An International Trade Flow Model With Zero Observations: an Extension of the Tobit Model

Jacob A. BIKKER
Econometric Research and Special Studies Department of the Nederlandsche Bank, Amsterdam.

Aart F. DE VOS
Department of Econometrics, Faculty of Economics, Free University, Amsterdam.

This article introduces a Bivariate normally distributed Probit Regression (BPR) model which explains data sets of economic phenomena including zero observations. It encompasses the Tobit model and models which ignore zero observations. The estimation procedure of the BPR model is very simple compared to related models.

Application to international trade flows makes clear that the BPR model is very important when the rounding threshold is relatively high. The Tobit model is rejected cogently. Most zero flows appear to be caused by economic factors, such as lack of profitability and not by statistical reasons, such as rounding to zero.

1. Introduction

Many phenomena in economics are positive (continuous) or zero (discrete). Examples are purchases of consumer durables, health care expenditure, labour force participation, (absolute values of) changes in asset holdings or in dividend payments, bilateral trade flows and foreign aid among countries. These zero observations cause problems in the estimation of log-linear models. Log-linear models cannot, of course, explain these zero-value phenomena. Linear models can, but the accumulation of observations at zero as well as the lack of negative observations points to a non-standard distribution. In practice, especially outside the more sophisticated academic research, but even in the literature, incorrect solutions such as ignoring zero observations or
replacement by - usually arbitrary - small constant values, are often encountered (e.g. LINNEMANN, 1966, JOHNSON and RAUSER, 1972, HU, 1972, WELCH, 1973, YOUNG and YOUNG, 1975, VAN BERGEIJK, 1990) in spite of the fact that the econometric literature has offered better solutions.

TOBIN (1958) has made an important contribution in this field by introducing a limited-dependent variable model which can generate a mass of zero observations. AMEMIYA (1973) provided a thorough statistical treatment for this model and FAIR (1977) devised an improvement over the laborious estimation procedure. Tobin specified a linear regression model to explain durable consumer expenditure, where the expenditure is zero when the dependent variable is negative. In fact, this is not a truthful representation of consumer behaviour, as consumers spend at least a fixed amount or nothing at all. Improvements of the model have been introduced by DAGENAIS (1969, 1975) and NELSON (1977) who replace the fixed threshold of the Tobit model by a stochastic one. In these models an action occurs when the potential magnitude of the action exceeds the threshold. However, in most situations the decision about the action does not depend on its magnitude, but upon a more or less independent process of weighing costs against benefits (HECKMAN, 1974, 1976). For example, in the case of bilateral trade flows it is likely that there is no minimum trade level but a decision about whether or not to carry on trade depending on its profitability. This profitability may relate to the volume of the trade flow due to scale effects, but depends also on other variables such as price differences between importing and exporting countries. An excellent survey on models of limited-dependent variables can be found in MADDALA (1983).

This article presents a new model, the Bivariate normally distributed Probit Regression (BPR) model, which explains the volume of a potential action by a regression model and the probability of its occurrence by a probit model. Such a model has been suggested by GOLDBERGER (1964, p. 252). The BPR model differs from Goldberger's as it allows dependency between the error terms of the two submodels. Such a correlation can be expected as included as well as deleted explanatory
variables might affect both the potential action and the probability of its occurrence. For instance, in the case of international trade flows this is likely as the profitability of potential trade flows tends to increase with their volume. It will be shown that ignoring this correlation can cause bias of the parameter estimates.

An important characteristic of the BPR model is that it encompasses the models of Tobin and Goldberger, thus permitting these models to be tested against the BPR model. The BPR model can be estimated easily whereas the estimation of the related models of Dagenais and Nelson is rather laborious and possible only under one or two rather crucial restrictions on their parameters.

Figures of bilateral trade flows often show a large number of zeros, especially between small countries at large distances. Trade flows can be zero due to economic factors, such as lack of profitability, or to statistical reasons, such as rounding to zero. Usually trade flows are explained without considering the probability of flows becoming zero. In this article the BPR model is applied to bilateral trade flows. Figures on trade between 72 or more countries including thousands of positive as well as thousands of zero flows for three years, 1959, 1974 and 1976, are examined. The analyses show that ignoring this probability of zero flows can distort the estimation of the model for positive trade flows. This certainly will occur when the trade figures are rounded and when the rounding threshold is relatively high. At the same time the analyses make clear that this statistical cause of zero flows is of minor significance compared to the economic cause. Finally, the effect on estimation results of replacing zero flows by small constants is analysed.

2. The Bivariate Normally Distributed Probit Regression Model

The main part of the BPR model consists of two equations, one for the probability of zero observations and one for the potential magnitude of a positive action:
\begin{equation}
\begin{aligned}
w_i &= \begin{cases} 
q_i & \text{if } d_i = \sum_{k=1}^{m} \gamma_k z_{ik} + v_i \leq 0 \\
y_i & \text{if } d_i > 0
\end{cases} \quad v_i \overset{d}{=} \text{IN}(0,1) \\
y_i &= \sum_{k=1}^{k} \beta_k x_{ik} + u_i \quad u_i \overset{d}{=} \text{IN}(0,\sigma^2)
\end{aligned}
\end{equation}

\(w_i\) is the observed phenomenon. The non-observed quantity \(d_i\) describes the decision process, the assessment of costs against benefits. In the case of lack of profitability the observed quantity \(w_i\) assumes the value \(q_i\). The values of \(q_i\) are known. In most applications \(q_i\) is zero or in a log-linear specification \(-\infty\). The variance of \(v_i\) is normalized at 1. This is no restriction as the parameters \(\gamma_k\) (\(k=1, \ldots, m\)) and \(v_i\) are determined, apart from a constant. Equation (2) explains the potential magnitude of the action \((y_i)\), and is observed only in the case of profitability. A multiplicative specification where \(y_i\) and \(x_{ik}\) are logarithms ensures that the potential action \(\exp(y_i)\) is positive. From (1) it follows that:

\begin{equation}
\begin{aligned}
P(w_i = q_i) &= P(v_i \leq -\sum_k \gamma_k z_{ik} = \phi(-\sum_k \gamma_k z_{ik}) \\
P(w_i \neq q_i) &= 1 - \phi(-\sum_k \gamma_k z_{ik}) = \phi(\sum_k \gamma_k z_{ik})
\end{aligned}
\end{equation}

\(\Phi\) denotes the cumulative standard normal distribution. As \(P(w_i = q_i)\) is independent of the parameters of (2), the parameters of the probit model can be estimated separately, as has been done by CRAGG (1971), AMEMIYA (1973) and POIRIER (1977). This procedure is justified when the error terms, \(u_i\) and \(v_i\), are uncorrelated. However, in many applications these error terms are correlated, since it may be expected that the profitability of the action increases with the magnitude of the potential action, in which case included and deleted explanatory variables of \(d_i\) and \(y_i\) will partly coincide. Therefore a bivariate distribution is chosen for \(u\) and \(v\):

\begin{equation}
\begin{aligned}
v_i &= \rho \frac{u_i}{\sigma} + \varepsilon_i \quad \varepsilon_i \overset{d}{=} \text{IN}(0,1-\rho^2); \quad \text{cov}(u_i, \varepsilon_i) = 0
\end{aligned}
\end{equation}
\( \rho \) is the correlation coefficient between \( u \) and \( v \). Equations (1), (2) and (5) constitute the BPR model. The likelihood of a non-zero observation is:

\[
(6) \quad P(w_i = y_i) = P(\sum \varepsilon \gamma z_i \varepsilon + v_i > 0) \ \phi((y_i - \sum \varepsilon \beta \varepsilon x_i \varepsilon)/\sigma)
\]

\[= P(\sum \varepsilon \gamma z_i \varepsilon + \rho u_i/\sigma + \varepsilon_i > 0) \ \phi((y_i - \sum \varepsilon \beta \varepsilon x_i \varepsilon)/\sigma)\]

with \( \phi \) the standard normal density function. The first right-hand side factor of (6) is the probability of an action, given \( u_i \). At first glance this may appear surprising as the decision process might be imagined to consist of two successive decisions: first carry on trade or not, and subsequently, in case of a positive decision, how much trade. The likelihood is written in reversed order. In reality the two decisions are taken simultaneously and not consecutively. Equation (6) can be rewritten as:

\[
(7) \quad P(w_i = y_i) = \Phi \left( \frac{(\sum \varepsilon \gamma z_i \varepsilon + \rho(y_i - \sum \varepsilon \beta \varepsilon x_i \varepsilon)/\sigma)}{(1-\rho^2)^{1/2}} \right) \phi((y_i - \sum \varepsilon \beta \varepsilon x_i \varepsilon)/\sigma)
\]

The estimation procedure is simplified by splitting up the bivariate quantity \( P(w_i=y_i) \) of (6) in the two independent univariate variables of (7). The logarithm of the complete likelihood function is:

\[
(8) \quad \ln L(\vec{\beta}, \vec{\gamma}, \rho, \sigma; \vec{w}, X, Z) = \sum \ln \Phi(-\sum \varepsilon \gamma z_i \varepsilon) \\
+ \sum_{i: w_i \neq q_i} \left[ \ln \Phi \{ (\sum \varepsilon \gamma z_i \varepsilon + \rho(y_i - \sum \varepsilon \beta \varepsilon x_i \varepsilon)/\sigma)(1-\rho^2)^{-1/2} \} \right] \\
- \frac{(y_i - \sum \varepsilon \beta \varepsilon x_i \varepsilon)^2}{2\sigma^2} - \ln \sigma - \frac{1}{2} \ln (2\pi)
\]

Vectors are denoted with \( \rightarrow \); \( X \) and \( Z \) are matrices with elements \( x_i \varepsilon \) and \( z_i \varepsilon \), respectively. An approximation of \( \ln \Phi \) enables the estimation of the parameters of (8) by a simple iterative modified OLS procedure (see appendix 1).
Diagram 1 can be used to illustrate some characteristics of the BPR model. The ellipses connect points \((u,v)\) with an equal probability density. The hatched area below the horizontal line \(-\Sigma \gamma z_i \ell\) contains all possible pairs of \((u,v)\) with zero values of \(d_i\), i.e. with zero flows. In the BPR model the marginal distribution of \(u_i\) is the integral over \(v_i\) over the non-hatched area of Diagram 1. This truncated distribution clearly has a positive expectation depending on \(\rho\). This can be proved when (5) is rewritten:

\[
(9) \quad u = \rho \sigma v_i + \xi_i \quad \xi_i \overset{\text{d}}{=} \text{IN}(0, \sigma^2 (1-\rho^2))
\]

So, for non-zero observations, \(u_i\) is equal to \(\rho \sigma \tilde{v}_i + \xi_i\) with \(\tilde{v}_i\) an error term with a distribution truncated from the left. This distribution is a function of \(-\Sigma \gamma z_i \ell\), so there is a correlation between \(z_i \ell\) and \(\tilde{v}_i\), and - as far as the variables \(z_i \ell\) and \(x_i \ell\) coincide - also between \(x_i \ell\) and \(u_i\), namely
via (9). These correlations have the same sign as \( \gamma_\ell \). If \( x_{i,\ell} = z_{i,\ell} \) for all \( i \) and \( \ell \), it can be shown that the absolute values of \( \beta_\ell \) for \( \ell = 2, \ldots, k \), have biased estimators when \( \rho \) is incorrectly ignored in the estimation procedure. This bias leads to overestimation when \( \beta_\ell \) and \( \gamma_\ell \) have equal signs and to underestimation when the signs are unequal. This stresses the importance of a specification allowing \( \rho \neq 0 \). Expectations of \( w_i \) and \( y_i \) follow from (9):

\[
E(y_i \mid w_i \neq q_i) = \sum \beta_\ell x_{i,\ell} + \rho \sigma E(\bar{\nu})
\]

\[
E(w_i) = P(w_i=q_i) q_i + P(w_i \neq q_i) E(y_i \mid w_i \neq q_i) = \Phi (-\sum \gamma_\ell z_{i,\ell} q_i + \Phi (\sum \gamma_\ell z_{i,\ell}) [\sum \beta_\ell x_{i,\ell} + \rho \sigma E(\bar{\nu})]
\]

These equations clearly show the bias when \( \rho \) is incorrectly assumed to be zero.

3. Related Models

3.1. The Tobit Model

The Tobit model can be considered to be a special case of the BPR model, if (i) the explanatory variables of the probit and the regression equation coincide completely, and (ii) one variable, say \( x_{i,1} \), is a constant. In this special case the decision about a transaction depends exclusively on its potential volume, that is \( d_i \) is equal to \( y_i \), apart from a constant. The following set of \( k+1 \) restrictions determines the Tobit model:

\[
\beta_\ell = \sigma \gamma_\ell \quad (\ell = 2, \ldots, k), \quad \beta_1 = \sigma \gamma_1 + r \quad \text{and} \quad \rho = 1
\]

\( r \) is the threshold. Under these restrictions (1) becomes:

\[
w_i = \begin{cases} 
q_i & \text{if } \sum \beta_\ell x_{i,\ell} + u_i < r \\
y_i & \text{elsewhere.}
\end{cases}
\]
The logarithm of the likelihood of the Tobit model is:

\[
\ln L(\vec{\beta}, \sigma; \vec{w}, X) = \sum_{i:w_i=q_i} \ln \Phi(r - \sum \beta_l x_{i,l})/\sigma) \\
- \sum_{i:w_i\neq q_i} [(y_i - \sum \beta_l x_{i,l})^2/(2\sigma^2) + \ln \sigma + 1/2 \ln (2\pi)]
\]

Appendix 2 gives an estimation procedure.

3.2. Models of Dagenais and Nelson

In the Threshold Regression model of DAGENAIIS (1969, 1975), the threshold \( r \) of the Tobit model (13) is replaced by a linear function of variables \( z_{i,k} \) plus a truncated normal distributed error. The truncation ensures a positive threshold. The equations of Dagenais are linear (not log-linear) and read in our notation as:

\[
w_i = q_i \quad \text{if} \quad y_i - d_i = \sum \beta_l x_{i,l} - \sum \delta_l z_{i,l} + u_i - \eta_i < 0 \\
w_i \quad \text{elsewhere}
\]

with \( y_i \) as in (2), \( \eta_i \) truncated from the left at \( -\sum \delta_l z_{i,l} \) and \( \text{cov}(u_i, \eta_i)=0 \). The truncated distribution leads to a very laborious estimation procedure. The model explains analogous observations as the BPR model but contains one additional parameter (\( \sigma_{\eta} \)). This may be the reason that the estimation gives rise to serious problems (DAGENAIIS, 1975, p. 279). It proves to be impossible to obtain plausible values of \( \rho \) and \( \sigma_{\eta}^2 \). Two a priori restrictions on the parameters are needed. We attribute this to numerical reasons.

The Censored Regression model of NELSON (1977) has much in common with Dagenais's model. Its distribution is not truncated because Nelson chooses a log-linear specification as in the BPR model to warrant non negative model values for the potential actions. Nelson's model also uses one parameter more than the BPR model and encounters
unresolvable estimation problems unless one restriction has been imposed. Nelson's suggestion to impose $\rho = 0$ has large consequences for the estimates as discussed earlier.

3.3. The Bivariate Normally Distributed Probit Threshold Regression Model

Non-zero actions which are smaller than half the unit in which the actions are expressed, are rounded to zero. So, in addition to the economic cause of zero flows, as mentioned earlier, there is a second mechanism which causes zero flows. To describe such a situation the BPR model should be extended with a non-stochastic threshold mechanism:

\[
(16) \quad w_i = \begin{cases} 
  y_i & \text{if } d_i = \sum \gamma \xi_i + v_i \geq 0 \text{ and } y_i = \sum \beta \xi_i + u_i \geq r_i \\
  q_i & \text{elsewhere}
\end{cases}
\]

with $r_i$ the rounding threshold of observation $i$. Equation (16), together with (2) and (5) constitutes the - what we call - Bivariate normally distributed Probit Threshold Regression (BPTR) model. The likelihood function is:

\[
(17) \quad L(\beta, \gamma, \sigma, \rho; w, X, Z, r) =
\]

\[
\prod_{i: w_i = q_i} \left[ 1 - \Phi_B((-r_i + \sum \beta \xi_i)/\sigma, \sum \gamma \xi_i, \rho) \right]
\]

\[
\prod_{i: w_i \neq q_i} \left[ \Phi_B((-r_i + \sum \beta \xi_i)/\sigma, \sum \gamma \xi_i, \rho) \phi((y_i - \sum \beta \xi_i)/\sigma) \right]
\]

with $\Phi_B$ the cumulated bivariate standard normal distribution with correlation coefficient $\rho$. In the BPTR model the points $(u, v)$ to the left of the vertical line $r_i - \sum \beta \xi_i$ in Diagram 1 are also related with zero observations, i.e. $w_i = q_i$. This BPTR model is a generalization of one of the models of CRAGG (1971), and BLUNDELL and MEGHIR (1987), who assume $\rho = 0$. The BPTR model also encompasses the BPR model: if $r_i$ goes to $-\infty$ rounding to zero does not occur.
4. A BPR Model For International Trade Flows

International trade flows have often been explained by the so-called gravity model. This model specifies that a trade flow from origin country i to destination country j is determined by supply conditions at the origin, by demand conditions at the destination and by stimulating or restraining forces relating to the specific trade flow between i and j. This model has the following form:

\[
X_{ij} = \beta_1 Y_i^{\beta_2} N_i^{\beta_3} Y_j^{\beta_4} N_j^{\beta_5} D_{ij}^{\beta_6} (P_{ij}^{uk})^{\beta_7} (P_{ij}^{f})^{\beta_8} \\
(P_{ij}^{p/b})^{\beta_9} \exp(u_{ij}) \\
u_{ij} \overset{d}{=} \text{IN}(0, \sigma^2)
\]

The trade flow from country i to country j (X_{ij}), is explained by GNP (Y_i and Y_j), population (N_i and N_j), distance (D_{ij}), and preference relations between the Commonwealth countries (P^{uk}), between France and its former colonies and among the former French colonies themselves (P^{f}), between Belgium and Zaire and between Portugal, Angola and Mozambique (P^{p/b}). The empirical results obtained with such models have always been judged very good. A large number of references can be found in BIKKER (1987).

The probability of a trade flow to become zero is described by

\[
X_{ij} = 0 \text{ if } \gamma_1 Y_i^{\gamma_2} N_i^{\gamma_3} Y_j^{\gamma_4} N_j^{\gamma_5} D_{ij}^{\gamma_6} (P_{ij}^{uk})^{\gamma_7} (P_{ij}^{f})^{\gamma_8} \\
(P_{ij}^{p/b})^{\gamma_9} \exp(v_{ij}) < 1 \\
v_{ij} \overset{d}{=} \text{IN}(0, 1)
\]

At this macro-economic level little is known about the profitability of potential bilateral trade transactions. As the probability of profitability will increase with the volume of the potential flow, the same explanatory variables are used in (18) and (19). This choice has the advantage that the BPR model encompasses the Tobit model, which allows the latter model...
to be tested.

Equations (18) and (19), together with \( \text{cov}(v_{ij}, u_{ij}) = \rho \), constitute the BPR model for bilateral trade flows. The BPR model has been applied to observations of trade flows between 72 countries. The figures date back to 1959 and 1976. These figures are described in BIKKER (1992). In both years 20 out of the 72x71=5112 trade flows are zero due to political causes. In 1959 2069 trade flows (40%) are zero due to economic or statistical causes, and in 1976 1548 trade flows (31%). The figures of 1959, used earlier by Linnemann (1966), are all rounded to the nearest one hundred thousand dollars (rounding threshold: 0.05 million dollars). The figures of 1976 have the same characteristics as those of 1959, apart from the rounding properties. They are expressed in several units, depending on the exporting country (rounding thresholds: 0.5, 0.05 and 0.005 million dollars). The fact that the trade figures are rounded makes the BPTR model, in principle, more suitable than the BPR model. However, the likelihood of the BPTR model includes a two-dimensional integral over the bivariate normal density, namely the shaded area of Diagram 1, which makes the estimation very laborious. Numerical estimation procedures did not generate satisfying results. In the next section we will prove that the effect of rounding is rather limited. Therefore no further efforts have been made to obtain more satisfying estimation results for the BPTR model.

Tables 1 and 2 give the estimation results for the BPR model, the Tobit model and the Independent Probit Regression (IPR) model (i.e. the BPR model with \( \rho = 0 \)). The results show that the correlation between \( u \) and \( v \) has changed between 1959 and 1976. In 1959 the BPR model with optimal \( \rho (\rho = 0.29) \) differs significantly from the IPR model. The value of the likelihood ratio test is 30, whereas the critical value is \( \chi^2_{0.05}(1) = 3.84 \). The corresponding \( \beta \)'s and \( \gamma \)'s also differ substantially. All the IPR estimates are biased: the mean underestimation is 8.7%. In 1976 the optimal \( \rho (\rho = 0.05) \) does not differ significantly from zero. The difference between 1959 and 1976 may be explained by the fact that the rounding threshold was relatively high in 1959 considering the 383% increase in the mean trade flow from 1959 to 1976. (The increase is even distorted downwards
by the 521 additional, usually small, non-zero trade flows in 1976 compared to 1959).

### Table 1: Estimation Results of the BPR Model Applied to Trade Flows of 1959

<table>
<thead>
<tr>
<th>Model</th>
<th>Independent Probit Regression</th>
<th>BPR(0.29)</th>
<th>Tobit model b)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Column</strong></td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td><strong>p</strong></td>
<td>0</td>
<td>0.29</td>
<td>1</td>
</tr>
<tr>
<td><strong>REGRESSION EQUATION</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>$\beta_1$</td>
<td>-8.43 (11.2)</td>
<td>-9.14 (12.2)</td>
</tr>
<tr>
<td>GNP</td>
<td>$\beta_2$</td>
<td>1.10 (36.7)</td>
<td>1.20 (40.0)</td>
</tr>
<tr>
<td>Population</td>
<td>$\beta_3$</td>
<td>-0.35 (11.7)</td>
<td>-0.39 (13.0)</td>
</tr>
<tr>
<td>GNP</td>
<td>$\beta_4$</td>
<td>0.98 (32.7)</td>
<td>1.06 (35.3)</td>
</tr>
<tr>
<td>Population</td>
<td>$\beta_5$</td>
<td>-0.28 (9.3)</td>
<td>-0.33 (11.0)</td>
</tr>
<tr>
<td>Distance</td>
<td>$\beta_6$</td>
<td>-0.76 (25.3)</td>
<td>-0.82 (27.3)</td>
</tr>
<tr>
<td>Preference: uk</td>
<td>$\beta_7$</td>
<td>0.85 (8.5)</td>
<td>0.92 (9.2)</td>
</tr>
<tr>
<td>f</td>
<td>$\beta_8$</td>
<td>2.19 (7.1)</td>
<td>2.39 (7.7)</td>
</tr>
<tr>
<td>p/b</td>
<td>$\beta_9$</td>
<td>5.11 (7.7)</td>
<td>5.38 (8.0)</td>
</tr>
<tr>
<td><strong>PROBIT EQUATION</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>$\gamma_1$</td>
<td>-5.18 (2.2)</td>
<td>-5.14 (4.3)</td>
</tr>
<tr>
<td>GNP</td>
<td>$\gamma_2$</td>
<td>1.07 (21.4)</td>
<td>1.09 (21.8)</td>
</tr>
<tr>
<td>Population</td>
<td>$\gamma_3$</td>
<td>-0.43 (10.8)</td>
<td>-0.44 (11.0)</td>
</tr>
<tr>
<td>GNP</td>
<td>$\gamma_4$</td>
<td>0.93 (18.6)</td>
<td>0.94 (18.8)</td>
</tr>
<tr>
<td>Population</td>
<td>$\gamma_5$</td>
<td>-0.52 (13.0)</td>
<td>-0.53 (13.3)</td>
</tr>
<tr>
<td>Distance</td>
<td>$\gamma_6$</td>
<td>-0.95 (15.8)</td>
<td>-0.93 (15.5)</td>
</tr>
<tr>
<td>Preference: uk</td>
<td>$\gamma_7$</td>
<td>0.65 (3.6)</td>
<td>0.70 (4.1)</td>
</tr>
<tr>
<td>f</td>
<td>$\gamma_8$</td>
<td>2.68 (3.9)</td>
<td>2.74 (4.1)</td>
</tr>
<tr>
<td>p/b</td>
<td>$\gamma_9$</td>
<td>3.36 (1.5)</td>
<td>3.18 (3.3)</td>
</tr>
<tr>
<td>standard deviation</td>
<td>$\sigma$</td>
<td>1.32</td>
<td>1.33</td>
</tr>
<tr>
<td>likelihood (in log.)</td>
<td></td>
<td>-6883</td>
<td>-6868</td>
</tr>
<tr>
<td>number of iterations</td>
<td></td>
<td>1</td>
<td>12</td>
</tr>
</tbody>
</table>

a) t-values in parentheses; the BPR t-values of $\beta$ are conditional on $\gamma$, $p$ and $\sigma$, etc., see appendix 1.

b) The values of $\gamma$ in this column have been calculated as follows:

$$\gamma_1 = (\beta_1 - \ln(0.05))/\sigma$$ and $$\gamma_2 = \gamma_2/\sigma$$ for $\gamma = 2, \ldots, 9$. 

TABLE 2: Estimation Results of the BPR Model Applied to Trade Flows of 1976

<table>
<thead>
<tr>
<th>Model</th>
<th>Independent Probit Regression</th>
<th>BPR(0.05)</th>
<th>Tobit model a)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Column 1</td>
<td>Column 2</td>
<td>Column 3</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>0.05</td>
<td>1</td>
</tr>
<tr>
<td>REGRESSION EQUATION</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>$\beta_1$</td>
<td>-2.68 (2.7)</td>
<td>-2.79 (2.8)</td>
</tr>
<tr>
<td>GNP (exporting country)</td>
<td>$\beta_2$</td>
<td>1.19 (39.7)</td>
<td>1.21 (40.3)</td>
</tr>
<tr>
<td>Population</td>
<td>$\beta_3$</td>
<td>-0.39 (9.8)</td>
<td>-0.40 (10.0)</td>
</tr>
<tr>
<td>GNP (importing country)</td>
<td>$\beta_4$</td>
<td>0.89 (29.7)</td>
<td>0.91 (30.3)</td>
</tr>
<tr>
<td>Population</td>
<td>$\beta_5$</td>
<td>-0.24 (6.0)</td>
<td>-0.25 (6.3)</td>
</tr>
<tr>
<td>Distance</td>
<td>$\beta_6$</td>
<td>-0.97 (24.3)</td>
<td>-0.98 (24.5)</td>
</tr>
<tr>
<td>Preference: uk</td>
<td>$\beta_7$</td>
<td>0.19 (1.4)</td>
<td>0.19 (1.4)</td>
</tr>
<tr>
<td>f</td>
<td>$\beta_8$</td>
<td>0.68 (1.6)</td>
<td>0.71 (1.7)</td>
</tr>
<tr>
<td>p/b</td>
<td>$\beta_9$</td>
<td>3.62 (4.1)</td>
<td>3.67 (4.1)</td>
</tr>
<tr>
<td>PROBIT EQUATION</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>$\gamma_1$</td>
<td>-2.17 (0.7)</td>
<td>-2.08 (0.6)</td>
</tr>
<tr>
<td>GNP (exporting country)</td>
<td>$\gamma_2$</td>
<td>0.76 (10.9)</td>
<td>0.77 (11.0)</td>
</tr>
<tr>
<td>Population</td>
<td>$\gamma_3$</td>
<td>-0.40 (6.7)</td>
<td>-0.40 (6.7)</td>
</tr>
<tr>
<td>GNP (importing country)</td>
<td>$\gamma_4$</td>
<td>0.76 (12.7)</td>
<td>0.76 (12.7)</td>
</tr>
<tr>
<td>Population</td>
<td>$\gamma_5$</td>
<td>-0.40 (6.7)</td>
<td>-0.40 (6.7)</td>
</tr>
<tr>
<td>Distance</td>
<td>$\gamma_6$</td>
<td>-0.78 (9.8)</td>
<td>-0.78 (9.8)</td>
</tr>
<tr>
<td>Preference: uk</td>
<td>$\gamma_7$</td>
<td>-0.11 (0.5)</td>
<td>-0.10 (0.5)</td>
</tr>
<tr>
<td>f</td>
<td>$\gamma_8$</td>
<td>1.85 (2.1)</td>
<td>1.88 (2.1)</td>
</tr>
<tr>
<td>p/b</td>
<td>$\gamma_9$</td>
<td>2.69 (0.9)</td>
<td>2.30 (0.7)</td>
</tr>
<tr>
<td>rounding threshold $r_1$ b)</td>
<td>$\gamma_{10}$</td>
<td>0.12 (3.0)</td>
<td>0.10 (2.5)</td>
</tr>
<tr>
<td>standard deviation</td>
<td>$\sigma$</td>
<td>1.77</td>
<td>1.77</td>
</tr>
<tr>
<td>likelihood (in log.)</td>
<td></td>
<td>-8957.4</td>
<td>-8957.1</td>
</tr>
<tr>
<td>number of iterations</td>
<td></td>
<td>1</td>
<td>4</td>
</tr>
</tbody>
</table>

a) The values of $\gamma$ in this column have been calculated as follows:
$$\gamma_\lambda = \beta_\lambda / \sigma$$ for $\lambda = 1, \ldots, 9$ and $\gamma_{10} = -1 / \sigma$

b) One additional explanatory variable, the rounding threshold $r_1$, has been used to enable the nesting of the BPR model. As the rounding threshold differs between exporting countries, a generalization of the restriction of (12) is necessary: $\beta_1 = \sigma \gamma_1 + r_1$ instead of $\beta_1 = \gamma_1 + r_1$.

VAN MAANEN (1988) uses the Tobit model to explain trade flows. However, the estimates of $\beta$ in Tables 1 and 2 make clear that the Tobit estimates are distorted by the incorrect assumption about $\rho$ ($\rho=1$ versus...
\( \rho = 0.29 \) (1959) or \( \rho = 0.05 \) (1976)). Their mean absolute value is higher by 50% (1959) and 131% (1976), respectively, than in the BPR model. The Tobit model is rejected cogently for both years.

The coefficients of the gravity model have their well-known values around 1 for the GNPs, -0.3 for the populations, strongly negative for distances and positive for the preference relations. Surprisingly, the probit parameters, the \( \gamma \)'s, are almost equal to the corresponding gravity parameters. They almost satisfy the relation \( \gamma_k = \beta_k / \sigma \) for \( k = 2, ..., 9 \). This confirms the assumption about the dependency of profitability on the potential volume of the trade flows. Apparently this does not hold for the so-called deleted explanatory variables embodied in the error terms, witness the low value of \( \rho \), especially in 1976. Probably the profitability is affected by all kind of micro-economic factors which do not influence the volume of the trade flows.

Comparison of the optimal models for 1959 and 1976 shows only limited changes over time in the \( \beta \)'s as well as in the \( \gamma \)'s (the latter multiplied by \( \sigma \)), except for the allocation parameters. The distance effect in 1976 is higher and, in line with expectations, the preference effects are all lower. These relatively small changes over time raise expectations that the BPR model can be used very well to predict future trade flows.

5. The Effect of Rounding to Zero

In this section the effect of rounding to zero will be analysed. The calculations are based on trade flows of 1974 between 80 countries, which are described in BIKKER (1987). The characteristics of this data set differ from those of 1959 and 1976. For instance, these figures have been registered by importing countries whereas the other figures were observed by exporting countries. Again 20 out of the 80x79=6320 trade flows are zero due to political causes and 2346 (37%) due to economic or statistical causes. The unit of the figures of 1974 is one million or one hundred thousand dollars, depending on the importing country. So the rounding threshold is $0.5 and 0.05 million, respectively. The effect of
### TABLE 3: Estimation Results of the BPR Model Applied to Trade Flows of 1974

<table>
<thead>
<tr>
<th>Model</th>
<th>Independent Probit Regression</th>
<th>BPR(0.03)</th>
<th>Tobit model a)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Column</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>ρ</td>
<td>0</td>
<td>0.03</td>
<td>1</td>
</tr>
</tbody>
</table>

#### REGRESSION EQUATION

<table>
<thead>
<tr>
<th></th>
<th>β₁</th>
<th>β₂</th>
<th>β₃</th>
<th>β₄</th>
<th>β₅</th>
<th>β₆</th>
<th>β₇</th>
<th>β₈</th>
<th>β₉</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-1.26 (17.0)</td>
<td>1.02 (56.7)</td>
<td>-0.19 (8.7)</td>
<td>1.02 (56.4)</td>
<td>-0.21 (9.0)</td>
<td>-0.94 (33.5)</td>
<td>1.16 (8.2)</td>
<td>2.47 (10.1)</td>
<td>5.86 (8.3)</td>
</tr>
<tr>
<td>GNP exporting country</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Population importing country</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### PROBIT EQUATION

<table>
<thead>
<tr>
<th></th>
<th>γ₁</th>
<th>γ₂</th>
<th>γ₃</th>
<th>γ₄</th>
<th>γ₅</th>
<th>γ₆</th>
<th>γ₇</th>
<th>γ₈</th>
<th>γ₉</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-0.23 (1.5)</td>
<td>0.81 (33.8)</td>
<td>-0.18 (7.6)</td>
<td>0.75 (25.7)</td>
<td>-0.34 (13.6)</td>
<td>-0.91 (21.7)</td>
<td>0.79 (4.3)</td>
<td>1.87 (7.0)</td>
<td>2.93 (2.1)</td>
</tr>
<tr>
<td>GNP exporting country</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Population importing country</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### Standard Deviation

<table>
<thead>
<tr>
<th>σ</th>
<th>1.38</th>
<th>1.38</th>
<th>2.03</th>
</tr>
</thead>
</table>

#### Likelihood (in log.)

<table>
<thead>
<tr>
<th>-9148.822</th>
<th>-9148.821</th>
<th>-9973</th>
</tr>
</thead>
</table>

#### Number of Iterations

<table>
<thead>
<tr>
<th>1</th>
<th>10</th>
<th>4</th>
</tr>
</thead>
</table>

---

a) The values of γ in this column have been calculated as follows:

γₖ = βₖ/σ for k = 1, ..., 9 and γ₁₀ = -1/σ

Rounding to zero can be illustrated clearly by an experimental calculation in which all non-zero trade flows below $0.5 million are rounded to zero. The number of zero flows then increases from 2346 (37%) to 2915 (46%). In this manner rounding as a cause of zero flows is stressed strongly. The estimation results based on the original and the censored
data set are shown in Tables 3 and 4, respectively. The parameter estimates are affected dramatically. Rounding to zero of 9% of the figures increases the optimal $\rho$ from 0.03 to 0.62.

### Table 4: Estimation Results of the BPR Model Applied to Trade Flows Larger than $\$0.5 \text{ Million (1974)}$

<table>
<thead>
<tr>
<th>Model</th>
<th>Independent Probit Regression</th>
<th>BPR(0.62)</th>
<th>Tobit model a)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Column</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>$\rho$</td>
<td>0</td>
<td>0.62</td>
<td>1</td>
</tr>
<tr>
<td>REGRESSION EQUATION</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>$\beta_1$</td>
<td>-0.59 (8.5)</td>
<td>-1.32 (18.9)</td>
</tr>
<tr>
<td>GNP</td>
<td>$\beta_2$</td>
<td>0.88 (52.0)</td>
<td>1.04 (60.9)</td>
</tr>
<tr>
<td>Population</td>
<td>$\beta_3$</td>
<td>-0.15 (6.9)</td>
<td>-0.19 (9.3)</td>
</tr>
<tr>
<td>GNP</td>
<td>$\beta_4$</td>
<td>0.85 (50.1)</td>
<td>1.00 (58.8)</td>
</tr>
<tr>
<td>Population</td>
<td>$\beta_5$</td>
<td>-0.12 (5.6)</td>
<td>-0.19 (9.2)</td>
</tr>
<tr>
<td>Distance</td>
<td>$\beta_6$</td>
<td>-0.81 (32.5)</td>
<td>-0.94 (36.0)</td>
</tr>
<tr>
<td>Preference: uk</td>
<td>$\beta_7$</td>
<td>0.91 (7.1)</td>
<td>1.07 (8.4)</td>
</tr>
<tr>
<td>$f$</td>
<td>$\beta_8$</td>
<td>2.12 (9.2)</td>
<td>2.47 (10.7)</td>
</tr>
<tr>
<td>$p/b$</td>
<td>$\beta_9$</td>
<td>5.07 (8.4)</td>
<td>5.88 (9.3)</td>
</tr>
<tr>
<td>PROBIT EQUATION</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>$\gamma_1$</td>
<td>-0.83 (10.4)</td>
<td>-0.96 (11.3)</td>
</tr>
<tr>
<td>GNP</td>
<td>$\gamma_2$</td>
<td>0.86 (39.0)</td>
<td>0.90 (37.4)</td>
</tr>
<tr>
<td>Population</td>
<td>$\gamma_3$</td>
<td>-0.21 (8.9)</td>
<td>-0.21 (9.2)</td>
</tr>
<tr>
<td>GNP</td>
<td>$\gamma_4$</td>
<td>0.84 (38.6)</td>
<td>0.87 (37.7)</td>
</tr>
<tr>
<td>Population</td>
<td>$\gamma_5$</td>
<td>-0.35 (14.8)</td>
<td>-0.35 (15.0)</td>
</tr>
<tr>
<td>Distance</td>
<td>$\gamma_6$</td>
<td>-0.94 (24.0)</td>
<td>-0.92 (22.5)</td>
</tr>
<tr>
<td>Preference: uk</td>
<td>$\gamma_7$</td>
<td>0.85 (4.7)</td>
<td>0.89 (5.8)</td>
</tr>
<tr>
<td>$f$</td>
<td>$\gamma_8$</td>
<td>2.01 (7.6)</td>
<td>2.22 (8.9)</td>
</tr>
<tr>
<td>$p/b$</td>
<td>$\gamma_9$</td>
<td>3.73 (3.0)</td>
<td>4.50 (2.4)</td>
</tr>
<tr>
<td>standard deviation</td>
<td>$\sigma$</td>
<td>1.18</td>
<td>1.25</td>
</tr>
<tr>
<td>likelihood (in log.)</td>
<td>-7494</td>
<td>-7432</td>
<td>-7714</td>
</tr>
<tr>
<td>number of iterations</td>
<td></td>
<td>1</td>
<td>5</td>
</tr>
</tbody>
</table>

a) The values of $\gamma$ in this column have been calculated as follows:

$$\gamma_1 = (\beta_1 - \ln(0.5))/\sigma$$

and

$$\gamma_k = \beta_k/\sigma \text{ for } k = 2, \ldots, 9.$$

This experimental calculation shows convincingly that the influence of rounding to zero in the original data set must have been negligible,
witness the low value of ρ (0.03) in that case. That demonstrates that the use of the BPR model instead of the BPTR model is justified. In principle this shift of the BPR model (ρ from 0.03 to 0.62) towards the Tobin model (ρ=1) shows the suitability of the latter model for rounded figures, when rounding is the only cause of zero flows. On the other hand, the results of Table 4 show how cogently the Tobin model is rejected even in this experiment: its log-likelihood is almost 300 points lower. This brings out the importance in actual reality of economic causes of zero flows which are ignored in the Tobit model. A test of the robustness of the model and its estimation procedure can be made by comparing the BPR parameters of Tables 3 and 4. The very high degree of conformity of the gravity parameter estimates, \( \beta \), shows that the BPR model can explain the observed trade flows in cases where the rounding threshold is high or is raised.

6. Replacing Zero Flows by Small Constants

Finally an illustration is given of the effect of replacing zero flows by small positive constants. This has been done by e.g. LINNEMANN (1966) with the data set of 1959, WELCH (1973) and VAN BERGEN (1990). Table 5 gives the IPR, BPR and Tobit estimates as well as Linnemann's estimates based on successively: sole positive trade flows, and positive trade flows plus selected zero flows replaced by $0.02 and 0.01 million, respectively. Linnemann used trade flows between 80 countries (3400 non-zero flows); in this article trade flows between 72 countries (3023 non-zero flows) are used, which explains the difference between columns 1 and 4. Like BIJKER (1982) Linnemann established that only a small percentage (15%) of the zero flows has an expected value below the rounding threshold of $0.05 million. Linnemann concluded that the smallest flows are overestimated by the regression model, based on non-zero flows. To redress this effect, 1532 zero flows are replaced by a constant, which pushes the regression plane downwards at the lower end, near the smallest flows. This leads to larger parameters, increasing the

---

1 Another discrepancy is that the preference relations are arranged differently.
difference between small and large flows. The downward move at the lower end of the regression plane also occurs in the BPR model. This approach with constants which is arbitrary, as noted by Linnemann himself, seems to overshoot the aim. His mean absolute estimates are 24% (Table 5, column 2) and 35% (column 3) too high, respectively.

**TABLE 5 : Estimation Results of the Independent Regression Model Compared With Those of the BPR Model**

<table>
<thead>
<tr>
<th>Estimation procedure</th>
<th>Linnemann's regression model</th>
<th>BPR model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>non-zero flows only</td>
<td>selected zeros replaced by</td>
</tr>
<tr>
<td></td>
<td>S 0.02</td>
<td>S 0.01</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Code of Linnemann</th>
<th>AC 1</th>
<th>AC 23</th>
<th>AC 17</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table</td>
<td>Table 5.1</td>
<td>Table 5.12</td>
<td>Table 5.11</td>
</tr>
<tr>
<td></td>
<td>This article, Table 1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Column</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>GNP exporting country</td>
<td>0.99</td>
<td>1.42</td>
<td>1.58</td>
<td>1.10</td>
<td>1.20</td>
<td>1.73</td>
</tr>
<tr>
<td>Population</td>
<td>-0.20</td>
<td>-0.39</td>
<td>-0.51</td>
<td>-0.35</td>
<td>-0.39</td>
<td>-0.63</td>
</tr>
<tr>
<td>GNP importing country</td>
<td>0.85</td>
<td>1.27</td>
<td>1.30</td>
<td>0.98</td>
<td>1.06</td>
<td>1.54</td>
</tr>
<tr>
<td>Population</td>
<td>-0.15</td>
<td>-0.46</td>
<td>-0.40</td>
<td>-0.28</td>
<td>-0.33</td>
<td>-0.69</td>
</tr>
<tr>
<td>Distance</td>
<td>-0.81</td>
<td>-1.17</td>
<td>-1.33</td>
<td>-0.76</td>
<td>-0.82</td>
<td>-1.20</td>
</tr>
</tbody>
</table>

7. Conclusions

The BPR model proves to be a useful model to explain sets of observations including zeros. The estimation results are plausible and there are no estimation problems as encountered for related models. The parameter estimates prove to be very sensitive to the correct specification with respect to \( p \), the coefficient of correlation between the errors of the two equations. This is an important conclusion because often models are used
with a given value of $\rho$, such as the independent probit regression model ($\rho=0$) or the Tobit model ($\rho=1$).

Application to a large number of bilateral trade flows makes clear that the BPR model is of great importance when the rounding threshold is relatively high as it is in the 1959 data set. For the data sets of the seventies the BPR model is not better than independent estimation of non-zero and zero flows. In all applications, even though there are (real or artificial) rounding thresholds, the Tobit model is rejected cogently. Rounding to zero in the seventies (and later) is not an important cause of zero observations. Most zero flows are caused by economic reasons, such as lack of profitability. The probability of zero flows can be explained by the same variables which determine the volume of the potential trade transaction.
Appendix 1: Estimation of the BPR Model

Approximation of LN Φ

Usually the estimation of limited-dependent variable models is very laborious. Here a simple iterative modified OLS procedure is introduced based on an approximation of the cumulative standard normal distribution function with a spline function:

\[
\ln \Phi(q) = \sum_{h=1}^{H} a_h (q-b_h)^2 I_h(q)
\]

The domain of the argument q is divided into H intervals. Two constants belong to each interval. \(I_h(q)\) is an indication function which is 1 if q lies in the h's interval and zero elsewhere. Intervals and constants used are (H=5):

\[(\infty, -3.4) \quad a_1 = 0.5 \quad b_1 = 0.5844\]
\[(-3.4, -1.7) \quad a_2 = 0.4 \quad b_2 = 1.0662\]
\[(-1.7, -0.1) \quad a_3 = 0.3 \quad b_3 = 1.4869\]
\[(-0.1, 2.0) \quad a_4 = 0.1665 \quad b_4 = 2\]
\[(2.0, \infty) \quad a_5 = 0\]

The approximation is very close which becomes clear when \(\ln \Phi(q)\) and its approximation are plotted.

Estimation Procedure

The likelihood of the BPR model (8) contains the parameters \(\vec{\beta}, \vec{\gamma}, \sigma\) and \(\rho\). The estimation procedure is iterative. Starting with \(\rho=0\), initial values of \(\vec{\beta}\) and \(\sigma\) are obtained by ordinary regression deleting the zero observations, and initial values of \(\vec{\gamma}\) are obtained by a simple probit estimation procedure. \(\rho\) is found by numerical search, optimising the relevant part of the likelihood:

\[L(\rho|\vec{\gamma}, \vec{\beta}, \sigma) \quad |\rho|<1\]
In each iteration \( \beta, \gamma, \sigma \) and \( \rho \) are successively estimated conditional on the estimates of the other parameters.

After substitution of (20) into (8), the log-likelihood of the BPR model is:

\[
\ln L(\beta, \gamma, \sigma, \rho; w, x, z) = -\sum_{i:w_i=q_i} \sum_h a_h (-\sum \gamma z_i \cdot \xi - bh)^2 \\
I_h (-\sum \gamma z_i \cdot \xi) - \sum_{i:w_i\neq q_i} [\sum_h a_h (c_i - bh)^2] \\
I_h (c_i) + (y_i - \sum \beta x_i \cdot \xi)^2 / (2\sigma^2) + \ln \sigma + 1/2 \ln (2\pi)
\]

with \( c_i = [\sum \gamma z_i \cdot \xi + \rho(y_i - \sum \beta x_i \cdot \xi)/\sigma]/(1-\rho^2)^{1/2} \)

**Estimation of \( \beta \) Conditional on \( \gamma, \sigma \) and \( \rho \)**

Only non-zero observations are relevant for the estimation of \( \beta \). The estimates of \( \beta \) can be obtained by ordinary regression, defining:

\[
y_i = \sum_h \sigma(2a_h)^{1/2} [(-\sum \gamma z_i \cdot \xi + \rho y_i/\sigma)/((1-\rho^2)^{1/2} - bh) \cdot I_h (c_i) \\
x_i \cdot \xi = \sum_h \rho(2a_h/(1-\rho^2))^{1/2} x_i \cdot \xi \cdot I_h (c_i)
\]

Regression of \( \begin{pmatrix} y_i \\ \xi \cdot \xi \end{pmatrix} \) on \( \begin{pmatrix} \gamma \\
\chi \end{pmatrix} \) yields the estimates of \( \beta \), conditional on estimates of \( \gamma, \sigma \) and \( \rho \) of the previous round in the iteration.

**Estimation of \( \gamma \) and \( \sigma \)**

The estimates of \( \gamma \) result from regression of \( y'' \) on \( X'' \) defining for non-zero observations:

\[
y_i'' = \sum_h (-bh\sigma + \rho/(1-\rho^2)^{1/2} (y_i - \sum \beta x_i \cdot \xi)) (2a_h)^{1/2} I_h (c_i)
\]
\[ x_{i,k} = - \sum_h \sigma (2a_h/(1-\rho^2))^{1/2} z_{i,k} I_h (c_i) \]

and for zero observations:

\[ y_i = - \sum_h b_h \sigma (2a_h)^{1/2} I_h (-\Sigma \mathbf{x}_i z_{i,k}) \]

\[ x_{i,k} = \sum_h \sigma (2a_h)^{1/2} z_{i,k} I_h (-\Sigma \mathbf{x}_i z_{i,k}) \]

\( \sigma \) is obtained from:

\[
\frac{\partial \ln L(\sigma; \mathbf{w}, \mathbf{X}, \mathbf{Z} | \mathbf{\gamma}, \mathbf{\beta}, \rho)}{\partial \sigma} =
\]

\[
\sum_{i:wi \neq qi} \sigma^2 - \sigma \sum_h \left( a_h \sum_{i:wi \neq qi} 2\rho (1-\rho^2)^{-1/2} \right)
\]

\[
[(1-\rho^2)^{-1/2} \Sigma \mathbf{x}_i z_{i,k} - b_h] (y_i - \Sigma \mathbf{x}_i z_{i,k}) I_h (c_i)
\]

\[
+ \sum_h a_h \sum_{i:wi \neq qi} \{ 2\rho^2 (y_i - \Sigma \mathbf{x}_i z_{i,k})^2 (1-\rho^2)^{-1} I_h (c_i) \}
\]

\[
+ (y_i - \Sigma \mathbf{x}_i z_{i,k})^2 = 0 = A\sigma^2 + B\sigma + C
\]

The last equation holds by definition. If \( B \) is negative, \( \sigma \) follows from:

\[ \sigma = (-B + (B^2 - 4AC)^{1/2})/(2A). \]
Appendix 2 : Estimation of the Tobit Model

The estimation procedure of the Tobit model is also based on the approximation of the cumulative standard normal distribution function. After substitution of (20) into (14) the loglikelihood of the Tobit model is:

$$\ln L(\beta, \sigma; w, X, r) = -\sum h \left( \sum_{i: w_i \neq q_i} a_h ((r_i - \sum \beta_i x_{i,x})/\sigma - b_h)^2 \right)$$

$$= I_h \left( \frac{(r_i - \sum \beta_i x_{i,x})}{\sigma} \right) - \sum_{i: w_i \neq q_i} \left( \frac{(y_i - \sum \beta_i x_{i,x})^2}{2\sigma^2} \right) + \ln \sigma + 1/2 \ln (2\pi)$$

Usually, as in TOBIN (1958), $r_i = r$ holds for all $i$.

For given $\sigma$ estimates of $\beta$ may be obtained by ordinary least squares if the dependent variable $y_i$ and the explanatory variables $x_{i,x}$ are defined as, for zero observations:

$$y_i = y_i \quad \text{and} \quad x_{i,x} = x_{i,x}$$

and for non-zero observations:

$$y_i = \sum_h (2a_h)^{1/2} (r_i - b_h \sigma) I_h \left( \frac{(r_i - \sum \beta_i x_{i,x})}{\sigma} \right)$$

$$x_{i,x} = \sum_h (2a_h)^{1/2} x_{i,x} I_h \left( \frac{(r_i - \sum \beta_i x_{i,x})}{\sigma} \right)$$

The maximum likelihood estimate of $\sigma$ given $\beta$ is obtained by putting $\partial \ln L/\partial \sigma = 0$. 


\[ \sum_{i:w_i \neq q_i} \sigma^2 + \sigma \sum_h \left( 2a_h \sum_{i:w_i=q_i} (r_i - \sum \beta \cdot x_i \cdot \varepsilon) b_h \right) i_h \left( \frac{(r_i - \sum \beta \cdot x_i \cdot \varepsilon)}{\sigma} \right) \]

\[ - \sum_h \left( 2a_h \sum_{i:w_i=q_i} (r_i - \sum \beta \cdot x_i \cdot \varepsilon)^2 b_h \right) i_h \left( \frac{(r_i - \sum \beta \cdot x_i \cdot \varepsilon)}{\sigma} \right) \]

\[ - \sum_{i:w_i \neq q_i} (y_i - \sum \beta \cdot x_i \cdot \varepsilon)^2 = 0 = A\sigma^2 + B\sigma + C \]

The last equation holds by definition. If \( B \) is negative, \( \sigma \) follows from:

(21) \[ \sigma = \frac{-B + (B^2 - 4AC)^{1/2}}{2A} \]

Maximum likelihood estimates of \( \sigma \) and \( \beta \) are obtained by iterating between the OLS regression and (21). A starting value for \( \sigma \) and \( \beta \) may be found by deleting the zero observations in the first round. It is likely that in most cases \( \sigma \) will mainly depend on the non-zero observations (the zero observations provide only rather vague information), so the iteration will quickly converge as was found in the estimation of the trade flow model.
REFERENCES


HECKMAN, J.J. (1976), The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models, *Annals of Economic and Social Measurement* 5, pp 475-492.

HU, T. (1971), The fitting of log-regression functions when some observations in the regressand are zero or negative, *American Journal of Agricultural Economics*.


