ROBUST ESTIMATION OF ENGEL CURVES WITH CENSORED DATA

Franco Peracchi*

University of California, Los Angeles

Econometric Research Program
Research Memorandum No. 333

October 1987

* This paper is based on a chapter of my dissertation at Princeton University. I am very grateful to my advisors, Angus Deaton, Whitney Newey, and Elvezio Ronchetti, for their generous guidance. I also thank Sule Ösler, Duncan Thomas, and the participants to the Princeton and UCLA econometric seminars for helpful comments.

Econometric Research Program
Princeton University
207 Dickinson Hall
Princeton, NJ 08544, USA

ABSTRACT

In this paper we compare the Tobit ML estimator with a number of semi-parametric and bounded-influence estimators for the censored regression model. The comparison is carried out on the basis of an empirical example, in which we estimate Engel curves using household budget data containing a significant fraction of reported zero expenditure. The ML estimator appears to be very sensitive to extreme observations and is way off in some cases. Semi-parametric and bounded-influence estimators are close to each other and look more reliable. However, bounded-influence estimators appear to be more precise, and provide diagnostic information that is useful to identify sources of model failures.

1. Introduction

Models of zero expenditure typically involve limited dependent variables, and are usually estimated by ML under assumptions such as normality and homoskedasticity. Unlike the linear regression model, however, failures of these distributional assumptions imply inconsistency of the ML estimates based on the misspecified distribution, and the resulting asymptotic bias may be very large.

The purpose of this paper is to compare the Tobit ML estimator for the censored regression model with two types of robust estimators. The first type are semi-parametric estimators based on the method of moments. Moments are chosen so as to guarantee consistency for a broad class of distributions of the observations. We consider Powell's (1984) censored least absolute deviation estimator (CLAD) and Powell's (1986) symmetrically censored least squares (SCLS) estimator.

The second type are optimal bounded-influence estimators of the type discussed in Hampel et al. (1986) and Peracchi (1987). These estimators cannot be improved upon simultaneously with respect to the criteria of efficiency at the assumed parametric model and protection against the bias that may arise because of local failures of the model assumptions. This property is not shared by the Tobit ML estimator, nor by the semi-parametric estimators that we consider. Optimal bounded-influence estimators may be interpreted as weighted ML estimators, and differ with respect to the choice of weight function.

The comparison between the various estimators is carried out on the basis of an empirical example, in which we estimate Engel curves using cross-section data from the Sudan. We address the following questions:

Is the normal Tobit model consistent with the data? Do robust estimators lead to different conclusions than ML and why? What are the differences between semi-parametric and bounded-influence estimators? What diagnostic information is provided by the various methods?

Our findings may be summarized as follows. The joint hypothesis of normality and Tobit specification is often at odds with the data. ML estimates tend to be very sensitive to extreme observations and may be way off in some cases. Robust estimates are close to each other and look more reliable. However, bounded-influence estimates appear to be more precise than CLAD and SCLS. Finally, robust weights provide useful information for identifying sources of model failures, in particular outliers and leverage points.

The plan of the paper is as follows. Section 2 discusses several models of zero expenditure. It also presents some of the estimators and tests used later in the empirical example. Section 3 introduces the class of bounded-influence estimators for the Tobit model. Section 4 contains the empirical results. Section 5 summarizes the conclusions.

2. Models of zero expenditure

Limited dependent variable models are widely used to model the income-expenditure relationship when household budget data contain a significant fraction of reported zero expenditures. In these cases it is commonly assumed that observed or recorded demand y is a function of an underlying latent variable y^*, which itself depends on a k-dimensional vector of explanatory variables, including income and other household characteristics. The conditional distribution of y^* given x is assumed to be centered about some linear combination of the x's. Thus the model for y^* takes the familiar regression form

$$(1) \qquad y^* = x'\beta_0 + r\,\sigma_0$$

where $(\beta_0, \sigma_0) \in \mathbb{R}^k \times \mathbb{R}_+$ are unknown parameters, and r is an unobservable disturbance. Given the relationship between y and y^*, which needs to be neither 1-1 nor continuous, and given a set of assumptions on the joint distribution of y^* and x, estimation and inference about the parameter of interest β_0 is carried out on the basis of a sample of N observations on $z = (y, x')'$.

The truncated regression model arises when the relationship between y and y^* is only defined for $y^* > 0$, in which case $y = y^*$. One interpretation of this model is that some or all the explanatory variables are unobservable when demand is equal to zero. The censored regression model arises when $y = \max(0, y^*)$. Unlike the truncated regression model, no information on the explanatory variables is lost when demand is equal to zero. The censored regression model is widely used, but there are two problems associated with it. The first is its validity as a representation of demand behavior. The second, common to all models involving limited dependent variables, is the lack of robustness of the ML estimator based on the normality assumption. We will consider these two problems in turn.

Cragg (1971) first pointed out that the censored (and truncated) regression model may not be a valid representation of demand, because it does not distinguish between the decision of purchasing a good and the decision of how much to purchase. As an alternative he suggested superimposing to it a binary censor. Thus $y = 1(w^* > 0) \max(0, y^*)$, where $1(A)$ is the indicator function of the event A, and w^* is a latent variable with $\Pr[w^* > 0] = \pi_0$, an unknown parameter which may differ across households. One interpretation of this model is that observed demand is zero either because the household genuinely does not consume the good, or because for some reason (e.g. infrequency of purchase or misreporting) a zero expenditure was recorded. The latter event occurs with probability π_0. The model reduces to censored regression when $\pi_0 = 1$ for all households. This restriction can easily be tested [see e.g. Deaton and Irish (1984)].

In a second model suggested by Cragg the position of the two 'hurdles' is reversed. First a binary censor w^* determines whether a purchase is to be made. The amount purchased is then determined by a truncated regression model. This is equivalent to replacing in the previous model the censored value of y^* with its truncated value. The model reduces to censored regression when $\Pr[w^* > 0] = \Pr[y^* > 0]$. This restriction can easily be tested [see e.g. Lin and Schmidt (1984)].

Both models of Cragg are versions of the Type 2 Tobit model of

Amemiya (1985) and are closely related to the "sample selectivity" models of labour economics.

Table 1 gives the likelihood score function $s(z,\theta)$ for all four models under the assumption that errors in (1) are iid with a standard normal distribution. For convenience model (1) is reparameterized by putting $\alpha = \beta/\sigma$ and $\gamma = 1/\sigma$. The score for α is always of the form $s_{\alpha}(z,\theta) = \lambda(z,\theta)\, x$, where the conditional expectation of the 'residual' $\lambda(z,\theta)$ given x will not be zero in general unless the errors in (1) are normal and homoskedastic. This implies that for all four models the ML estimator of α based on the normality assumption will not be consistent in general when the errors in (1) are non-normal or heteroskedastic.

For the truncated and censored regression model, this lack of robustness of the normal ML estimator has been investigated in a few, simple cases [see e.g. Hurd (1979), Arabmazar and Schmidt (1981) and (1982), Goldberger (1983) and Kiefer and Skoog (1984)]. A general finding is that the bias under non-normality or heteroskedasticity can be very large, particularly if the scale parameter σ_0 is unknown and the degree of censoring is high. The Tobit ML estimator is generally less biased than the truncated regression estimator, which suggests that the limit observations should be used when available. However, most studies only consider the problem of estimating the mean of a population. Further, they restrict attention to symmetric distributions and do not investigate explicitly the relationship between the bias and the tail behavior of the error distribution. This is unfortunate, because the bias depends also on the distribution of the regressors, and is likely to be more severe for non-symmetric or thick-tailed distributions.

The normality assumption may be tested in various ways. One is to nest the normal distribution in a larger parametric family and then construct a score test of the restrictions implied by normality. For example, the test of Bera, Jarque and Lee (1984) is based on the Pearson family, but other choices of nesting family are posssible [see e.g. Ruud (1984)]. Although designed against one specific alternative, tests of this type have power against a number of misspecification alternatives.

A second way is to construct general specification tests based on the comparison of two estimators that are both consistent at the assumed model, but have different probability limits when the model is misspecified. The various tests differ in the choice of what estimators to compare. For example, Nelson (1981) compares a consistent estimator of the covariance of x and y with the efficient estimator based on the assumption of a normal censored regression (Tobit) model. Ruud (1984) compares the Probit and Tobit ML estimators of α. Chesher, Lancaster and Irish (1985) compare alternative estimators of the information matrix for the Tobit model, as suggested by White (1982). All these tests can be interpreted as conditional moment tests [Newey (1985)]. They are not specifically designed to test normality and may lack power against certain alternatives.

A third possibility is to use graphical methods based on some non-parametric estimator of the error distribution, such as the Kaplan-Meier estimator for the censored regression case [see e.g. Chesher, Lancaster and Irish (1985)].

Estimators that are consistent under weak distributional assumptions have been proposed for the truncated and censored regression models. In this paper we will consider two semi-parametric estimators based on the method of moments, where the particular choice of moments guarantees consistency for a broad class of distributions. Interestingly, neither estimator requires a knowledge of the scale parameter σ_0, which may help explain their robustness properties. However, neither estimator is consistent when the censored regression model is misspecified.

The censored least absolute deviation (CLAD) estimator [Powell (1984)] is consistent and asymptotically normal provided that the conditional error distribution has median zero (homoskedasticity is not required), but estimating its AVM requires estimating the error density at the origin. When it exists, the CLAD score function is given by

$$\eta(z,\beta) = 1(x'\beta > 0) \text{ sign } (y - x'\beta) \ x.$$

The symmetrically censored least squares (SCLS) estimator [Powell (1986)] is consistent and asymptotically normal under the somewhat stronger assumption that the conditional error distribution is symmetric about zero (again homoskedasticity is not required). When it exists, the SCLS score function is given by

$$\eta(z,\beta) = 1(x'\beta > 0) \ [\min (y, 2x'\beta) - x'\beta] \ x.$$

A version of this estimator is also available for the truncated regression model. A test of the symmetry assumption has been proposed by Newey (1987). Newey's test compares the SCLS estimator with a more efficient estimator under symmetry.

Monte Carlo evidence [Paarsch (1984), Powell (1986)] indicates that both the CLAD and SCLS estimator may be very inefficient relative to the ML estimator based on a correctly specified model. This raises the question of whether too much information is ignored in order to attain consistency under very general conditions. Furthermore, neither estimator has a bounded influence function (IF), and therefore neither estimators is bias-robust in the sense of Hampel (1974). It is therefore interesting to compare them with some bounded-influence estimators for the censored regression model.

3. Optimal bounded-influence estimators for the Tobit model

In this Section we assume that the distribution of the observations is in a neighborhood of the normal censored regression (Tobit) model. This information is exploited to construct 'optimal bounded-influence' estimators of $\theta_0 = (\alpha_0', \gamma_0)'$. These estimators have minimum asymptotic mean square error (MSE) among all estimators that are consistent at the assumed model and have a bounded IF. This implies that optimal bounded-influence estimators cannot be improved upon simultaneously with respect to the criteria of efficiency at the assumed model and protection against the bias that may arise because of local failures of the model assumptions.

It can be shown [see e.g. Peracchi (1987)] that when the estimator's sensitivity [the sup-norm of its IF with respect to some metric] is bounded by a given constant c, and the sensitivity and the MSE are both defined in the metric of a pd matrix B, an optimal bounded-influence estimator $\tilde{\theta}$ of θ_0 may be based on the score function

$$\eta(z,\theta) = w(z,\theta)\ [s(z,\theta) - a(\theta)] \tag{2}$$

where $s(z,\theta)$ is the Tobit likelihood score and $w(z,\theta)$ is a scalar weight function defined by

$$w(z,\theta) = \min\left\{ 1, \frac{c}{\|P(\theta)^{-1}\ [s(z,\theta) - a(\theta)]\|_B} \right\}. \tag{3}$$

The various estimators differ with respect to the choice of the metric B and the sensitivity bound c. The symmetric, pd matrix $P(\theta)$ and the vector $a(\theta)$ are roots of the system of equations

$$(4) \qquad E_\theta\, w(z,\theta)\, [s(z,\theta) - a] = 0$$

$$(5) \qquad E_\theta\, w(z,\theta)\, [s(z,\theta) - a]\, s(z,\theta)' - P = 0,$$

where expectations are taken with respect to the assumed Tobit model. A solution to (4)-(5) exists provided that c is large enough [see Peracchi (1987) for a necessary lower bound].

The optimal bounded-influence estimator $\tilde{\theta}$ may be interpreted as a weighted ML estimator, where the weight function $w(\cdot,\theta)$ depends on the matrix B. When $B = I_p$, $\tilde{\theta}$ is the Tobit analogue of the regression estimator of Hampel (1978) and Krasker (1980). When $B = AV(\tilde{\theta},F_\theta)^{-1}$ we obtain the analogue of the regression estimator of Krasker and Welsch (1982). Other choices of B will be discussed later. The vector $a(\theta)$ is a bias correction term that depends on the assumed model F_θ and ensures that $\tilde{\theta}$ is (Fisher) consistent for θ at F_θ. Geometrically, the likelihood score for one observation is shrunk to satisfy the bounded-influence constraint, and twisted to ensure consistency at the assumed model.

If the Tobit model is correctly specified, the optimal bounded-influence estimator $\tilde{\theta}$ can be shown to be consistent and asymptotically normal with AVM equal to $P_0^{-1}\, Q_0\, P_0'^{-1}$, where $P_0 = E_0\, \eta(z,\theta_0)\, s(z,\theta_0)'$ and $Q_0 = E_0\, \eta(z,\theta_0)\, \eta(z,\theta_0)'$. The proof is a straightforward application of results of Huber (1967) and Amemiya (1985).

When the assumption of iid normal errors in (1) is violated, $\tilde{\theta}$ is generally inconsistent. However, and this is the main reason for using

$\tilde{\theta}$, if misspecification occurs because the assumed model is contaminated with probability ε by some other distribution, then the norm of the asymptotic bias of $\tilde{\theta}$ (in the metric of B) cannot exceed εc, provided that ε is small [see Hampel et al. (1986) for details]. A consistent estimate of the AVM of $\tilde{\theta}$ in the case of heteroskedasticity or non-normality is given by $\tilde{P}_N^{-1} \tilde{Q}_N \tilde{P}_N'^{-1}$, where $\tilde{P}_N = N^{-1} \Sigma_{n=1}^{N} (\partial/\partial\theta') \eta(z_n,\tilde{\theta}_N)$ and $\tilde{Q}_N = N^{-1} \Sigma_{n=1}^{N} \eta(z_n,\tilde{\theta}_N) \eta(z_n,\tilde{\theta}_N)'$.

Often the distribution of the regressors is unknown, and so the vector $a_0 = a(\theta_0)$ and the pd matrix $P_0 = P(\theta_0)$ have to be estimated jointly with θ_0. One possibility is to solve the equation system

$$(6) \qquad N^{-1} \Sigma_n \min\left\{ 1, \frac{c}{\|P^{-1}[s(z_n,\theta) - a]\|_B} \right\} [s(z_n,\theta) - a] = 0$$

$$(7) \qquad N^{-1} \Sigma_n E_\Phi\, \omega_n(r,\theta,a,P)\ [\sigma_n(r,\theta) - a] = 0$$

$$(8) \qquad N^{-1} \Sigma_n \{E_\Phi\, \omega_n(r,\theta,a,P)\ [\sigma_n(r,\theta) - a]\ \sigma_n(r,\theta)'\} - P = 0,$$

where r is a random variable with a standard normal df. The function

$$\sigma_n(r,\theta) = \begin{bmatrix} [1(r > -x_n'\alpha)\ r - 1(r \le -x_n'\alpha)\ \lambda(x_n'\alpha)]\ x_n \\ 1(r > -x_n'\alpha)\ (x_n'\alpha + r)\ /\ \gamma \end{bmatrix}$$

corresponds to the Tobit score, $\lambda(u)$ is the normal Mill's ratio, and $\omega_n(r,\theta,a,P) = \min\ \{1,\ c\ /\ \|P^{-1}\ [\sigma_n(r,\theta) - a]\|_B\}$. Let π denote the vector of $p(p + 1)/2$ distinct elements of the symmetric matrix P, and let $\pi_0 = \pi(\theta_0)$ denote the corresponding population value. Stacking (6), (7) and the p(p + 1)/2 distinct equations of (8) together gives a system of 2p + p(p + 1)/2 equations in the unknown vectors θ, a and π, namely

(9) $$N^{-1} \Sigma_{n=1}^{N} \varphi(z_n,\mu) = 0,$$

where $\mu = (\theta', a', \pi')'$. Let $\tilde{\mu}$ be an M-estimator defined as a root of equation (9). It can be shown that $\tilde{\mu}$ is consistent if the Tobit model is correctly specified, and is asymptotically normal with AVM equal to $U_0^{-1} V_0 U_0'^{-1}$, where $U_0 = (\partial/\partial\mu') E_0 \varphi(z,\mu_0)$ and $V_0 = E_0 \varphi(z,\mu_0) \varphi(z,\mu_0)'$.

The AVM of $\tilde{\mu}$ is generally not block-diagonal. Thus, unlike the linear regression case, estimation of (a, π) does affect the asymptotic distribution of $\tilde{\theta}$. In particular, when the error distribution is correctly specified, the AVM of $\tilde{\theta}$ exceeds $P_0^{-1} Q_0 P_0'^{-1}$ by the psd matrix $P_0^{-1} C(\mu_0,F_0) AV((\tilde{a},\tilde{\pi}),F_0) C(\mu_0,F_0)' P_0'^{-1}$, where $C(\mu_0,F_0)$ is the expectation of the gradient of the score for θ with respect to (a, π). When the error distribution is misspecified, the AVM of $\tilde{\theta}$ may still be estimated consistently by the p×p top-left sub-matrix of $\tilde{U}_N^{-1} \tilde{V}_N \tilde{U}_N'^{-1}$, where $\tilde{U}_N = N^{-1} \Sigma_{n=1}^{N} (\partial/\partial\mu') \varphi(z_n,\tilde{\mu}_N)$ and $\tilde{V}_N = N^{-1} \Sigma_{n=1}^{N} \varphi(z_n,\tilde{\mu}_N) \varphi(z_n,\tilde{\mu}_N)'$.

The computation of $\tilde{\theta}$ may be quite expensive, but considerable simplifications can be obtained by exploiting the arbitrariness of the metric in which the norm of the IF is defined. Here we propose two possibilities. The first is to choose $B = P(\theta)^2$. Although not very natural, this metric is convenient from the point of view of computation, since it eliminates the need of solving for the matrix P at each iteration. However, the resulting sensitivity measure is not invariant under a reparameterization of the model. One choice that leads to invariance is $B = P(\theta) J(\theta)^{-1} P(\theta)$, where $J(\theta)$ is the information matrix associated with the parametric model F_θ. The resulting weight function, which is also computationally simple, rescales the recentered likelihood score whenever its norm, in the metric of the information matrix, is greater than the given bound c. The estimators based on these two choices of weight function are denoted by BI1 and BI2 respectively.

Table 2 summarizes the score function for each of the estimators that we consider. BI0 is the estimator based on a score of the form (2), with $a(\theta)$ given by (4) and with weight function given by

$$w(z,\theta) = \min \{1, c / \|s(z,\theta)\|\}.$$

The BI0 estimator is simplest to compute because $w(z,\theta)$ does not depend on $a(\theta)$. It is easy to verify that the BI0 estimator is consistent at the Tobit model, has a bounded IF and is asymptotically normal. Therefore, it should provide good starting values for one-step versions of the H-K and K-W estimators. The method of Bickel (1975) can be used to show that these one-step estimators are asymptotically equivalent to the fully iterated estimators.

4. Empirical application

The data set that we are going to analyze is taken from the 1978-80 Household Income and Expenditure Survey of the Sudan [for a description of the data see Deaton and Case (1985)]. It was chosen as an example of the type of data that are often used by economists. The sources of contamination may be several - misreporting by individual households, coding and punching errors, data manipulation at the editing stage - but the actual amount of contamination is unknown.

The original data set contains observations from different regions of the Sudan. To keep the model as simple as possible, we only consider the subset of 268 observations from the Nile region. We estimate Engel curves for 3 commodities with a non-negligible fraction of reported zero expenditures, namely clothing and footwear ('clothing'), transport services and repairs ('transport'), and tobacco products ('tobacco'). The degree of censoring is differnt for the various commodities and is equal to 8.2% for clothing, 23.9% for transport, and 31.7% for tobacco.

We consider a number of popular models of Engel curves for an individual commodity i:

WL: $$w_i = a_i(p) + b_i(p) \ln x$$

QWL: $$w_i = a_i(p) + b_i(p) \ln x + d_i(p) (\ln x)^2$$

LES: $$p_i q_i = a_i(p) + b_i(p) x$$

QES: $$p_i q_i = a_i(p) + b_i(p) x + d_i(p) x^2$$

where w_i is the budget share, $p_i q_i$ is total expenditure on commodity i, p is the vector of prices of all commodities and x is total outlay. WL is the so-called Working-Leser form. QWL adds to WL the square of ln x. LES and QES are the Engel curves corresponding respectively to the linear and quadratic expenditure system. All curves belong to the general class of Engel curves considered by Gorman (1981), and are all theory consistent in the sense that each of them can be derived by Shephard's Lemma from some nice cost function. QWL and QES may be interpreted as second order approximations, based respectively on powers of ln x and of x, to an arbitrary Engel curve.

Demographic and area effects are introduced in the analysis by expressing income in per capita terms, and by assuming that for each model the intercept a_i depends linearly on a number of household characteristics: household size, a household composition effect (number of household members less than 14 years old), and an area dummy (DRUR) with a value of one for households living in rural areas and zero for households living in urban areas. This approach may be restrictive, because demographic and area effects may affect the whole set of parameters.

For each commodity and functional form we first consider a number of tests for normality and conditional symmetry of the error distribution and for the Tobit specification. The normality assumption is tested against the general Pearson family using the score test of Bera, Jarque and Lee (1984). The specification tests of Nelson (1981) and Ruud (1984) are also considered. The Tobit specification is tested against Cragg's (1971) first model using the score test of Deaton and Irish (1984), and against Cragg's second model using the score test of Lin and Schmidt

(1984). The joint hypothesis of conditional symmetry and Tobit specification is tested as in Newey (1987).

The Deaton-Irish test essentially compares a consistent estimate of Pr (y > 0) with the efficient estimate based on the Tobit model, namely $N^{-1} \Sigma_{n=1}^{N} \hat{\Phi}_n$, where $\hat{\Phi}_n = \Phi(x_n'\hat{\alpha}_{ML})$. The Lin-Schmidt test compares a consistent estimate of E (xy | y > 0) with the efficient estimate for the Tobit model, namely $N^{-1} \Sigma_{n=1}^{N} 1(y_n > 0) [x_n'\hat{\alpha}_{ML} + \hat{\phi}_n/\hat{\Phi}_n] x_n$, where $\hat{\phi}_n = \phi(x_n'\hat{\alpha}_{ML})$. Thus both tests are related to Nelson's specification test, and have power against a variety of alternatives, including heteroskedasticity and non-normality.

All tests are based on the maintained hypothesis that Engel curves are correctly specified. Therefore, they all have power against misspecification arising from omitted variables or an incorrect functional form. All score test statistics are computed in an asymptotically equivalent form as N times the uncentered R^2 in the regression of a column of ones on the likelihood score for the unrestricted model (evaluated at the restricted estimates). Under the null hypothesis, all test statistics except the Deaton-Irish statistic have an asymptotic χ^2 distribution. The number of degrees of freedom is equal to 2 for the Bera-Jarque-Lee test, and to the number of regressors for all other tests. The Deaton-Irish statistic is the square root of the score test statistic, with the same sign as the average score for π. This statistic is asymptotically normal under the null hypothesis. As noted by Deaton and Irish (1984), if the average score for π is negative and significantly different from zero, this is evidence against both Tobit and Cragg's first model.

The various test statistics are presented in Table 4. The hypothesis of a Tobit model is strongly rejected in all cases, except the WL form for clothing. The results for the WL and LES forms are generally consistent with those for QWL and QES. Nelson's test tends to reject less than the others. This may be a consequence of its low power, as suggested by Ruud (1984). The Ruud and the Lin-Schmidt statistics are very close and always lead to rejection. The Deaton-Irish statistic is always positive, which indicates rejection of the Tobit model, but not in the direction of Cragg's first model. The conditional symmetry hypothesis is almost always accepted. The exceptions are transport and the LES form for clothing.

Misspecification may also be detected by less formal procedures. Following Chesher, Lancaster and Irish (1985), we use the Tobit ML residuals to compute the Kaplan-Meier estimate of the distribution function of the errors. If the model is correctly specified the Kaplan-Meier estimate is consistent, and therefore the plot of its inverse normal transform against the ordered ML residuals should be close to a 45-degree line. In agreement with our earlier findings, there is some evidence for normality only in the case of the WL form for clothing [Figures 1a,b]. However, even in this case, the Kaplan-Meier estimate has somewhat fatter tails than normal. In the other cases, especially the expenditure equations, the Kaplan-Meier estimate looks often like the double-exponential distribution [see e.g. the case of the LES form for clothing in Figures 1c,d], which makes the CLAD estimator an interesting alternative to ML.

Thus, formal tests and graphical procedures all indicate that misspecification is likely to be present in most of the cases that we consider. However, it is hard to determine the exact nature of the misspecification, and in particular, whether it is due to failures of the Tobit specification or simply to failures of the normality assumption. To what extent is this going to affect estimation of the parameter of interest, namely the income elasticity of demand?

Here we compare the results obtained for 8 different estimators: the Tobit ML estimator, the 5 bounded-influence estimators discussed in Section 3, and Powell's CLAD and SCLS estimators. Details on the computation are given in the Appendix. We also computed one-step versions of the H-K (Hampel-Krasker type) and K-W (Krasker-Welsch type) estimators, starting at the BI0 estimator, but somewhat surprisingly they behaved rather poorly and therefore are not presented here.

Estimated standard errors for the ML estimator are consistent under heteroskedasticity and non-normality. The AVM for bounded-influence estimators was estimated in the two ways suggested in Section 3. The first estimate is consistent if we condition on the given set of regressors, which is not unreasonable in the present context. The second estimate is fully consistent but is difficult to compute, because of the need of differentiating numerically a very complicated score function. The computation is simplest in the BI1 case, because only the recentering vector a_0 has to be estimated. For this estimator we present unconditional t-ratios for the WL and LES case, along with the conditional ones. No clear pattern emerges, and the results appear to be very sensitive to the choice of the displacement for numerical differentiation and the order of quadrature for numerical integration. For this reason, in what follows we only report the conditional standard errors.

For the CLAD covariance estimates we consider different window widths for the non-parametric estimator of the error density at the origin. The estimates are very sensitive to the degree of smoothing. We report results for three choices, corresponding respectively to setting $c_0 = 0.5$, 1.0 and 2.0 in equation (5.5) of Powell (1984).

Table 6 presents estimates of the income elasticity of demand evaluated at the median income. First consider the WL and LES forms. In the case of clothing estimates do not change much between models and estimation techniques, and are rather precise. In the case of transport the differences between ML and all other estimates tend to be large. Semi-parametric and bounded-influence estimates are generally close (with the exception of the H-K estimator, perhaps because of the numerical problems encountered in this case), but bounded-influence estimates tend to be more precise. In the case of tobacco again we find large differences between ML and all other estimators, but all estimates are very unprecise.

In the QWL case, estimated elasticities are close to the ones for the WL and LES forms. In the QES case, however, ML estimates just blow up. For all goods ML estimates are at least twice as big as for the other specifications. The SCLS estimates are not reported because of divergence of the algorithm. The other estimates are also larger than for all previous specifications, but the increase is much less pronounced.

Our results are illustrated in Figure 2, that compares the shape of estimated Engel curves for transport in the case of a median household with 7 household members, 4 adults and 3 children, living in a rural area. The differences between ML and the other estimates are also very big in the case of tobacco, but are smaller in the case of clothing.

Specification tests based on the difference between ML and bounded-influence estimates satisfy the conditions for powerful tests [see e.g. Ruud (1984)], namely a large difference between estimators under the alternative model, and a relatively efficient alternative estimator to ML. On the other hand, tests based on the difference between the ML and the CLAD or SCLS estimator satisfy the first but not the second conditions and therefore should be less powerful.

The specification test statistics are presented in Table 7. The test statistics are computed as N times the uncentered R^2 in the regression of a column of ones on the likelihood score and the influence function for the subset of regression parameters, both evaluated at the ML estimates. Under the null hypothesis of correct specification these statistic have an asymptotic χ^2 distribution with the number of degrees of freedom equal to the number of regression parameters. We find this artificial regression form more convenient than Hausman (1978) original form, because the covariance matrix for the difference between the two contrasts, even when constrained to be psd, as suggested in Newey (1985), is typically singular.

Equality of the regression coefficients is typically rejected for clothing and transport, but not for tobacco. In the case of CLAD and SCLS rejection occurs less frequently, essentially because of the larger standard errors of these estimates. Interestingly, equality of the whole parameter vector is almost always rejected in the case of bounded-influence estimators. This partly reflects the fact that ML estimates of the scale parameter are usually larger and less precise than bounded-influence estimates.

Can we detect what are the influential data points in the sample? One possibility is to use diagnostic methods of the type proposed by Belsely, Kuh and Welsch (1980) and Cook and Weisberg (1982) for the linear model. These methods are all based on deleting a subset of observation at a time and then comparing the resulting estimates with the ones based on the full sample. A subset of observations is deemed to be influential if this difference is large. Usually only methods based on the deletion of single observations are applied, because of the combinatorial problems arising with multiple deletion. However, even single deletion methods may be quite expensive for non-linear estimators, in particular when the number of observations is high. Moreover, it is not clear what is the best way of summarizing all the information contained in the differences.

Another possibility is to examine, for a given estimator, the norm of the IF for each observation. An influential observation is one for which the norm of the IF is big. In large samples, this is equivalent to deleting one observation at a time and then computing the norm of the difference in the estimates with respect to the full sample [see e.g. Efron (1982)].

Neither method is entirely satisfactory for detecting outliers. On the one hand, it is well known that single deletion methods may fail to reveal the pattern of multiple outliers [see e.g. Atkinson (1986)]. On the other hand, outliers may not be detected by methods that are based on estimators that are not robust.

As an example, consider the relationship between log per-capita income and budget share on transport. To each point in the scatter we superimpose the norm of the IF of the Tobit ML estimator, the IF being evaluated at the ML estimates [Figure 3] Two very influential points are clearly revealed. However, the norm of the IF is not very large for the points in the cluster on the top-right of the scatter. These points correspond to households with an unusually high expenditure share on transport (15% or more). It is their presence that explains why ML estimates of elasticity are so large, especially in the QES case.

As an alternative to the previous methods we recommend using the weights from bounded-influence estimation. An influential observation is now one which receives a small weight. The weights are jointly computed with the robust parameter estimates and no additional calculation is required. They are easy to interpret given the weighted ML nature of bounded-influence estimators, and summarize all the information on the influence of one observation in a single scalar number in the interval [0,1]. Figure 4 is a graphical illustration of the use of the robust weights. The scatter of the log per capita income and the budget share on transport is presented again, this time with the robust weights superimposed. Notice that the two very influential observations of Figure 3 are heavily downweighted, but so are now the points in the cluster on the top-right of the scatter. The various estimators differ in the weight assigned to each of these observations. The H-K estimator downweights more the former set of points, while the other bounded-influence estimators downweight more the latter.

5. Conclusions

This study shows that estimated Engel curves may differ significantly depending on the choice of estimation technique. This does not surprise, given the systematic rejection of the assumptions of normal errors and Tobit specification. What may surprise, however, is the extent to which ML may differ from the other estimates as an effect of a few extreme observations.

We found that semi-parametric and bounded-influence estimates tend to be close to each other, but the latter appear to be more precise and lead to tests that appear to be more powerful. It would be interesting to verify these indications with a full scale Monte Carlo study.

This study also demonstrates the feasibility of bounded-influence estimation outside the context of the linear regression model. In our view, using bounded-influence estimators in applied work offers several advantages. First, it ensures protection against influential observations and local failures of parametric assumptions. Second, the difference with respect to ML estimates provides the basis for specification tests that have power against a variety of alternatives. Third, as diagnostics for outliers and influential observations, robust weights appear to be do better than the usual data deletion procedures. The price one has to pay by using these estimators is a loss of efficiency with respect to ML if the assumed model is indeed correct. However, and this is yet another advantage, the investigator can choose the efficiency loss that he/she is willing to tolerate.

Appendix

The computation of bounded-influence estimates proceeds as follows:

1. Start with $\theta^{(0)} = \hat{\theta}_{ML}$ and $a^{(0)} = 0$.

2. Choose $A^{(1)} = I_p$ for the BI0 and BI1 estimators, $A^{(1)} = J(\theta^{(0)})^{-1/2}$ for the BI2 estimator, $A^{(1)} = P(\theta^{(0)})^{-1}$ for the H-K estimator and $A^{(1)} = Q(\theta^{(0)})^{-1/2}$ for the K-W estimator [$P(\theta^{(0)})$ is the solution to (11) for a given $a^{(0)}$, and $Q(\theta)$ is defined in Proposition 1].

3. Given $A^{(1)}$, compute $a^{(1)}$ as

$$a^{(1)} = \left[\Sigma_n E_\Phi \min\left\{ 1, \frac{c}{\|A^{(1)}[s_n(r,\theta^{(0)}) - b^{(0)}]\|} \right\} \right]^{-1}$$

$$\times \Sigma_n E_\Phi \min\left\{ 1, \frac{c}{\|A^{(1)}[s_n(r,\theta^{(0)}) - b^{(0)}]\|} \right\} s_n(r,\hat{\theta}^{(0)})$$

where $b^{(0)}$ is equal to 0 for the BI1 estimator and is equal to $a^{(0)}$ otherwise. The normal integrals are evaluated numerically by using the Gauss-Legendre subroutine in Quandt (1986).

4. Given $A^{(1)}$, $a^{(1)}$ and $b^{(1)}$, compute $\theta^{(1)}$ by solving

$$\Sigma_n \min\left\{ 1, \frac{c}{\|A^{(1)}[s(z_n,\theta) - b^{(1)}]\|} \right\} [s(z_n,\theta) - a^{(1)}] = 0.$$

This is done by using the Newton-Raphson algorithm NEWRAP in GQOPT.

5. Given $\theta^{(1)}$, compute $A^{(2)}$, $a^{(2)}$, $b^{(2)}$and $\theta^{(2)}$ as in Step 2 to 4, and iterate. Convergence of this algorithm is not guaranteed.

The sensitivity bound c is chosen so as to obtain an average weight of about 95%. When $c = \infty$ all bounded-influence estimators that we consider are the same as the ML estimator, with an average weight equal to unity. Thus, our choice of the sensitivity bound may be interpreted as resulting in an efficiency loss of about 5% when the Tobit model is indeed correct. The % of downweighted observations varies depending on the specification and, to a lesser extent, the type of estimator. Typically is between 10 and 15% for the WL and QWL forms, and is somewhat lower for LES and QES. In the latter case, however, the value of the minimum weight is much smaller, which indicates the presence of highly influential observations.

The convergence criterion requires the maximal change in any of the parameter estimates to be less than 10^{-4}. Convergence is typically attained after 5 to 10 iterations of the outer loop. We had numerical problems with the H-K estimator, in particular for the QWL and QES specifications, and we do not report results for these two cases. For the other bounded-influence estimators, sometimes the algorithm cycled between two values very close to each other. In these cases convergence was always reached by weakening the tolerance to 10^{-3}.

The CLAD estimates are computed by iteratively reweighted LS with weight function given by $w(y,x,\beta) = 1(x'\beta) \min\{|y - x'\beta|^{-1}, \varepsilon^{-1}\}$, where ε is positive and small. The SCLS estimates are computed by the iterative LS algorithm mentioned in Powell (1986). The convergence criterion requires the maximal change in any of the parameter estimates to be less than 10^{-5}. CLAD estimates typically need more iterations to converge. In a few cases the limit of 100 iterations was reached without convergence.

REFERENCES

Amemiya, T. (1985), *Advanced Econometrics*, Harvard University Press, Cambridge, Mass.

Arabmazar and Schmidt (1981) "Further evidence on the robustness of the Tobit estimator to heteroskedasticity", *Journal of Econometrics*, 17, 253-258.

Arabmazar and Schmidt (1982), "An investigation of the robustness of the Tobit estimator to non-normality", *Econometrica*, 50, 1055-1063.

Atkinson, A.C. (1986), "Masking unmasked", *Biometrika*, 73, 533-541.

Belsley, D.A., Kuh, E. and Welsch, R.E. (1980), *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*, Wiley, New York.

Bera, A.K., Jarque, C.M. and Lee, L.F. (1984), "Testing for the normality assumption in limited dependent variable models", *International Economic Review*, 25, 563-578.

Bickel, P.J. (1975), "One-step Huber estimates in the linear model", *Journal of the American Statistical Association*, 70, 428-434.

Chesher, A., Lancaster, T., and Irish, M. (1985), "On detecting the failure of distributional assumptions", Discussion Paper 85/165, Department of Economics, University of Bristol.

Cragg (1972), "Some statistical models for limited dependent variables with applications to the demand for durable goods", *Econometrica*, 39, 829-844.

Deaton, A. and Irish, M. (1984), "Statistical models for zero expenditures in houseld budgets", *Journal of Public Economics*, 23, 59-80.

Efron, B. (1982), *The Jackknife, the Bootstrap and Other Resampling Plans*, SIAM, Philadelphia.

Goldberger, A. (1983), "Abnormal selection bias", in S. Karlin, T. Amemiya, and L.A. Goodman (eds.), *Studies in Econometrics, Time*

Series and Multivariate Statistics, Academic Press, New York.

Gorman (1980), "Some Engel curves", in A.S. Deaton (ed.), *Essays in Theory and Measurement of Consumer Behavior*, Cambridge University Press, Cambridge, U.K.

Hampel, F.R. (1974), "The influence curve and its role in robust estimation", *Journal of the American Statistical Association*, 69, 383-393.

Hampel, F.R. (1978), "Optimally bounding the gross-error- sensitivity and the influence of position in factor space", *Proceedings of the ASA Statistical Computing Section*, ASA, Washington, D.C.

Hampel, F.R., Ronchetti, E., Rousseeuw, P.J., and Stahel, W.A. (1986), *Robust Statistics: The Approach Based on Influence Functions*, Wiley, New York.

Hausman, J.A. (1978), "Specification tests in econometrics", *Econometrica*, 46, 1251-1272.

Huber, P.J. (1967), "The behavior of maximum likelihood estimates under non-standard conditions", *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, Vol. 1, 221-33.

Hurd, M. (1979), "Estimation in truncated samples when there is heteroskedasticity", *Journal of Econometrics*, 11, 247-258.

Kiefer, N. and Skoog, G. (1984), "Local asymptotic specification error analysis", *Econometrica*, 873-885.

Krasker, W.S. (1980), "Estimation in linear regression models with disparate data points", *Econometrica*, 48, 1333-1346.

Krasker, W.S. and Welsch, R.E. (1982), "Efficient bounded-influence regression estimation", *Journal of the American Statistical Association*, 77, 595-604.

Lin, T.F. and Schmidt, P. (1984), "A test of the Tobit specification against an alternative suggested by Cragg", *Review of Economics and Statistics*, 66, 174-177.

Nelson, F.D. (1981), "A test for misspecification in the censored normal model", *Econometrica*, 49, 1317-1329.

Newey, W.K. (1985), "Maximum likelihood specification testing and conditional moment tests", *Econometrica*, 53, 1047-1069.

Newey, W.K. (1987), "Specification tests for distributional assumptions in the Tobit model", *Journal of Econometrics*, 34, 125-145.

Paarsch, H.J. (1984), A Monte Carlo comparison of estimators for censored regression models", *Journal of Econometrics*, 24, 197-213.

Peracchi, F. (1987), Bounded Influence Methods in Econometrics with an Application to the Censored Regression Model, Ph.D. Thesis, Princeton University.

Powell, J.L. (1984), "Least absolute deviations estimation for the censored regression model", *Journal of Econometrics*, 25, 303-325.

Powell, J.L. (1986), "Symmetrically trimmed least squares estimation for Tobit models", *Econometrica*, 54, 1435-1460.

Quandt, R.E. (1986), "Econometrics of disequilibrium", Princeton University, mimeo.

Ruud, P.A. (1984), "Tests of specification in econometrics", *Econometric Reviews*, 3, 211-242.

White, H. (1982), "Maximum-likelihood estimation of misspecified models", *Econometrica*, 48, 817-838.

Table 1

Normal likelihood score ($r = \gamma y - x'\alpha$).

1. Truncated regression model

$$s_\alpha(z,\theta) = 1(y > 0) \left[r - \frac{\phi(x'\alpha)}{\Phi(x'\alpha)} \right] x$$

$$s_\gamma(z,\theta) = 1(y > 0) \left[\frac{1}{\gamma} - ry \right] .$$

2. Censored regression model

$$s_\alpha(z,\theta) = \left[1(y > 0)\ r - 1(y = 0) \frac{\phi(x'\alpha)}{1 - \Phi(x'\alpha)} \right] x$$

$$s_\gamma(z,\theta) = 1(y > 0) \left[\frac{1}{\gamma} - ry \right] .$$

3. Cragg's 1st model

$$s_\alpha(z,\theta) = \left[1(y > 0)\ r - 1(y = 0) \frac{\pi\ \phi(x'\alpha)}{1 - \pi\ \Phi(x'\alpha)} \right] x$$

$$s_\pi(z,\theta) = 1(y > 0) \frac{1}{\pi} - 1(y = 0) \frac{\Phi(x'\alpha)}{1 - \pi\ \Phi(x'\alpha)}$$

$$s_\gamma(z,\theta) = 1(y > 0) \left[\frac{1}{\gamma} - ry \right] .$$

4. Cragg's 2nd model

$$s_\alpha(z,\theta) = 1(y > 0) \left[r - \frac{\phi(x'\alpha)}{\Phi(x'\alpha)} \right] x$$

$$s_\pi(z,\theta) = 1(y > 0) \frac{1}{\pi} - 1(y = 0) \frac{1}{1 - \pi}$$

$$s_\gamma(z,\theta) = 1(y > 0) \left[\frac{1}{\gamma} - ry \right] .$$

Table 2

Bounded-influence estimators for the censored regression model.

All bounded-influence estimators that we consider are based on a score function of the form

$$\eta(z,\theta) = \min\left\{ 1, \frac{c}{\|A(\theta)\ [s(z,\theta) - b(\theta)]\|} \right\} [s(z,\theta) - a(\theta)]$$

where

Estimator	$b(\theta)$	$A(\theta)$	Metric on the IF
BI0	0	I_p	-
BI1	$a(\theta)$	I_p	$P(\theta)^2$
BI2	$a(\theta)$	$J(\theta)^{-1/2}$	$P(\theta)\ J(\theta)^{-1}\ P(\theta)$
H-K (Hampel-Krasker)	$a(\theta)$	$P(\theta)^{-1}$	I_p
K-W (Krasker-Welsch)	$a(\theta)$	$Q(\theta)^{-1/2}$	$P(\theta)\ Q(\theta)^{-1}\ P(\theta)$

Note: The p×1 vector $a(\theta)$ and the p×p matrix $P(\theta)$ are solutions to equations (11)-(12) in the text, and $Q(\theta) = E_\theta\ \eta(Z,\theta)\ \eta(Z,\theta)'$.

Table 3

Definition and summary statistics for the variables in the data set.

SHXCLOTH: % share of total exp. on clothing and footwear.
SHXTRANS: % share of total exp. on transport services and repairs.
SHXTOBAC: % share of total exp. on tobacco products.
XCLOTH : household expenditure on clothing and footwear.
XTRANS : household expenditure on transport services and repairs.
XTOBAC : household expenditure on tobacco products.
LXPC : log of total expenditure per household member.
LXPCSQ : square of LXPC.
XPC : total expenditure per household member.
XPCSQ : square of XPC.
HHSIZE : number of household members.
LT14 : household members less than 14 years old.

Variable	Min	Max	Median	MAD (1)
SHXCLOTH	0.00	22.35	4.0	92.24
SHXTRANS	0.00	30.43	0.97	0.97
SHXTOBAC	0.00	12.27	0.52	0.52
XCLOTH	0.00	52.63	4.00	2.73
XTRANS	0.00	73.53	0.94	0.94
XTOBAC	0.00	19.44	0.34	0.34
LXPC	7.97	11.68	9.41	0.31
LXPCSQ	63.45	136.48	88.54	5.77
XPC	2.88	118.46	12.20	3.73
XPCSQ	8.29	14032.43	148.89	84.30
HHSIZE	2	16	7	2
LT14	0	9	3	1

(1) Median absolute deviation from the median.

Table 4

Tests for normality, symmetry and Tobit specification.

	Ruud	Nelson	BJL	L-S	D-I	Newey
Clothing						
WL	14.1b	4.01a	4.22a	14.1b	1.68a	6.88a
LES	93.2	71.5	36.6	94.6	4.06	19.6
QWL	15.9b	7.21a	4.78a	16.0b	1.69a	8.73a
QES	99.0	127.0	20.2	100.6	0.95a	9.14a
Transport						
WL	110.3	40.6	60.8	111.8	10.2	16.4
LES	161.5	80.7	37.8	162.5	12.7	12.0b
QWL	118.9	61.0	59.6	119.4	10.6	15.5
QES	171.5	184.3	70.0	178.0	11.9	c
Tobacco						
WL	48.9	11.1b	45.8	49.0	6.74	6.39a
LES	70.7	23.8	61.5	71.1	7.22	6.84a
QWL	50.2	17.3	46.7	50.2	6.80	3.94a
QES	75.5	82.0	54.5	77.4	6.37	c

a: Asymptotic p-value greater then .05.

b: Asymptotic p-value between .01 and .05.

c: Not available.

Table 5

T-ratios of the BI1 estimates. A: Conditional on the given set of regressors. B: Unconditional, 20-point Gauss-Legendre integration. C: Unconditional, 40-point Gauss-Legendre integration.

	Clothing		Transport		Tobacco	
	WL	LES	WL	LES	WL	LES
CNST						
A	3.41	5.47	4.92	5.77	2.47	.207
B	.589	10.3	11.8	.956	.420	.127
C	3.53	1.45	.674	22.0	.436	.767
INCOME[a]						
A	4.23	7.75	5.29	6.23	2.12	.425
B	.719	13.8	12.7	.966	.364	.231
C	8.10	2.19	.754	25.0	.369	.624
HHSIZE						
A	1.50	4.99	2.93	5.42	1.11	2.90
B	1.07	17.6	5.75	3.80	2.52	6.22
C	1.47	.844	2.77	11.3	2.15	6.13
LT14						
A	1.24	.055	.374	.118	1.91	1.95
B	1.37	.184	.859	.107	1.16	7.63
C	.437	.028	.094	.207	4.76	2.74
DRUR						
A	1.03	.222	2.41	2.29	1.80	2.40
B	1.21	.656	5.45	2.00	2.43	6.86
C	.133	.051	.424	1.93	.486	2.74

a: LXPC for WL and XPC for LES.

Table 6

Income elasticity of demand evaluated at the median
(standard errors in parentheses).

	Clothing				Transport				Tobacco			
	WL	LES	QWL	QES	WL	LES	QWL	QES	WL	LES	QWL	QES
Tobit	1.61	1.23	1.71	2.64	5.85	6.73	5.70	13.24	.620	1.12	.593	3.27
	(.122)	(.475)	(2.48)	(.362)	(.745)	(2.67)	(26.7)	(1.77)	(.660)	(.815)	(13.4)	(1.67)
BI0	1.54	1.60	1.62	2.08	3.10	3.63	3.46	4.60	-.025	.253	-.073	1.42
	(.127)	(.207)	(2.18)	(.186)	(.383)	(.582)	(21.2)	(.377)	(.484)	(.581)	(10.2)	(2.16)
BI1	1.54	1.60	1.62	2.08	3.09	3.63	3.47	4.29	-.027	.254	-.079	1.42
	(.128)	(.207)	(2.31)	(.186)	(.394)	(.582)	(22.8)	(.379)	(.484)	(.583)	(10.2)	(2.15)
BI2	1.58	1.62	1.62	2.00	2.97	2.60	3.05	3.09	-.022	-.026	-.119	.276
	(.117)	(.199)	(3.33)	(.486)	(.317)	(.394)	(12.5)	(1.53)	(.471)	(.267)	(9.80)	(1.41)
H-K	1.63	1.54	a	a	4.51	6.99	a	a	.375	.817	a	a
	(.118)	(.318)			(.569)	(2.95)			(.544)	(.877)		
K-W	1.59	1.71	1.63	2.09	3.20	3.30	3.21	4.19	.031	.314	.022	1.59
	(.119)	(.186)	(3.70)	(.763)	(.347)	(.459)	(12.7)	(1.09)	(.495)	(.654)	(10.5)	(2.14)
SCLS	1.55	.948	1.66	3.25	3.68	2.38	3.21	b	-1.42	-.084	-3.38	b
	(.127)	(.368)	(.124)	(.531)	(.895)	(1.40)	(.561)		(.853)	(1.30)	(3.05)	
CLADc	1.61	1.39	1.61	2.11	3.13	2.97	3.14	4.69	.170	0.010	.245	.655
	(.082)	(.710)	(.099)	(.319)	(.532)	(2.41)	(2.17)	(.915)	(.716)	(1.52)	(.741)	(.690)
	(.115)	(.227)	(.109)	(.484)	(.420)	(3.95)	(.429)	(.943)	(.628)	(2.01)	(.504)	(.982)
	(.113)	(.170)	(.114)	(.573)	(.502)	(1.85)	(.645)	(1.42)	(.794)	(.309)	(.667)	(.942)

a: Not computed.

b: Algorithm failed to converge.

c: Standard errors corresponding respectively to c_0 = 0.5, 1.0 and 2.0 in equation (5.5) of Powell (1984).

Table 7

Specification tests on the regression parameters.

	BI0	BI1	BI2	H-K	K-W	SCLS	CLAD
Clothing							
WL	20.0	19.3	18.3	3.82a	15.8	2.02a	1.55a
LES	98.8	98.9	111.9	19.5	105.1	7.14a	19.8
QWL	7.86a	7.81a	20.4	c	16.7	1.03a	3.64a
QES	30.1	30.1	56.4	c	44.7	27.7	18.9
Transport							
WL	253.9	257.0	228.4	24.5	123.1	23.6	20.7
LES	171.2	175.4	164.5	100.0	86.2	38.4	13.9b
QWL	101.3	120.1	203.6	c	43.5	32.3	21.9
QES	16.2b	17.0	56.6	c	13.0b	54.4	32.1
Tobacco							
WL	7.19a	7.51a	8.98a	16.8	13.0b	16.9	15.2
LES	7.58a	7.69a	2.61a	24.8	8.13a	26.7	30.8
QWL	19.4	19.8	14.3b	c	12.0a	18.7	15.6b
QES	20.0	20.2	1.02a	c	7.90a	27.7	21.8

a: Asymptotic p-value greater than .05.

b: Asymptotic p-value between .01 and .05.

c: Not available.

FIGURE 1a

DEP VBL = SHXCLOTH, WORKING-LESER FORM
KAPLAN-MEIER ESTIMATE OF THE DF OF THE STANDARDIZED RESIDUALS

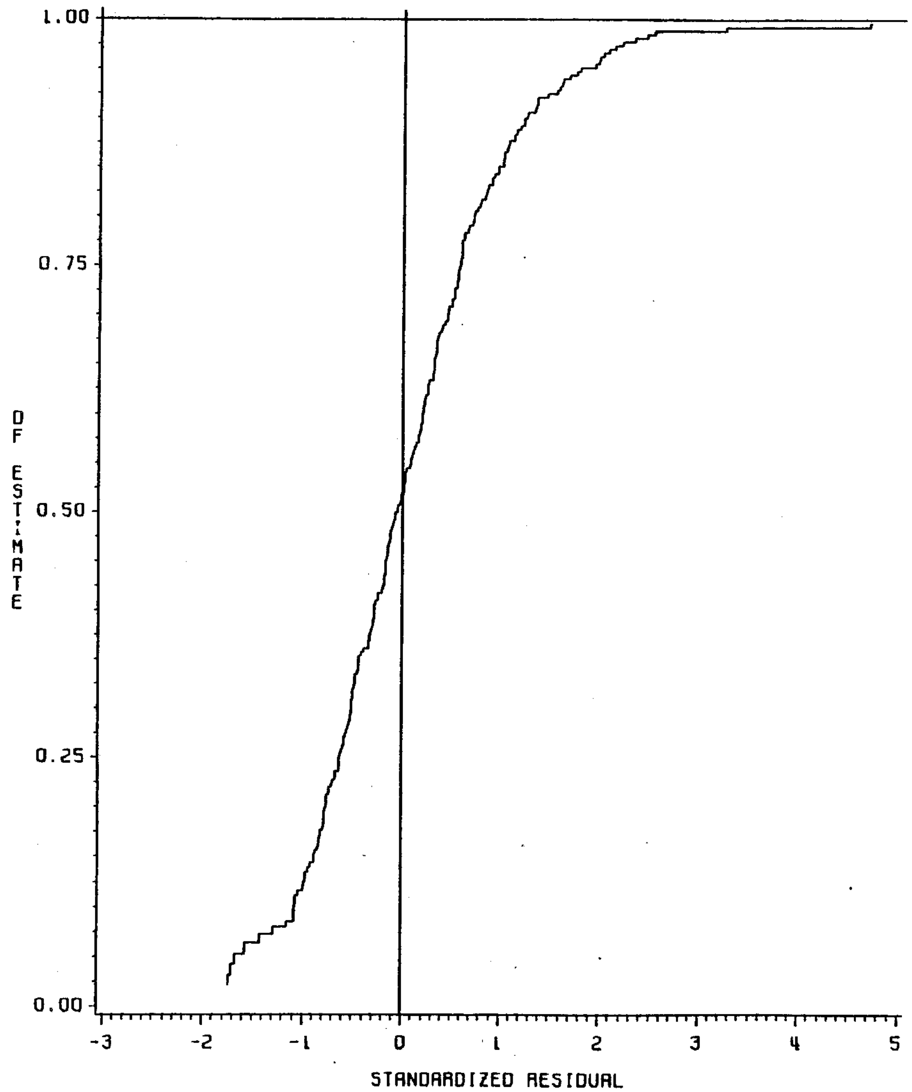


FIGURE 1b

DEP VBL = SHXCLOTH, WORKING-LESER FORM
INVERSE NORMAL TRANSFORMATION OF THE KAPLAN-MEIER ESTIMATE

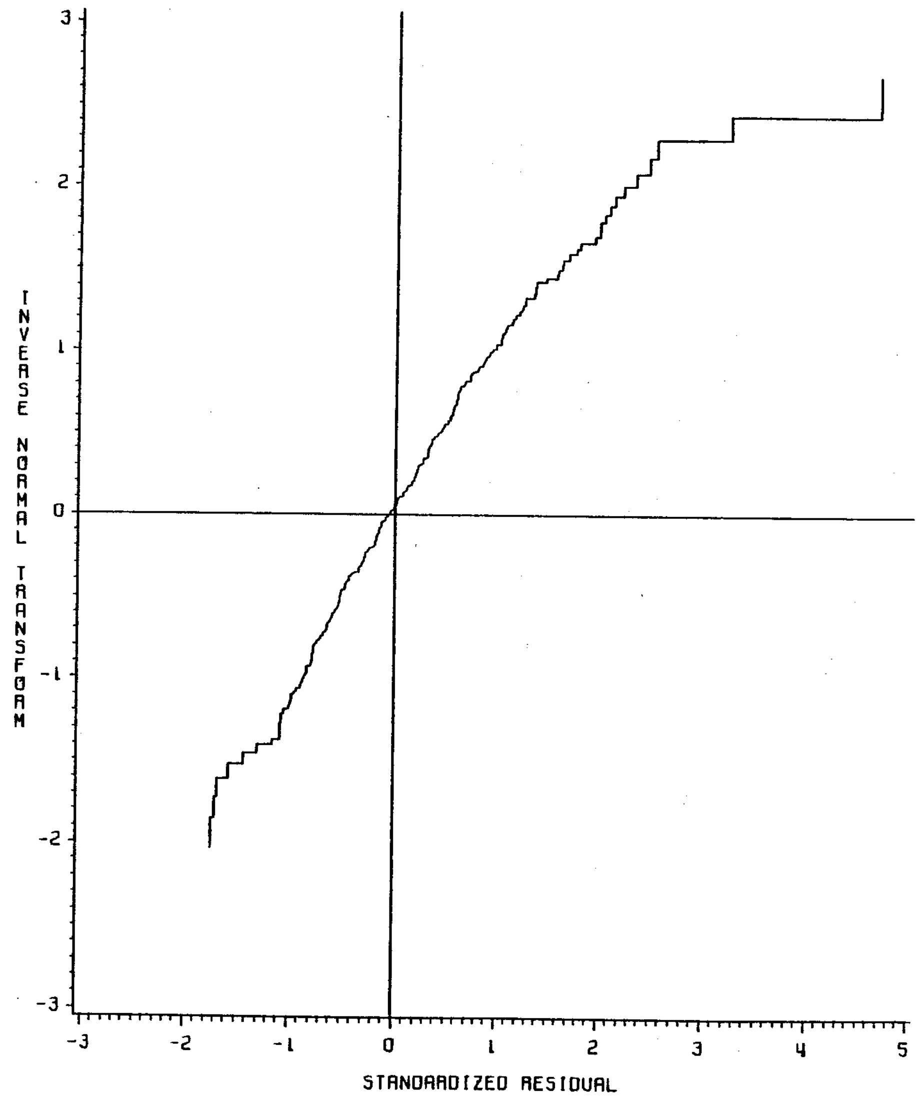


FIGURE 1c

DEP VBL = XCLOTH, LINEAR EXPENDITURE SYSTEM
KAPLAN-MEIER ESTIMATE OF THE DF OF THE STANDARDIZED RESIDUALS

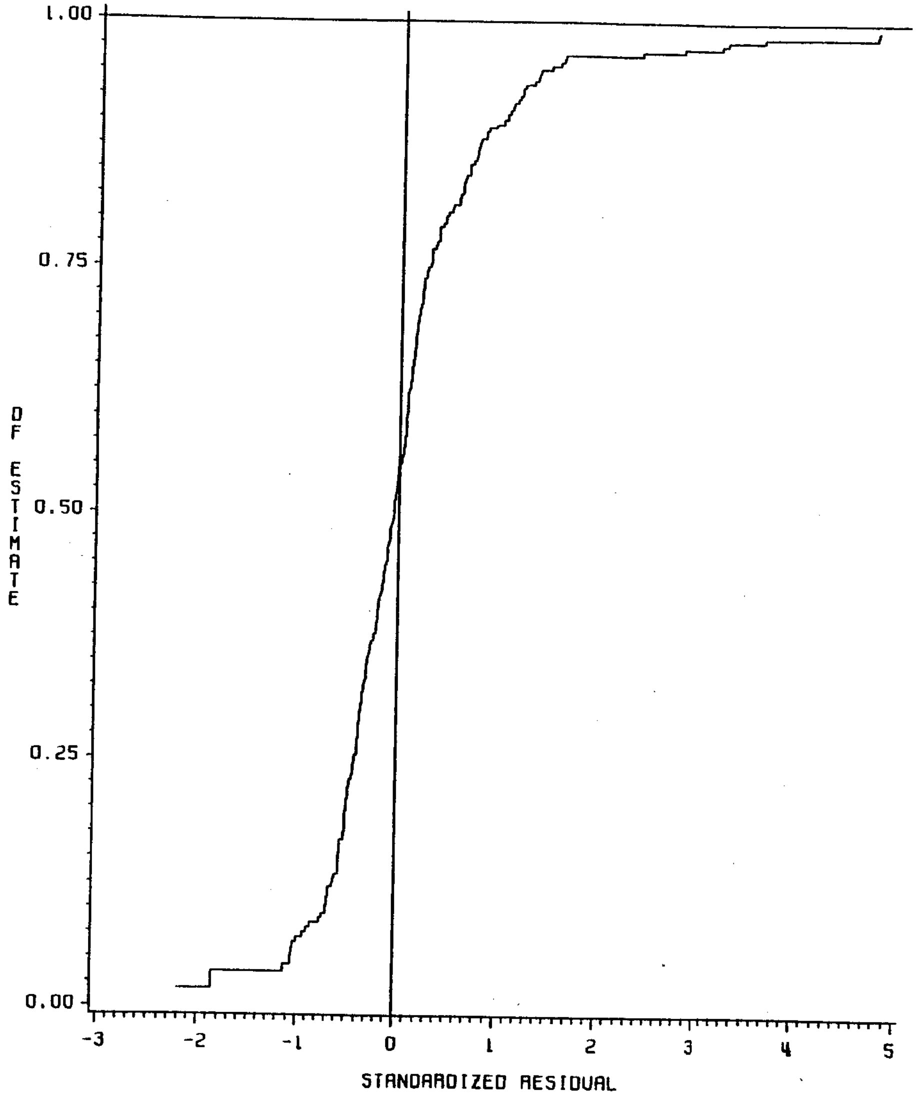


FIGURE 1d

DEP VBL = XCLOTH, LINEAR EXPENDITURE SYSTEM
INVERSE NORMAL TRANSFORMATION OF THE KAPLAN-MEIER ESTIMATE

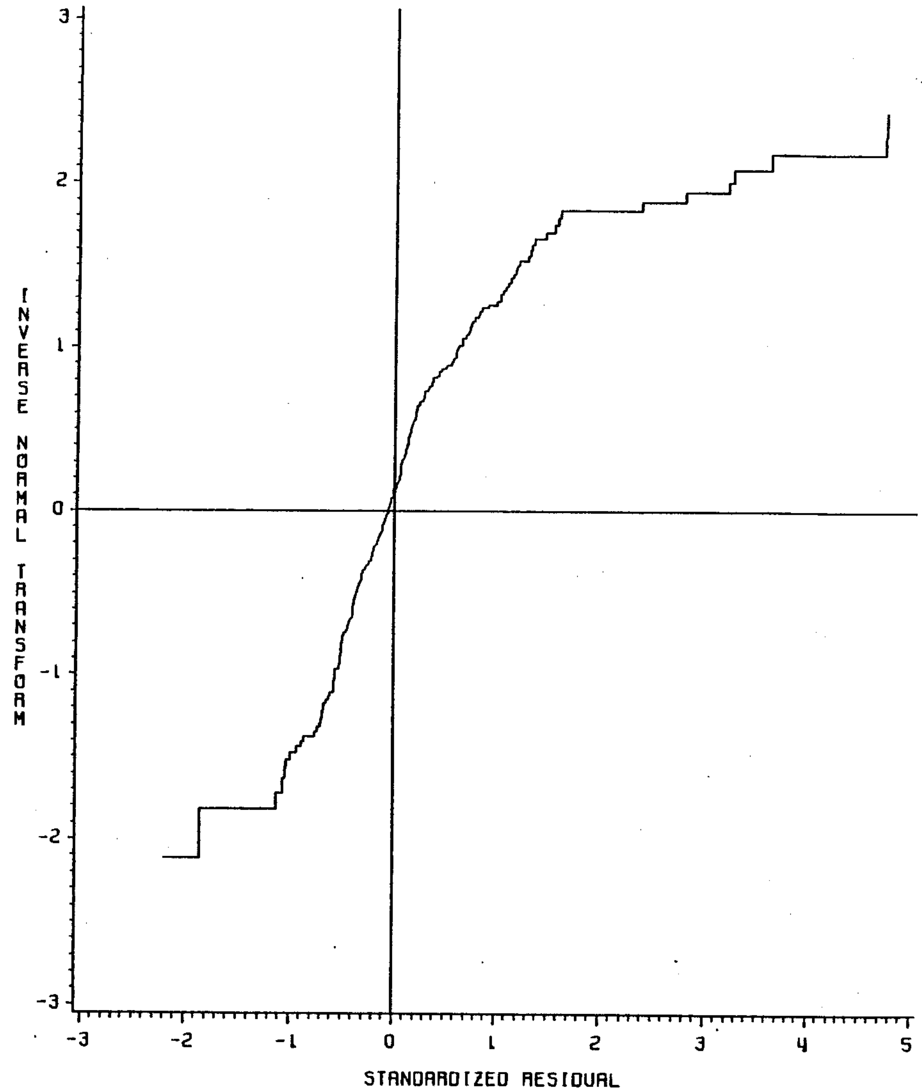


FIGURE 2a

ESTIMATED ENGEL CURVES FOR TRANSPORT
WORKING-LESER FORM
LEGEND: ML=1, BI1=2, BI2=3, LAD=4

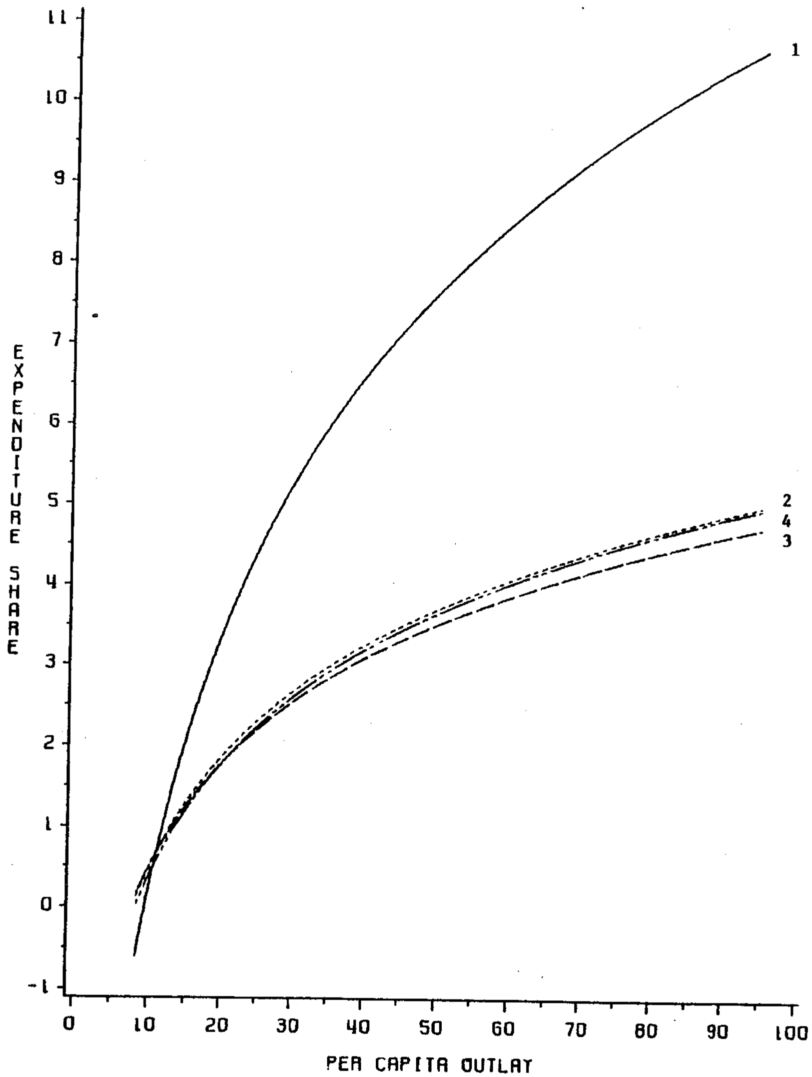


FIGURE 2b

ESTIMATED ENGEL CURVES FOR TRANSPORT
LINEAR EXPENDITURE SYSTEM
LEGEND: ML=1, BI1=2, BI2=3, LAD=4

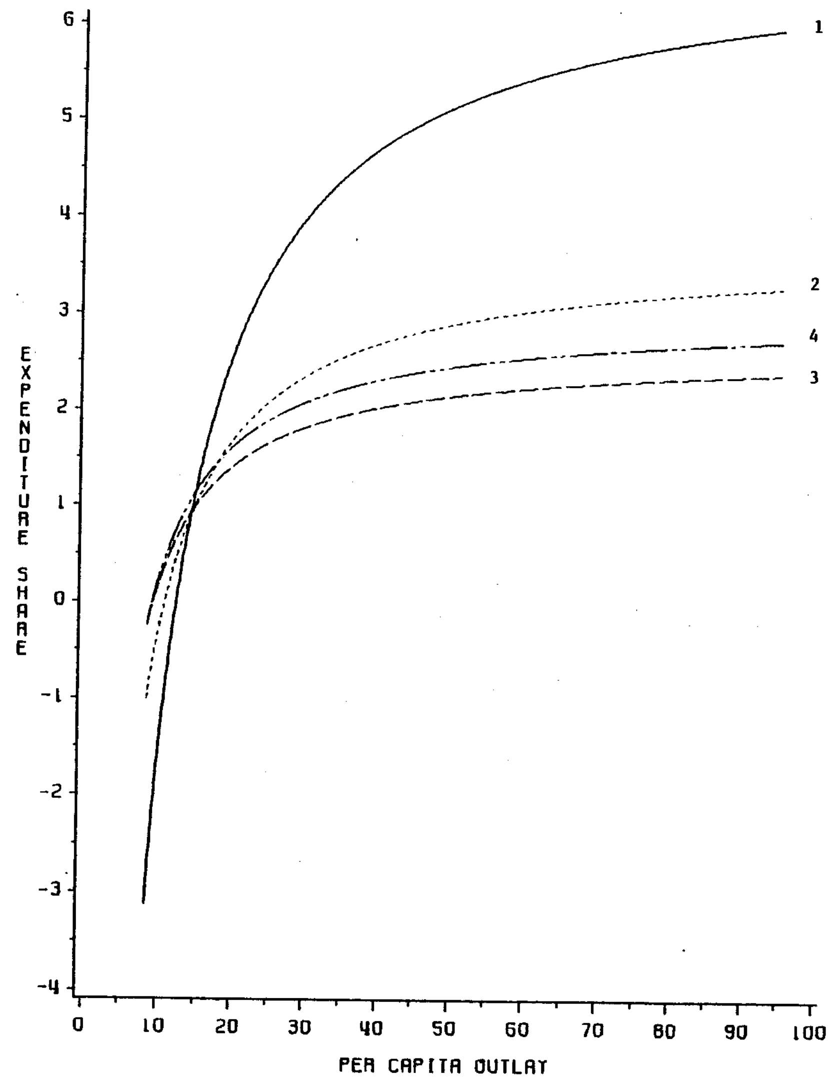


FIGURE 2c

ESTIMATED ENGEL CURVES FOR TRANSPORT
QUADRATIC WORKING-LESER FORM
LEGEND: ML=1, BI1=2, BI2=3, LAD=4

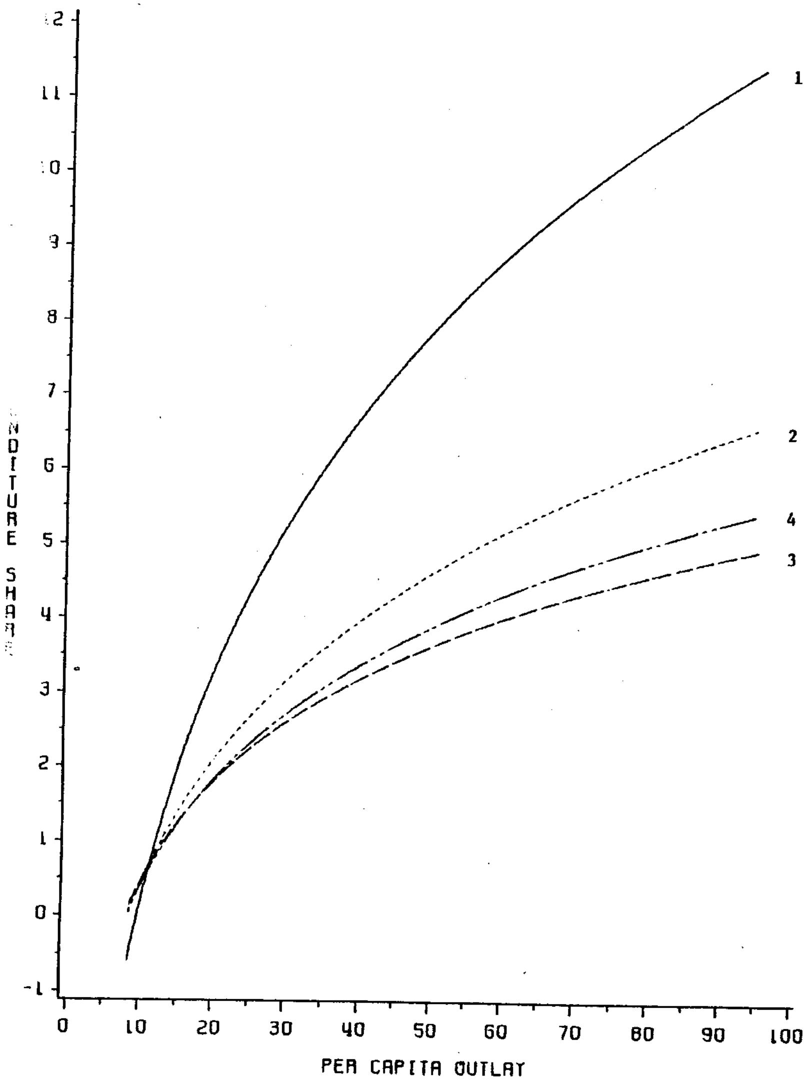


FIGURE 2d

ESTIMATED ENGEL CURVES FOR TRANSPORT
QUADRATIC EXPENDITURE SYSTEM
LEGEND: ML=1, BI1=2, BI2=3, LAD=4

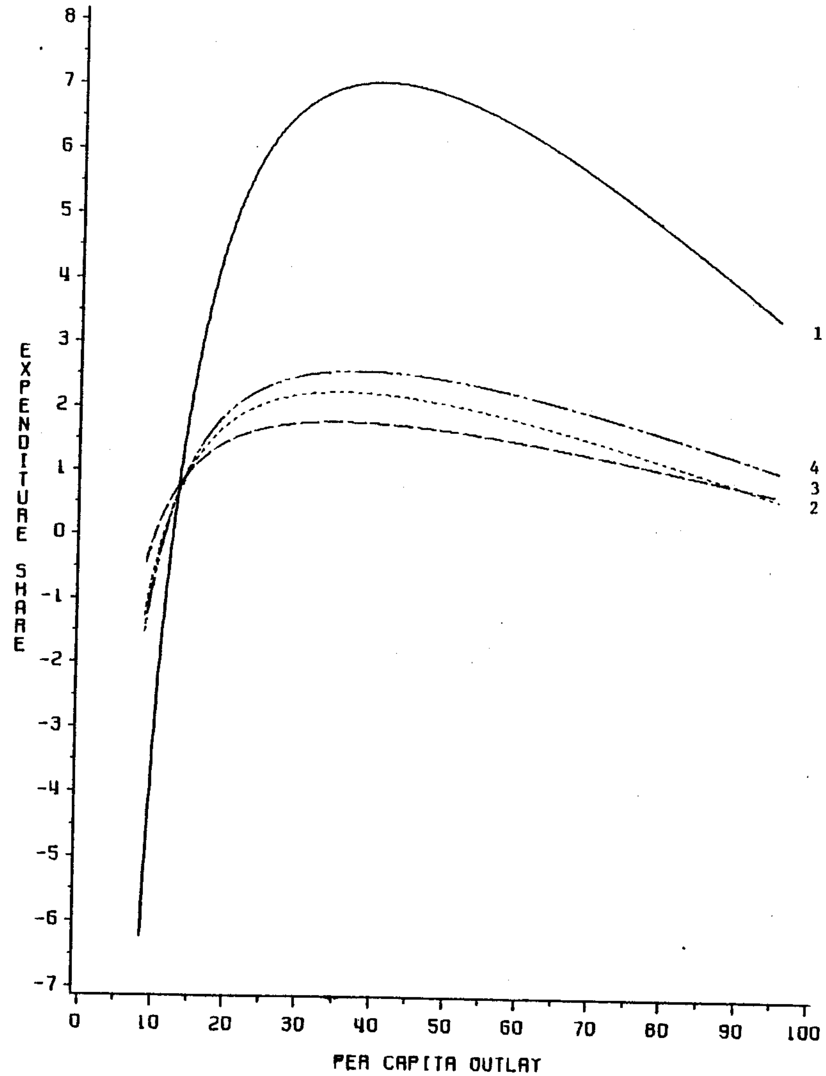
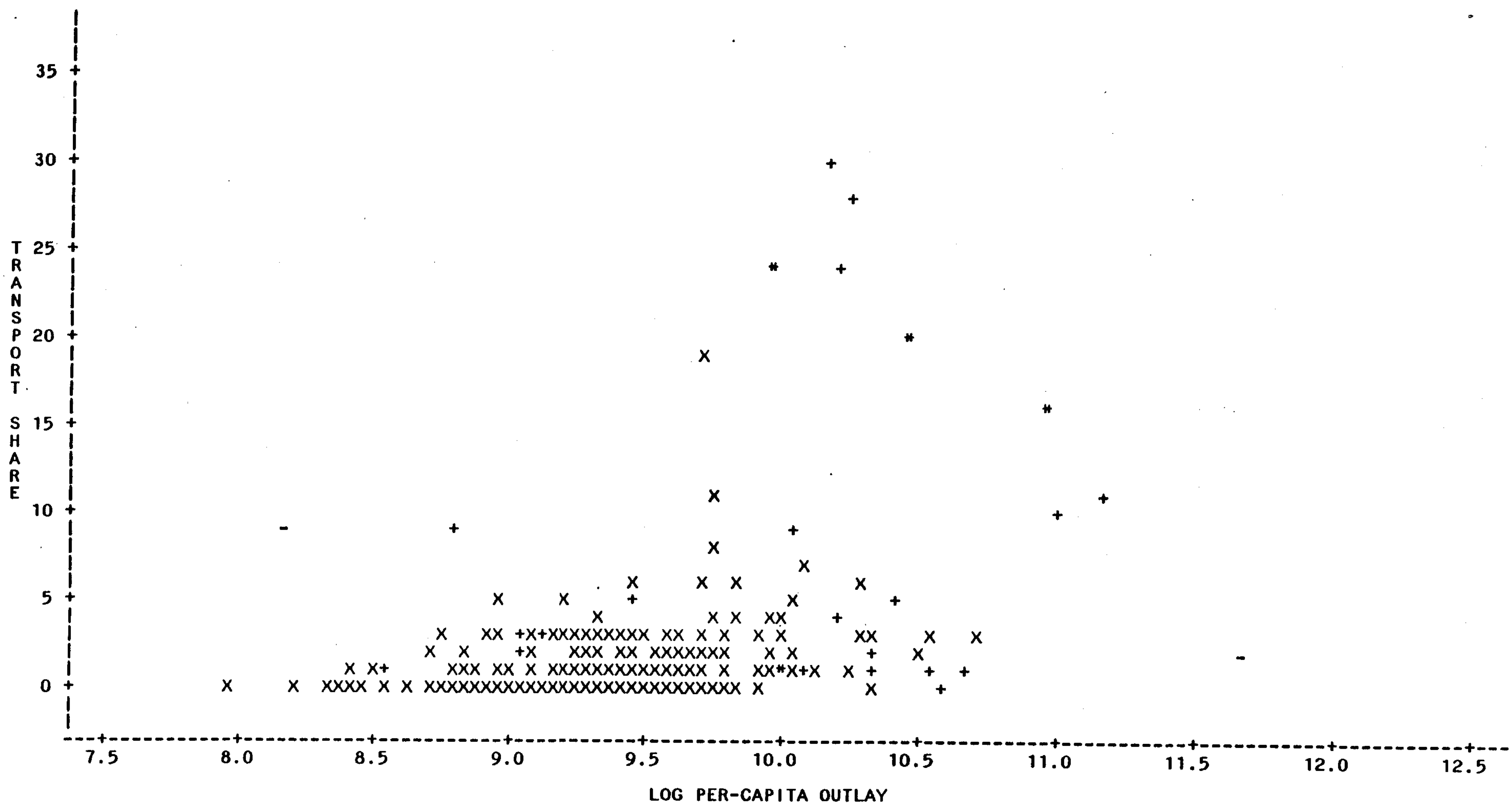


Figure 3

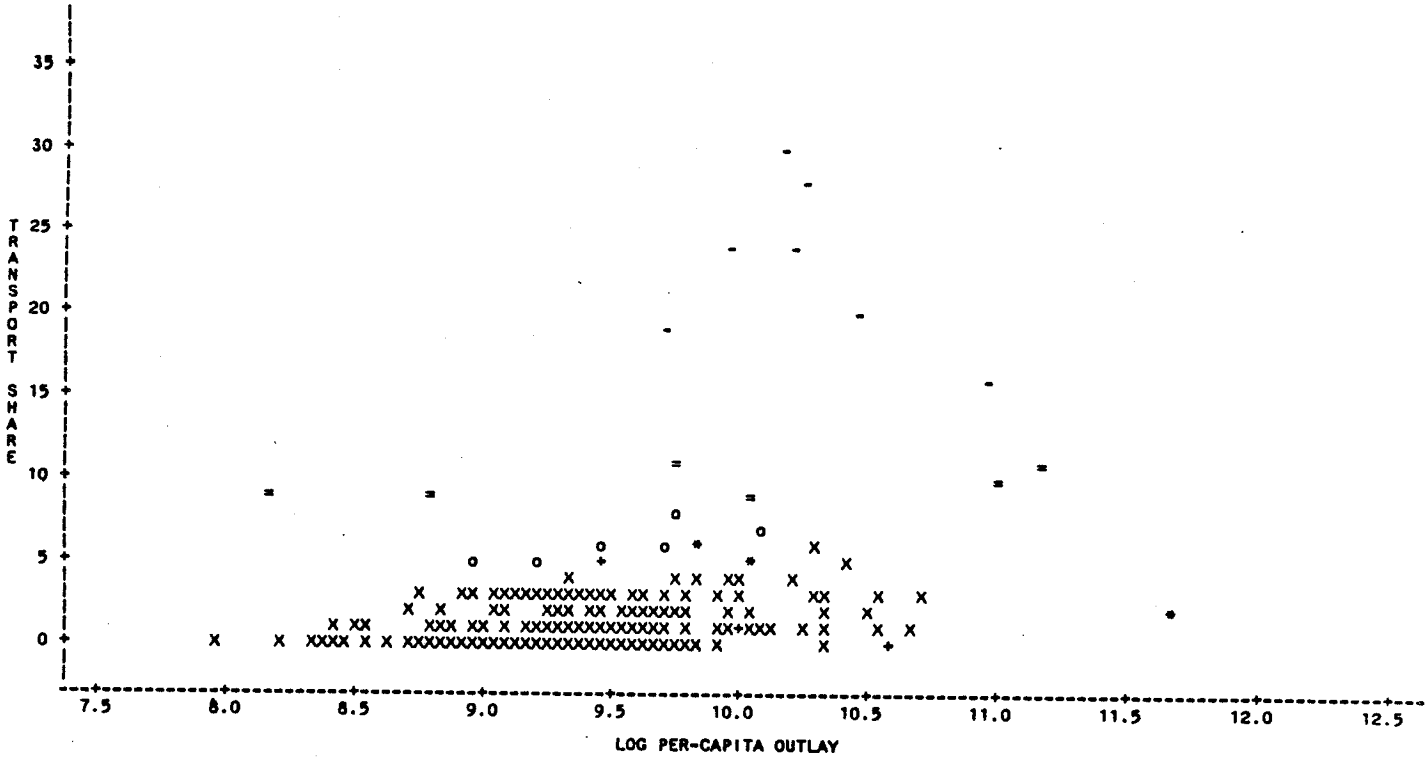
PLOT OF TRANSPORT SHARE VS. PER CAPITA OUTLAY
NORM OF TOBIT-ML INFLUENCE FUNCTION SUPERIMPOSED. WORKING-LESER FORM



SYMBOL	TNORM	SYMBOL	TNORM	SYMBOL	TNORM
XXXXXX	0.52839 - 25.35586	******	75.01079 - 124.66573	======	174.32067 - 223.97560
++++++	25.35586 - 75.01079	oooooo	124.66573 - 174.32067	------	223.97560 - 248.80307

Figure 4a

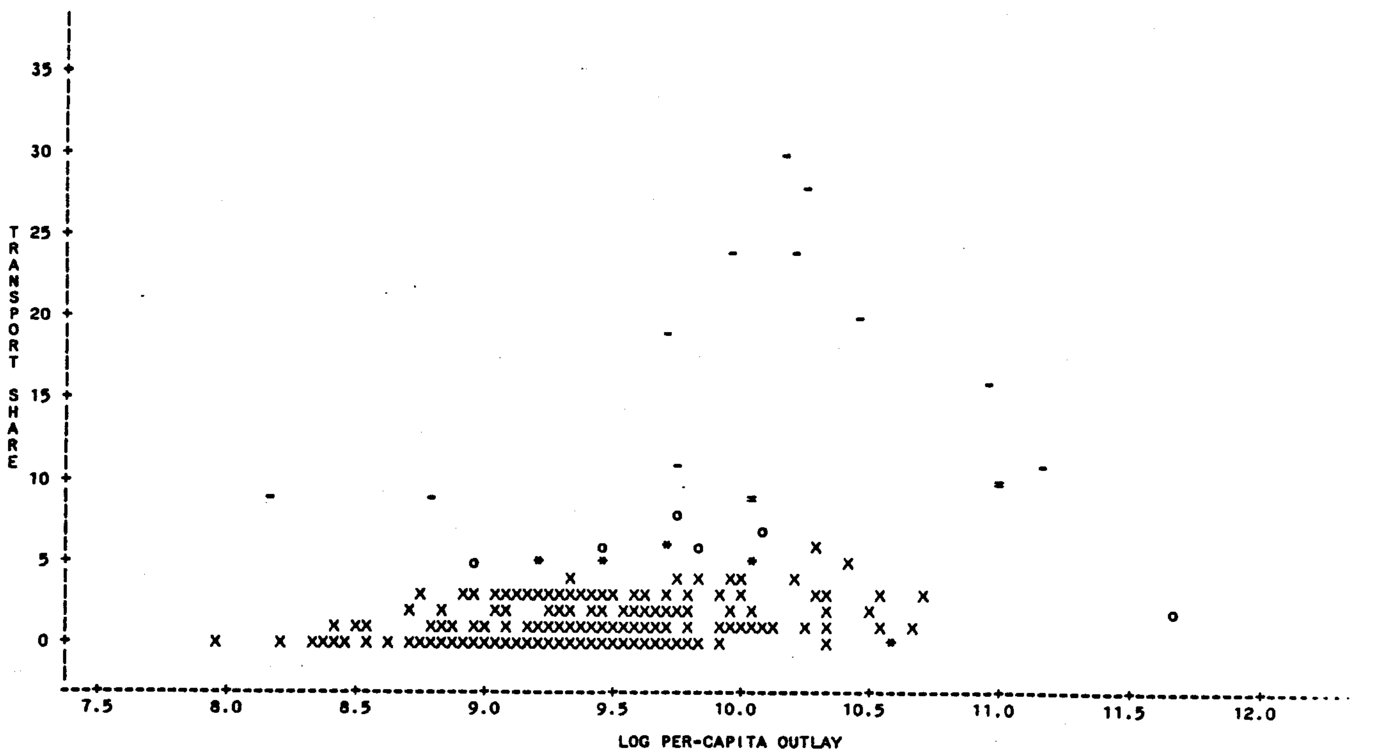
PLOT OF TRANSPORT SHARE VS. LOG PER-CAPITA OUTLAY
BI1 WEIGHTS SUPERIMPOSED. WORKING-LESER FORM



SYMBOL	BI1W	SYMBOL	BI1W	SYMBOL	BI1W
------	0.00 - 0.24	oooooo	0.50 - 0.74	++++++	0.90 - 0.99
======	0.25 - 0.49	******	0.75 - 0.89	XXXXXX	1.00

Figure 4b

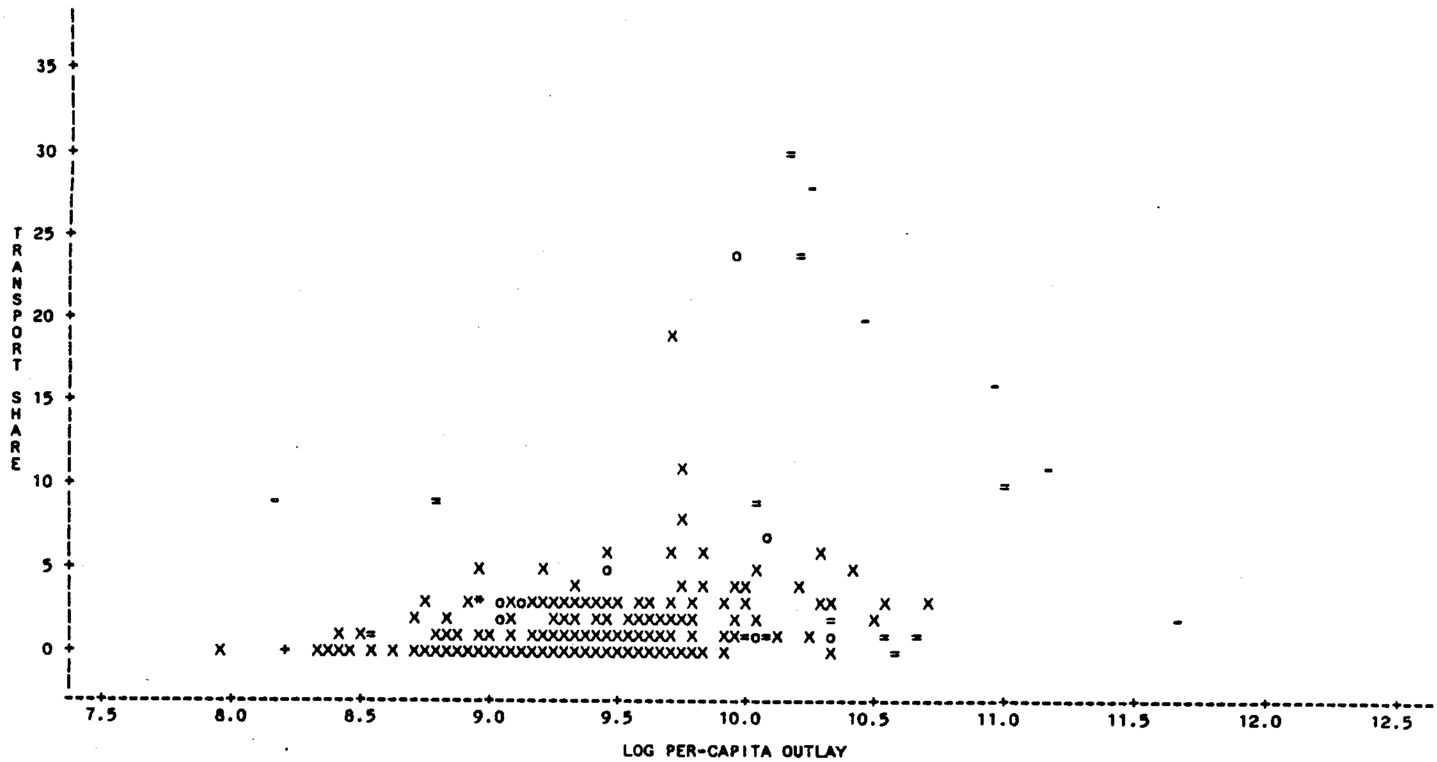
PLOT OF TRANSPORT SHARE VS. LOG PER-CAPITA OUTLAY
BI2 WEIGHTS SUPERIMPOSED. WORKING-LESER FORM



SYMBOL	BI2W	SYMBOL	BI2W	SYMBOL	BI2W
------	0.00 - 0.24	oooooo	0.50 - 0.74	++++++	0.90 - 0.99
======	0.25 - 0.49	******	0.75 - 0.89	XXXXXX	1.00

Figure 4c

PLOT OF TRANSPORT SHARE VS. LOG PER-CAPITA OUTLAY
H-K WEIGHTS SUPERIMPOSED. WORKING-LESER FORM

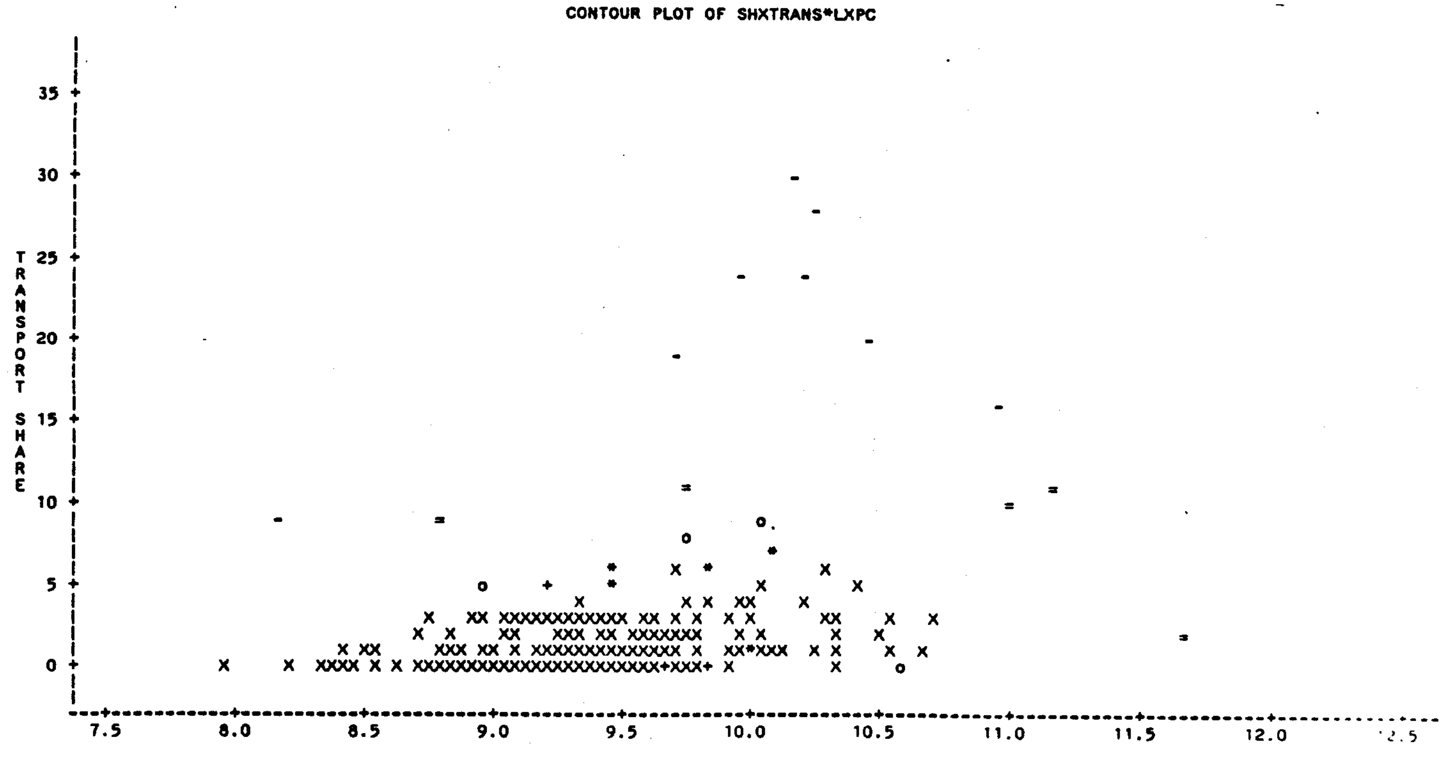


SYMBOL	HKW	SYMBOL	HKW	SYMBOL	HKW
------	0.00 - 0.24	oooooo	0.50 - 0.74	++++++	0.90 - 0.99
======	0.25 - 0.49	******	0.75 - 0.89	XXXXXX	1.00

Figure 4d

PLOT OF TRANSPORT SHARE VS. LOG PER-CAPITA OUTLAY
K-W WEIGHTS SUPERIMPOSED. WORKING-LESER FORM

CONTOUR PLOT OF SHXTRANS*LXPC



SYMBOL	KWW	SYMBOL	KWW	SYMBOL	KWW
------	0.00 - 0.24	oooooo	0.50 - 0.74	++++++	0.90 - 0.99
======	0.25 - 0.49	******	0.75 - 0.89	XXXXXX	1.00