihood ratio tests seem to be reasonably but not entirely consistent, and both have too high a probability of type 1 error.

Table 11

Asymptotic Standard Error Performance, LES, GLES

	LES						GLES					
	$\hat{\rho}=0$		$\hat{\rho}=1$		$\hat{\rho}$ estimated		$\hat{\rho}=0$		$\hat{\rho}=1$		$\hat{\rho}$ estimated	
	N1	N2	N1	N2	N1	N2	N1	N2	N1	N2	N1	N2
q_2	26*	50*	50*	70*	62	86*	24*	46*	44*	62*	46*	76*
q_3	22*	38*	52	66*	50*	82*	12*	32*	38*	50*	44*	66*
γ_1	34*	52*	70	80*	54*	82*	20*	42*	60	72*	54*	70*
γ_2	34*	56*	72	94	58	86*	16*	34*	64	72*	54*	74*
γ_3	30*	56*	56	70*	54*	84	22*	40*	54*	68*	52*	72*
σ							18*	42*	48*	68*	46*	66*
ρ					12*	52*					16*	40*

Table 12

Asymptotic Standard Error Performance, Translog

	$\hat{\rho}=0$		$\hat{\rho}=1$		$\hat{\rho}$ estimated	
	N1	N2	N1	N2	N1	N2
α_1	10*	26*	76	98	36*	62*
α_2	10*	20*	72	94	36*	64*
β_{11}	46*	64*	76	98	54*	94
β_{22}	46*	66*	72	86*	66	92
β_{12}	34*	62*	72	94	62	84*
β_{13}	40*	60*	72	90	54*	86*
β_{23}	24*	40*	76	90	74	88*
ρ					32*	66*

Estimates of the LES and GLES do not always make economic sense (see Section 2 above). If an econometrician estimated one of those systems and found that σ or any of the q_i's were negative, or that any of the shares at the midpoint were negative, or that utility was undefined at the midpoint, he would surely dismiss the estimates as absurd, and assume that the system was incorrectly specified. The number of times that this would have happened in each of our experiments in which the LES or GLES was estimated has been tabulated in Table 13. Note that the case in which the true system is not LES (see section 7 below) is included in the table. Table 13 emphasizes how bad the estimates are when the systems are estimated in differences, but shows that implausible estimates are not restricted to that case.

In order to check that our results did not depend heavily on the choice of ρ, the translog and LES were estimated on data generated with a true ρ of .98, using all three estimating techniques. The results were similar to those for $\rho = .9$, and hence are not tabulated. In particular, we had expected estimating in first differences to perform well when ρ was .98, but it actually performed badly. Estimating ρ seemed to be the best technique for both systems, but estimating in levels again worked reasonably well. We are thus confident that the choice of ρ is not crucial to any of our results, provided that ρ is not too small.

7. The Effects of Misspecification

Few economists believe that observed demand shares are generated by models as simple as the LES and the translog. Nevertheless, they may estimate such models anyway, in the hope that they will thereby get reliable estimates

Table 13

Plausibility of LES and GLES Estimates

True ρ	Estimated ρ	True and Estimated Systems	Percent of Estimates that were Unacceptable
0	0	LES - LES	0
0	0	LES - GLES	0
.9	0	LES - LES	8
.9	0	LES - GLES	18
.9	0	Tlog - LES	2
.9	0	Tlog - GLES	2
.9	1	LES - LES	64
.9	1	LES - GLES	66
.9	1	Tlog - LES	42
.9	1	Tlog - GLES	54
.9	est.	LES - LES	4
.9	est.	LES - GLES	2
.9	est.	Tlog - LES	40
.9	est.	Tlog - GLES	16
.98	0	LES - LES	16
.98	1	LES - LES	30
.98	est.	LES - LES	10

of elasticities and shares. In order to see whether this is a reasonable hope, we estimated the LES and GLES using data generated by the translog, and the translog using data generated by the LES. Nine different experiments were performed, with ρ equal to .9 in all cases, and the usual three error specifications were used for all three systems. Tables 14, 15 and 16 tabulate measures of central tendency and bias which may be compared to those in Tables 5, 6 and 7, and Tables 18, 19 and 20 tabulate measures of dispersion which may be compared to those in Tables 8, 9 and 10. The comparison is generally very unfavorable to the misspecified systems.

Just how severely misspecification causes bias can be seen dramatically in Table 17, which tabulates the percentage of the time that the fitted value fell below the true value for all price and income elasticities, shares and ρ, for all three systems both correctly and incorrectly specified, with ρ always estimated. All the share estimates from the misspecified systems are severely biased, and so are many of the elasticity estimates. Rather remarkably, the LES, which is the most restrictive system, produces less biased estimates than the GLES or the translog, which is supposed to be able to approximate any utility function. It is also of interest that ρ was always over-estimated when the translog was fitted to LES shares, since it was almost always under-estimated when the translog was fitted to translog shares. One may speculate that the very high values of ρ which are sometimes achieved in practice reflect misspecification.

Our results suggest that demand systems such as the LES and the translog cannot be expected to perform well when they did not generate the data. It would

Table 14

Location of Estimates, LES on Translog-Generated Shares

$\rho=.9$

		$\hat{\rho}=0$			$\hat{\rho}=1$			$\hat{\rho}$ estimated		
	Truth	Mean	Median	Bias	Mean	Median	Bias	Mean	Median	Bias
ρ	.900							.899	.974	38
η_{11}	-.580	-.593	-.610	58	-.488	-.483	40	-.606	-.541	48
η_{22}	-2.33	-1.66	-1.64	16*	-1.01	-1.64	22*	7.90	-2.43	52
η_{33}	-.831	-.757	-.771	26*	-.806	-.807	30*	-.782	-.768	40
μ_1	.522	.581	.571	34*	.242	.202	88*	.405	.425	86*
μ_2	2.88	2.46	2.46	76*	1.27	1.93	76*	-18.0	2.89	48
μ_3	.536	.723	.694	2*	.376	.484	60	.498	.519	60
w_1	.300	.307	.306	26*	.287	.248	62	.324	.315	32*
w_2	.200	.188	.184	80*	.126	.092	58	.061	.124	92*
w_3	.500	.505	.507	32*	.587	.558	46	.616	.541	14*

Table 15

<u>Location of Estimates, GLES on Translog-generated Shares</u>

$\rho=.9$

		$\hat{\rho}=0$			$\hat{\rho}=1$			$\hat{\rho}$ estimated		
	Truth	Mean	Median	Bias	Mean	Median	Bias	Mean	Median	Bias
ρ	.900							.823	.818	60
η_{11}	-.580	-.611	-.631	58	-.466	-.561	46	-.654	-.653	60
η_{22}	-2.33	-1.73	-1.72	18*	.182	-1.04	36	-1.19	-2.58	66*
η_{33}	-.830	-.751	-.777	30*	-.827	-.816	44	-.762	-.777	36
μ_1	.522	.579	.560	34*	.296	.415	80*	.452	.443	74*
μ_2	2.88	2.50	2.59	78*	-2.03	1.10	64	.548	3.16	36
μ_3	.536	.697	.687	14*	.550	.571	44	.509	.509	74*
w_1	.300	.307	.306	22*	.310	.312	46	.321	.314	24*
w_2	.200	.191	.187	76*	-.035	.006	86*	.117	.177	88*
w_3	.500	.502	.503	38	.725	.674	16	.563	.514	20*

Table 16

Location of Estimates, Translog on LES-generated Shares

$\rho=.9$

		$\hat{\rho}=0$			$\hat{\rho}=1$			$\hat{\rho}$ estimated		
	True Value	Mean	Median	Bias	Mean	Median	Bias	Mean	Median	Bias
ρ								.977	.979	0*
η_{11}	-.467	-.950	-.899	92*	-.923	-.983	90*	-.847	-.729	84*
η_{22}	-1.00	-2.48	-2.35	96*	-.898	-.999	48	-1.11	-1.09	78*
η_{33}	-.640	-1.00	-1.00	96*	-.401	-.825	68*	-.442	-.847	86*
μ_1	.667	.625	.626	54	.895	1.03	10*	.924	.840	16*
μ_2	2.00	2.99	2.88	8*	1.14	1.11	96*	1.18	1.20	100*
μ_3	.800	.456	.463	94*	.789	.917	32*	.686	.959	16*
w_1	.300	.300	.297	56	.043	-.026	62	.159	.253	80*
w_2	.200	.199	.196	62	.864	.791	32*	.600	.335	2*
w_3	.500	.501	.502	42	.093	.118	68*	.241	.415	94*

Table 17

Bias in Elasticity and Share Estimates

$\rho = .9$, $\hat{\rho}$ estimated

True: Est.:	Translog Translog	Translog LES	Translog GLES	LES Translog	LES LES	LES GLES
η_{11}	58	48	60	84*	52	48
η_{21}	48	26*	36	0*	48	56
η_{31}	44	32*	20*	48	50	44
η_{12}	58	40	34*	44	56	50
η_{22}	44	52	66*	78*	44	50
η_{32}	56	64	80*	28*	54	50
η_{13}	42	20*	10*	48	52	48
η_{23}	54	38	70*	0*	54	56
η_{33}	52	40	36	86*	46	50
μ_1	54	86*	74*	16*	50	58
μ_2	52	48	36	100*	52	48
μ_3	44	60	74*	16*	46	44
w_1	48	32*	24*	80*	54	52
w_2	62	92*	88*	2*	56	64
w_3	46	14*	20*	94*	48	44
ρ	90*	38	60	0*	82*	92*

Table 18

Dispersion of Estimates, LES on Translog-Generated Shares

	$\hat{\rho}=0$				$\hat{\rho}=.9$ $\rho=1$				$\hat{\rho}$ estimated			
	Mid-Spread	Range	RMSE	K-S	Mid-Spread	Range	RMSE	K-S	Mid-Spread	Range	RMSE	K-S
ρ									.173	.455	.121	.003
η_{11}	.233	1.02	.199	.490	.313	1.22	.278	.713	.373	1.22	.258	.719
η_{22}	.901	2.85	.958	.991	2.37	28.2	4.89	.052	4.16	327.	46.9	.000
η_{33}	.146	.558	.140	.962	.095	.516	.105	.720	.239	.519	.152	.747
μ_1	.161	.525	.138	.982	.464	.680	.367	.006	.138	.617	.181	.192
μ_2	.685	1.85	.644	.979	3.49	44.3	7.33	.027	6.66	715.	101.	.000
μ_3	.180	.552	.225	.281	.621	.813	.335	.007	.115	.757	.140	.456
w_1	.015	.047	.012	.858	.212	.565	.133	.278	.060	.331	.068	.146
w_2	.016	.054	.018	.473	.581	1.23	.342	.334	.207	.879	.223	.157
w_3	.012	.036	.010	.960	.429	1.01	.279	.419	.212	.931	.198	.026

Table 19

Dispersion of Estimates, GLES on Translog-Generated Shares

$\rho = .9$

	$\hat{\rho}=0$				$\hat{\rho}=1$				$\hat{\rho}$ estimated			
	Mid-Spread	Range	RMSE	K-S	Mid-Spread	Range	RMSE	K-S	Mid-Spread	Range	RMSE	K-S
ρ									.255	.530	.164	.340
η_{11}	.295	.924	.224	.999	.346	2.26	.445	.106	.379	1.19	.251	.755
η_{22}	1.17	3.20	.992	.991	4.10	168.	21.1	.000	1.10	105.7	12.5	.000
η_{33}	.197	.519	.146	.765	.115	.557	.106	.921	.235	.702	.179	.982
μ_1	.166	.609	.145	.959	.289	1.81	.401	.084	.148	.859	.144	.931
μ_2	.747	2.19	.672	.883	6.17	319.	41.3	.000	1.18	215.3	25.7	.000
μ_3	.234	.635	.218	.978	.160	.746	.175	.176	.103	.503	.103	.346
w_1	.015	.047	.013	.912	.189	.605	.139	1.000	.027	.377	.056	.017
w_2	.018	.057	.017	.245	.327	1.12	.350	.950	.099	.605	.150	.007
w_3	.017	.050	.011	.921	.332	1.13	.342	.610	.080	.686	.133	.001

Table 20

<u>Dispersion of Estimates, Translog on LES-Generated Shares</u>

$\rho = .9$

	$\hat{\rho}=0$				$\hat{\rho}=1$				$\hat{\rho}$ estimated			
	Mid-Spread	Range	RMSE	K-S	Mid-Spread	Range	RMSE	K-S	Mid-Spread	Range	RMSE	K-S
ρ									.041	.081	.081	.113
η_{11}	.562	1.39	.602	.797	.503	34.7	4.10	.000	.420	11.3	1.38	.002
η_{22}	.919	4.57	1.74	.837	.066	6.64	.84	.000	.168	.744	.179	.842
η_{33}	.270	.912	.418	.994	.653	12.9	2.12	.000	.382	34.2	4.42	.000
μ_1	.404	1.38	.306	.893	.392	32.8	3.61	.000	.307	5.13	.729	.005
μ_2	1.03	3.52	1.22	.721	.334	12.5	1.65	.000	.187	1.81	.862	.002
μ_3	.334	.723	.403	.811	.433	6.41	.849	.028	.202	13.2	1.81	.000
w_1	.017	.061	.012	.477	1.17	3.88	.878	.906	.197	2.34	.368	.006
w_2	.018	.067	.014	.405	1.65	10.32	1.77	.214	.632	2.03	.609	.007
w_3	.018	.055	.012	.999	.962	7.43	1.12	.181	.363	1.02	.399	.003

clearly be of interest to experiment with other systems and with data generated in a more complicated way, but that would exceed the scope of this chapter.

8. Conclusions

The experiments performed in this paper have attempted to replicate as far as possible the situation faced by researchers attempting to estimate demand systems from aggregate time series data. We hope that our conclusions will influence future research in this area. They are as follows.

1. If the true system is known and if the true error terms do not exhibit autocorrelation, it appears to be possible to get excellent estimates.

2. If the true system is not known, estimates are likely to be biased and unreliable.

3. Estimating in first differences should always be avoided. When there is autocorrelation the best procedure is probably to estimate ρ , but estimating in levels is not bad, and may actually give the best share estimates. Prior information about ρ may be very valuable.

4. Estimated asymptotic standard errors are usually too low when there is autocorrelation, sometimes by a substantial margin.

5. The distributions of parameter and other estimates are often far from normal when there is autocorrelation, and extreme values are likely to be encountered.

Footnotes

* We are greatly indebted to Professor R.E. Quandt for helpful advice at all stages of this paper's preparation. We should also like to thank O.C. Ashenfelter and S.M. Goldfeld for valuable suggestions. Any errors that remain are ours alone. All computations were performed on an IBM 360/91 belonging to Princeton University.

1. Abbott and Ashenfelter [1974] and Parks [1969] each estimated several different demand systems using the same data, and found that elasticities varied substantially with the model estimated.

2. Some examples are: Abbott and Ashenfelter [1974], 39 observations; Kiefer [1975], 39 observations; Lluch and Williams [1975], 43 observations; MacKinnon [1975], 38 observations.

3. Examples of the former include Brown and Heien [1972], and Wales [1971]; examples of the latter include Abbott and Ashenfelter [1974] and Kiefer [1975].

References

Abbott, Michael and Orley C. Ashenfelter, "Labor Supply, Commodity Demand, and the Allocation of Time," Industrial Relations Section Working Paper #57, Princeton University, November 1974.

Barten, A.P., "Maximum Likelihood Estimation of a Complete System of Demand Equations," European Economic Review, 1, 1969, pp. 7-73.

Berndt, E.R. and N.E. Savin, "Estimation and Hypothesis Testing in Singular Equation Systems with Autoregressive Disturbances," Econometrica, forthcoming 1975.

Brown, A. and A. Deaton, "Models of Consumer Behavior: A Survey," Economic Journal, 82, 1972, pp. 1145-1236.

Brown, M. and D. Heien, "The S-Branch Utility Tree: A Generalization of the Linear Expenditure System," Econometrica, 40, 1972, pp. 737-747.

Christensen, L.R. and Dale Jorgenson and Lawrence J. Lau, "Transcendental Logarithmic Utility Functions," Institute for Mathematical Studies in the Social Sciences Technical Report #94, Stanford University, March 1973.

Goldfeld, S.M. and R.E. Quandt, Nonlinear Methods in Econometrics, Amsterdam, North-Holland, 1972.

Kendall, M.G. and A. Stuart, The Advanced Theory of Statistics, Vol. 2, New York, Hafner, 1961.

Kiefer, Nicholas M., "Quadratic Utility, Labor Supply and Commodity Demand," 1975.

Lluch, C. and R. Williams, "Consumer Demand Systems and Aggregate Consumption in the U.S.A.: An Application of the Extended Linear Expenditure System," Canadian Journal of Economics, 8, 1975, pp. 49-66.

MacKinnon, James G., "Estimating the Linear Expenditure System and its Generalizations," 1975.

Parks, R.W., "Systems of Demand Equations: An Empirical Comparison of Alternative Functional Forms," Econometrica, 37, 1969, pp. 629-650.

Quandt, R.E., "A Note on Amemiya's Nonlinear Two-Stage Least Squares Estimators," Econometric Research Program, Princeton University, Memorandum #178, 1975.

Tukey, J.W., Exploratory Data Analysis, Vol. 1. Limited Preliminary Edition, New York, McGraw-Hill, 1971.

Wales, Terence J., "A Generalized Linear Expenditure Model of the Demand for Nondurable Goods in Canada," Canadian Journal of Economics, 4, 1971, pp. 1471-1484.

Appendix

The income and price series that were used in all the experiments reported on in this paper are listed in Table A1. These series were generated by random processes of the following type:

$$p_t^i = p_{t-1}^i (g^i + e_t^i + a^i v_t) \ .$$

g^1 was 1.038, g^2 was 1.042, g^3 was 1.03 and g^4 (for income) was 1.06. a^1, a^2 and a^3 were 1.0, and a^4 was 1.5. v_t was normally distributed with mean 0 and variance .0004. e_t^i was normally distributed with mean 0 and variance $.004[(g^i-1)/(g^4-1)]^2$. All series were then normalized to unity at observation 21.

Table A1

Income and Price Series

OBS. NO.	INCOME	PRICE 1	PRICE 2	PRICE 3
1	0.4826434188	0.6045165103	0.4919483239	0.6351125227
2	0.4978299400	0.6167759478	0.5147707013	0.6544333350
3	0.5282387831	0.6368720140	0.5409417521	0.6736539433
4	0.5320597814	0.6408252114	0.5427287832	0.6681351526
5	0.5416917579	0.6526848978	0.5533529507	0.6880148021
6	0.5258896341	0.6522707199	0.5756706249	0.6957046010
7	0.5428726165	0.6538655089	0.5973213641	0.6961745820
8	0.5675151048	0.6710305370	0.6014462881	0.7068263661
9	0.6111275552	0.6990311078	0.6354573704	0.7360304757
10	0.6378763315	0.7242941669	0.6587750185	0.7517335907
11	0.7095817908	0.7839113079	0.7131350589	0.8108584362
12	0.7419041107	0.8219697274	0.7402003404	0.8278253226
13	0.7429461816	0.8361590007	0.7644212013	0.8414826498
14	0.7581331759	0.8568834328	0.7911097829	0.8566750955
15	0.7944493542	0.8679485667	0.8157872732	0.8670957486
16	0.3299321306	0.8687197010	0.8249883268	0.8929878367
17	0.8719066969	0.9098058484	0.8906151011	0.9374514025
18	0.9087000913	0.9410872797	0.9386111208	0.9727064460
19	0.9330066199	0.9591130362	0.9707463920	0.9864336214
20	0.9859055356	0.9762876921	0.9635471580	1.0028449069
21	1.0000000000	1.0000000000	1.0000000000	1.0000000000
22	1.0864547846	1.0267593036	1.0492333232	1.0334960087
23	1.1827186056	1.0731057541	1.1039573306	1.0614933903
24	1.1773937310	1.0590655314	1.0996421693	1.0518193878
25	1.2392871709	1.1229799790	1.1885546958	1.0898486693
26	1.2975118640	1.1117031498	1.2005588166	1.0780654056
27	1.3521659766	1.1078174028	1.2059354759	1.0935070733
28	1.4711872051	1.1741978369	1.2311217618	1.1317301952
29	1.5267519019	1.2110563326	1.2730187866	1.1669866054
30	1.5692298035	1.2455384605	1.3545664416	1.2093885851
31	1.7017670337	1.3122133052	1.4109880332	1.2497278427
32	1.8148178253	1.3719849341	1.5137185049	1.2824789895
33	1.8470124826	1.3565331234	1.5600526034	1.2609382053
34	1.9211384217	1.4350591998	1.5792824356	1.2977671910
35	1.9227209464	1.4862545524	1.6142118144	1.3064517036
36	2.0084080157	1.5065083411	1.6317899646	1.3004790064
37	2.1101382967	1.5686036886	1.6860173216	1.3158425133
38	2.3160238818	1.7206465497	1.7770283490	1.3984663556
39	2.4718281874	1.8456920422	1.8642437427	1.4667840713
40	2.6324755309	1.9008523332	1.9935325202	1.4997394076