CentER
for
Economic Research

Discussion paper

[redacted]

TILBURG UNIVERSITY

No. 8948

MAXIMUM SCORE ESTIMATION IN THE ORDERED RESPONSE MODEL

by Peter Kooreman
and Bertrand Melenberg

October, 1989

Maximum Score Estimation in the Ordered Response Model

Peter Kooreman*
Bertrand Melenberg*

October 1989

Tilburg University
Department of Economics
P.O.B. 90153
5000 LE Tilburg
The Netherlands

*Financial support has been provided by the Netherlands Organization for Scientific Research (NWO) and the Royal Netherlands Academy of Arts and Sciences (KNAW).

Abstract

We extend Manski's maximum score estimator to the ordered response case and prove that the generalized estimator is strongly consistent.

1. Introduction

This paper is concerned with semiparametric estimation of the regression coefficients of a linear model when the dependent variable is grouped, i.e. is only observed to fall in certain known intervals on a continuous scale, its actual value remaining unobserved. Such a model is usually referred to as the ordered response model. A typical example of a variable for which often only grouped information is available is income. Many respondents are often either unable or unwilling to provide a precise measure of their income, and therefore many micro data sets only contain bracketwise information on income; see, for example, Cramer (1969), Stewart (1983) and Kapteyn et al. (1988). The usual estimation method for the ordered response model is maximum likelihood, under the assumption that the error terms are independent, and identically normally distributed. Stewart (1983) gives a convenient algorithm to obtain the maximum likelihood estimates in this case.

As in many other models with limited dependent variables, the maximum likelihood estimator is generally inconsistent if the underlying distributional assumptions are not correct. Therefore, it is important to search for alternative estimators which are consistent under weaker distributional assumptions.

During the last decade several semiparametric estimators have been proposed for models with limited dependent endogenous variables; see Robinson (1988) for a review. However, most of these methods are designed for either the binary response model or the censored regression model, and are not straightforwardly applicable to the ordered response case.

In this paper we extend Manski's Maximium Score Estimator (MSCORE; see Manski (1975, 1985)) to the ordered response case. An attractive property of MSCORE and its generalization is that it does not require the error terms to be identically distributed. Our extension builds on the fact that MSCORE can be interpreted as a least absolute deviations estimator. An important difference with the binary case is that the ordered response model does not require a normalization of the parameter vector.

Section 2 introduces the generalized estimator and in section 3 we prove its consistency.

2. Maximum Score estimation in the ordered response model

Consider the following binary response model

$$\begin{cases} y_i^* = x_i'\beta + \varepsilon_i \\ \\ y_i = 1 \text{ if } y_i^* > 0 \\ \quad\; 0 \text{ otherwise} \end{cases} \qquad i=1,\ldots,N \tag{2.1}$$

It is assumed that the error term ε_i has a unique median $Med(\varepsilon_i|x_i)=0$. x_i is a K-vector of explanatory variables and β is a conformable vector of true parameters.

Define, for some vector b, the score of observation i to be equal to 1 if y_i^* and $x_i'b$ have the same sign and 0 otherwise. Manski has proved that under mild conditions the (normalized) b which maximizes the sum of the scores is a consistent estimator of the true parameter vector β[1].

Our generalization is based on the fact that, as noted by Manski (1985), the MSCORE estimator is equivalent to a Least Absolute Deviations estimator, obtained by minimizing

$$\frac{1}{N} \sum_i |y_i - I(x_i'b>0)| \tag{2.2}$$

w.r.t. b. Here I(E)=1 if event E occurs and 0 otherwise.

Let $t(.):\mathbb{R}\to\mathbb{R}$ be a non-decreasing measurable function. If $Med(t(y^*)|x)$ is unique, there holds

$$Med(t(y^*)|x)=t(Med(y^*|x)) \tag{2.3}$$

So we have

$$Med(y_i|x_i)=Med(I(y_i^*>0)|x_i)=I(Med(y_i^*|x_i)>0)=I(x_i'\beta>0) \tag{2.4}$$

Thus $I(x_i'\hat{\beta}>0)$, with $\hat{\beta}$ being the (not necessarily unique) maximum score estimate, can be interpreted as an estimate of the median of $y_i|x_i$.

This median interpretation of the maximum score estimator suggests the following generalization to the ordered response case. Suppose that the dependent variable is grouped in M non-overlapping known[2)] intervals $(A_0,A_1]$, $(A_1,A_2]$,...,(A_{M-1},A_M) where $A_{m-1}<A_m$ for m=1,...,M; possibly, $A_0=-\infty$ and $A_M=\infty$. Define the non-decreasing transformation

$$z_i = \sum_{m=0}^{M-1} w_m I(y_i^*>A_m) \tag{2.5}$$

where w_m, m=0,...,M-1, are non-negative weights. Again using (2.3) it follows under certain conditions that

$$Med(z_i|x_i) = \sum_{m=0}^{M-1} w_m I(x_i'\beta>A_m) \tag{2.6}$$

Now the generalized maximum score estimator is obtained by minimizing

$$\frac{1}{N}\sum_i\Big|z_i-\sum_{m=1}^{M-1} w_m I(x_i'b>A_m)\Big| \tag{2.7}$$

with respect to b.

In the next section we prove that this estimator is strongly consistent.

3. Strong consistency of the generalized maximum score estimator

Our proof will be along the lines of the proof for the binary case given in Manski (1985).

Let us impose

Assumption 1. There exists a unique $\beta \in \mathbb{R}^K$ such that $\mathrm{Med}(y_i^* | x_i) = x_i'\beta$.

Assumption 2. a) The support of F_x is not contained in any proper linear subspace of $\mathbb{R}^K$. b) There exists at least one $k \in \{1, \ldots, K\}$ such that $\beta_k \neq 0$ and such that, for almost every value of $\tilde{x} = (x_1, x_2, \ldots, x_{k-1}, x_{k+1}, \ldots, x_K)$, the distribution of x_k conditional on $\tilde{x}$ has everywhere positive Lebesgue density.

Assumption 3. (y_i^*, x_i), $i=1,\ldots,N$ is a random sample from F_{yx}. For each i, (z_i, x_i) is observed.

Assumption 4. The bounds A_m, $m=0,1,\ldots,M$ are known constants and $M>2$. The weights w_m are non-negative, with at least two of them being strictly positive.

Assumption 5. The parameter space B is a compact subset of $\mathbb{R}^K$ and contains β.

Assumptions 1,2 and 3 are (almost[3]) identical to assumptions made in Manski (1985). Assumptions 4 and 5 distinguish the ordered response model from the binary response model. Note that assumption 5 is less stringent than in the binary case. There β is identified only up to scale so that the parameter space is taken to be the unit hypersphere in $\mathbb{R}^K$.

Define $g(v) = \sum_{m=0}^{M-1} w_m I(v > A_m) = \sum_{m=1}^{M} r_m I(A_{m-1} < v < A_m)$ where $r_m = \sum_{j=0}^{m-1} w_j$ for $m=1,\ldots,M$.

We first prove the following lemma.

Lemma 1 (Identification). Let $X_b = \{x \in \mathbb{R}^K;\ g(x'b) \neq g(x'\beta)\}$ and let $R(b) = \int_{X_b} dF_x$. Under assumptions 2 and 4, $R(b) > 0$ for all $b \in B$.

Proof. Let $k=K$ satisfy part b) of assumption 2 and consider the case in which $\beta_K>0$ (the case $\beta_K<0$ is symmetric). Without loss of generality we may assume that $A_{m_1}=0$ and $A_{m_2}=Q>0$ for some $m_1\neq m_2$; $m_1,m_2\in\{1,\ldots,M-1\}$. Let, for any $b\in\mathbb{R}^K$, $\tilde{b}=(b_1,\ldots,b_{K-1})'$ so $b=(\tilde{b}',b_K)'$. From Lemma 2 in Manski (1985) it follows immediately that $R(b)>0$ for all $b\in\mathbb{R}^K$ such that $b_K\leq 0$.

Now consider the case $b_K>0$. We can write

$$
\begin{aligned}
R(b) = {} & P[g(x'b)\neq g(x'\beta)] = \\
& P[x'b<0;0<x'\beta<Q] + P[x'b<0;x'\beta>Q] + \\
& P[0<x'b<Q;x'\beta<0] + P[0<x'b<Q;x'\beta>Q] + \\
& P[x'b>Q;x'\beta<0] \quad + P[x'b>Q;0<x'\beta<Q] = \\
= {} & P[x_K<-\tilde{x}'\tilde{b}/b_K;-\tilde{x}'\tilde{\beta}/\beta_K<x_K<(Q-\tilde{x}'\tilde{\beta})/\beta_K] + \\
& P[x_K<-\tilde{x}'\tilde{b}/b_K;x_K>(Q-\tilde{x}'\tilde{\beta})/\beta_K] + \\
& P[-\tilde{x}'\tilde{b}/b_K<x_K<(Q-\tilde{x}'\tilde{b})/b_K;x_K<-\tilde{x}'\tilde{\beta}/\beta_K] + \\
& P[-\tilde{x}'\tilde{b}/b_K<x_K<(Q-\tilde{x}'\tilde{b})/b_K;x_K>(Q-\tilde{x}'\tilde{\beta})/\beta_K] + \\
& P[x_K>(Q-\tilde{x}'\tilde{b})/b_K;x_K<-\tilde{x}'\tilde{\beta}/\beta_K] + \\
& P[x_K>(Q-\tilde{x}'\tilde{b})/b_K;-\tilde{x}'\tilde{\beta}/\beta_K<x_K<(Q-\tilde{x}'\tilde{\beta})/\beta_K]
\end{aligned}
\tag{3.1}
$$

Under part b) of assumption 2, at least one term on the r.h.s. of (3.1) is positive for almost any $\tilde{x}$ such that $-\tilde{x}'\tilde{\beta}/\beta_K\neq-\tilde{x}'\tilde{b}/b_K$ and/or $(Q-\tilde{x}'\tilde{\beta})/\beta_K\neq(Q-\tilde{x}'\tilde{b})/b_K$. However, we have $-\tilde{x}'\tilde{\beta}/\beta_K=-\tilde{x}'\tilde{b}/b_K$ and $(Q-\tilde{x}'\tilde{\beta})/\beta_K=(Q-\tilde{x}'\tilde{b})/b_K$ only for $\tilde{x}$ being in a (K-2)-dimensional subspace of $\mathbb{R}^{K-1}$. By part a) of assumption 2, the probability that $\tilde{x}$ is in such a subspace is less than one. Therefore, β is also identified relative to all b for which $b_K>0$. $\square$

The crucial difference with the binary case is that there $R(b)>0$ as long as $P[-\tilde{x}'\tilde{\beta}/\beta_K=-\tilde{x}'\tilde{b}/b_K]<1$. Since $P[-\tilde{x}'\tilde{\beta}/\beta_K=-\tilde{x}'\tilde{b}/b_K]=1$ if b is a scalar multiple of β, β is then identified only up to an arbitrary scale factor. In the ordered response case, however, $R(b)>0$ as long as $P[-\tilde{x}'\tilde{\beta}/\beta_K=-\tilde{x}'\tilde{b}/b_K$ and $(Q-\tilde{x}'\tilde{\beta})/\beta_K=(Q-\tilde{x}'\tilde{b})/b_K]<1$. Clearly, this is true even if b is a scalar multiple of β (except for $\tilde{x}$ being in a (K-2)-dimensional subspace of $\mathbb{R}^{K-1}$).

Theorem 1 (Strong consistency). Let assumptions 1 through 5 hold. Then the estimator defined by minimizing (2.7) with respect to b is strongly consistent for β.

Proof: The sample score function can be written as

$$S_N(b) = \frac{1}{N}\sum_i |z_i - g(x_i'b)| =$$

$$= \sum_{m=1}^{M-1}\sum_{j=m+1}^{M} |r_m-r_j|\{P_N(A_{m-1}<y^*<A_m;A_{j-1}<x'b<A_j) +$$

$$+ P_N(A_{j-1}<y^*<A_j;A_{m-1}<x'b<A_m)\} \qquad (3.2)$$

where P_N denotes the empirical probability distribution of (y_i^*,x_i). Similarly, we write the population score function as

$$S(b) = E|g(y^*)-g(x'b)| =$$

$$= \sum_{m=1}^{M-1}\sum_{j=m+1}^{M} |r_m-r_j|\{P(A_{m-1}<y^*<A_m;A_{j-1}<x'b<A_j) +$$

$$+ P(A_{j-1}<y^*<A_j;A_{m-1}<x'b<A_m)\} \qquad (3.3)$$

A. Under assumption 3, it follows by application of a Glivenko-Cantelli type result, similarly as in Lemma 4 in Manski (1985), that each right hand side term in the above decomposition of S_N converges to the corresponding term in the decomposition of S, uniformly in b. Hence,

$$S_N(b) \xrightarrow{a.s.} S(b)$$

uniformly over $b \in B \subset \mathbb{R}^K$.

B. Applying similar arguments as in Lemma 5 in Manski (1985), it follows under assumptions 1 and 2, that S(b) is continuous at all b for which $b_k \neq 0$.

C. We have to show that S(b) attains a unique global minimum at β. Using the fact that

$$E|g(y^*)-g(x'b)| =$$

$$\int_{\text{all } x} \left[\int_{\{t>g(x'b)\}} [1-2F_{z|x}(t)]dt + \int_{\mathbb{R}} F_{z|x}(t)dt \right] dF_x \tag{3.4}$$

this is equivalent to showing that the right hand side of (3.4) exceeds

$$\int_{\text{all } x} \left[\int_{\{t>g(x'\beta)\}} [1-2F_{z|x}(t)]dt + \int_{\mathbb{R}} F_{z|x}(t)dt \right] dF_x$$

for all $b \neq \beta$.

Consider, for a given x, the case where $x'\beta \neq A_m$ for all m, and $g(x'b) \neq g(x'\beta)$. If $g(x'b) > g(x'\beta))$, then

$$\int_{\{t>g(x'b)\}} [1-2F_{z|x}(t)]dt \quad > \int_{\{t>g(x'\beta)\}} [1-2F_{z|x}(t)]dt \tag{3.5}$$

will hold since $1-F_{z|x}(t)<0$ for $t>g(x'\beta)$. But if $g(x'b)<g(x'\beta)$, (3.5) will also hold since $1-F_{z|x}(t)>0$ for $t<g(x'\beta)$.

From assumption 2, it follows that, for all m, $P(x'\beta=A_m)$ has probability zero, whereas Lemma 1 guarantees that $X_b=\{x\in\mathbb{R}^K;\ g(x'b)\neq g(x'\beta)\}$ has positive probability. So, we conclude that $S(\beta)<S(b)$ for all $b\neq\beta$, $b\in B$.

Combining parts A, B and C and applying theorem 2 in Manski (1983), it follows that

$$P\{\lim_{N\to\infty} \sup_{b\in\hat{B}_N} \|b-\beta\| = 0\} = 1.$$

where $\hat{B}_N$ is the set of solutions to $\min_{b\in B} S_N(b)$. □

Footnotes

1) In Manski's 1975 article, the maximum score estimator was introduced for the more general multinomial discrete response model. However, the results in that paper do not carry over to the ordered response case. For example, the assumption in Manski (1975) of independence of the error terms across alternatives is clearly not appropriate here.

2) If the interval bounds are not known, β is identified only up to scale.

3) Manski (1985) makes the additional assumption $0<P(y\geq 0|x)<1$, a.e. F_x. This assumption (c.q. its analogue in the ordered response case) does not seem to be necessary. Suppose $P(A_{m-1}<y^{*}<A_m|x)=1$ a.e. F_x ($A_{m-1}\neq-\infty$ and $A_m\neq\infty$). Since $Med(y^{*}|x)=x'\beta$, it must be the case that $P(A_{m-1}<x'\beta<A_m|x)=1$. But this contradicts part b) of assumption 2 (following the same arguments as in the proof of Lemma 1).

References

Cramer, J.S. (1969), Empirical Econometrics, Amsterdam, North-Holland.

Kapteyn, A., P. Kooreman and R. Willemse (1988), "Some methodological issues in the implementation of subjective poverty definitions", The Journal of Human Resources, 23, pp. 222-242.

Manski, C.F. (1975), "Maximum score estimation of the stochastic utility model of choice", Journal of Econometrics, 3, pp.205-228.

Manski, C.F. (1983), "Closest empirical distribution estimation", Econometrica, 51, pp. 305-319.

Manski, C.F. (1985), "Semi-parametric analysis of discrete response; asymptotic properties of the maximum score estimator", Journal of Econometrics, 27, pp.313-333.

Robinson, P.M. (1988), "Semiparametric Econometrics: a Survey", Journal of Applied Econometrics, 3, pp. 35-51.

Stewart, M.B. (1983), "On Least Squares estimation when the dependent variable is grouped", Review of Economic Studies, 50, pp. 737-753.

Discussion Paper Series, CentER, Tilburg University, The Netherlands:

No.	Author(s)	Title
8801	Th. van de Klundert and F. van der Ploeg	Fiscal Policy and Finite Lives in Interdependent Economies with Real and Nominal Wage Rigidity
8802	J.R. Magnus and B. Pesaran	The Bias of Forecasts from a First-order Autoregression
8803	A.A. Weber	The Credibility of Monetary Policies, Policy-makers' Reputation and the EMS-Hypothesis: Empirical Evidence from 13 Countries
8804	F. van der Ploeg and A.J. de Zeeuw	Perfect Equilibrium in a Model of Competitive Arms Accumulation
8805	M.F.J. Steel	Seemingly Unrelated Regression Equation Systems under Diffuse Stochastic Prior Information: A Recursive Analytical Approach
8806	Th. Ten Raa and E.N. Wolff	Secondary Products and the Measurement of Productivity Growth
8807	F. van der Ploeg	Monetary and Fiscal Policy in Interdependent Economies with Capital Accumulation, Death and Population Growth
8901	Th. Ten Raa and P. Kop Jansen	The Choice of Model in the Construction of Input-Output Coefficients Matrices
8902	Th. Nijman and F. Palm	Generalized Least Squares Estimation of Linear Models Containing Rational Future Expectations
8903	A. van Soest, I. Woittiez, A. Kapteyn	Labour Supply, Income Taxes and Hours Restrictions in The Netherlands
8904	F. van der Ploeg	Capital Accumulation, Inflation and Long-Run Conflict in International Objectives
8905	Th. van de Klundert and A. van Schaik	Unemployment Persistence and Loss of Productive Capacity: A Keynesian Approach
8906	A.J. Markink and F. van der Ploeg	Dynamic Policy Simulation of Linear Models with Rational Expectations of Future Events: A Computer Package
8907	J. Osiewalski	Posterior Densities for Nonlinear Regression with Equicorrelated Errors
8908	M.F.J. Steel	A Bayesian Analysis of Simultaneous Equation Models by Combining Recursive Analytical and Numerical Approaches

No.	Author(s)	Title
8909	F. van der Ploeg	Two Essays on Political Economy (i) The Political Economy of Overvaluation (ii) Election Outcomes and the Stockmarket
8910	R. Gradus and A. de Zeeuw	Corporate Tax Rate Policy and Public and Private Employment
8911	A.P. Barten	Allais Characterisation of Preference Structures and the Structure of Demand
8912	K. Kamiya and A.J.J. Talman	Simplicial Algorithm to Find Zero Points of a Function with Special Structure on a Simplotope
8913	G. van der Laan and A.J.J. Talman	Price Rigidities and Rationing
8914	J. Osiewalski and M.F.J. Steel	A Bayesian Analysis of Exogeneity in Models Pooling Time-Series and Cross-Section Data
8915	R.P. Gilles, P.H. Ruys and J. Shou	On the Existence of Networks in Relational Models
8916	A. Kapteyn, P. Kooreman and A. van Soest	Quantity Rationing and Concavity in a Flexible Household Labor Supply Model
8917	F. Canova	Seasonalities in Foreign Exchange Markets
8918	F. van der Ploeg	Monetary Disinflation, Fiscal Expansion and the Current Account in an Interdependent World
8919	W. Bossert and F. Stehling	On the Uniqueness of Cardinally Interpreted Utility Functions
8920	F. van der Ploeg	Monetary Interdependence under Alternative Exchange-Rate Regimes
8921	D. Canning	Bottlenecks and Persistent Unemployment: Why Do Booms End?
8922	C. Fershtman and A. Fishman	Price Cycles and Booms: Dynamic Search Equilibrium
8923	M.B. Canzoneri and C.A. Rogers	Is the European Community an Optimal Currency Area? Optimal Tax Smoothing versus the Cost of Multiple Currencies
8924	F. Groot, C. Withagen and A. de Zeeuw	Theory of Natural Exhaustible Resources: The Cartel-Versus-Fringe Model Reconsidered

No.	Author(s)	Title
8925	O.P. Attanasio and G. Weber	Consumption, Productivity Growth and the Interest Rate
8926	N. Rankin	Monetary and Fiscal Policy in a 'Hartian' Model of Imperfect Competition
8927	Th. van de Klundert	Reducing External Debt in a World with Imperfect Asset and Imperfect Commodity Substitution
8928	C. Dang	The D_1-Triangulation of R^n for Simplicial Algorithms for Computing Solutions of Nonlinear Equations
8929	M.F.J. Steel and J.F. Richard	Bayesian Multivariate Exogeneity Analysis: An Application to a UK Money Demand Equation
8930	F. van der Ploeg	Fiscal Aspects of Monetary Integration in Europe
8931	H.A. Keuzenkamp	The Prehistory of Rational Expectations
8932	E. van Damme, R. Selten and E. Winter	Alternating Bid Bargaining with a Smallest Money Unit
8933	H. Carlsson and E. van Damme	Global Payoff Uncertainty and Risk Dominance
8934	H. Huizinga	National Tax Policies towards Product-Innovating Multinational Enterprises
8935	C. Dang and D. Talman	A New Triangulation of the Unit Simplex for Computing Economic Equilibria
8936	Th. Nijman and M. Verbeek	The Nonresponse Bias in the Analysis of the Determinants of Total Annual Expenditures of Households Based on Panel Data
8937	A.P. Barten	The Estimation of Mixed Demand Systems
8938	G. Marini	Monetary Shocks and the Nominal Interest Rate
8939	W. Güth and E. van Damme	Equilibrium Selection in the Spence Signaling Game
8940	G. Marini and P. Scaramozzino	Monopolistic Competition, Expected Inflation and Contract Length
8941	J.K. Dagsvik	The Generalized Extreme Value Random Utility Model for Continuous Choice