E C O N O M I C S B U L L E T I N

Heteroskedasticity, the single crossing property and ordered response models

Andreas C. Drichoutis
Agricultural University of Athens

Panagiotis Lazaridis
Agricultural University of Athens

Rodolfo M. Nayga, Jr.
Texas AMUniversity

Abstract

Heteroskedasticity in ordered response models has not garnered enough attention in the literature. Econometric software packages do not handle this problem satisfactorily either. We provide formulas to calculate heteroskedasticity corrected marginal effects and discrete changes using an approach that deals with single crossing property, a very restrictive assumption of ordered response models.

We wish to thank Ximing Wu for helpful comments on an earlier draft of this paper.
Citation: Drichoutis, Andreas C., Panagiotis Lazaridis, and Rodolfo M. Nayga, Jr., (2006) "Heteroskedasticity, the single crossing property and ordered response models." *Economics Bulletin*, Vol. 3, No. 31 pp. 1-6
Submitted: November 20, 2006. **Accepted:** November 24, 2006.
URL: http://economicsbulletin.vanderbilt.edu/2006/volume3/EB-06C10009A.pdf

1. Introduction

Ordered response models originated from the biometrics literature and their appearance in the social sciences is attributed to McKelvey and Zavoina (1975). Since then, many applications and extensions of these models have appeared in the economics literature. Most recent is Boes and Winkelmann's (2005) attempt to develop more flexible models that will overcome the restrictions inherent in standard models (i.e., the single crossing property - that is the signs of the marginal effects can only change once when moving from the smallest to the largest category).

Despite the growing popularity of ordered response models, the literature provides no clear way of accounting for heteroskedasticity in these models. Very few econometric packages can also account for heteroskedasticity, and even for those that could, the software generally cannot distinguish cases where heteroskedasticity is created by the same covariates that are included in the model. Consequently, this issue is often ignored in model estimation. Usually if one manages to derive heteroskedasticity corrected marginal effects, these are presented by softwares in two tables: one for the variable in the model and one for the variable in the heteroskedastic term. Clearly, the more appropriate way would be to account for the simultaneous variation of the variable in the model and in the heteroskedastic term.

Since the literature is limited and vague in this area, we show how to derive the appropriate formulas to account for heteroskedasticity in ordered response models. Furthermore, we also show that accounting for heteroskedasticity provides a more flexible analysis of marginal effects since the restrictive single crossing property vanishes.

2. The standard ordered response model

Assume there is a latent variable y^* ranging from $-\infty$ to $+\infty$ and is mapped to an observed variable y. The y variable is thought as providing incomplete information about the underlying y^* according to the measurement equation:

$$y_i = m \quad \text{if } \tau_{m-1} \le y_i^* < \tau_m \quad \text{for} \quad m=1 \text{ to } J \tag{1}$$

The τ 's are called thresholds and the extreme categories 1 and J are defined by open-ended intervals with $\tau_0 = -\infty$ and $\tau_J = +\infty$. The structural model is:

$$y_i^* = \mathbf{b'x}_i + \varepsilon_i \tag{2}$$

where $\mathbf{b}$ is a vector of structural coefficients. The observation mechanism results from a complete censoring of the latent dependent variable as follows:

$$\begin{aligned} y_i \ &= 0 \text{ if } y_i^* \le \tau_0, \\ &= 1 \text{ if } \tau_0 < y_i^* \le \tau_1, \\ &= 2 \text{ if } \tau_1 < y_i^* \le \tau_2, \\ &\cdots \\ &= J \text{ if } y_i^* \le \tau_{J-1}. \end{aligned} \tag{3}$$

Maximum likelihood estimation (ML) can be used to estimate the regression of y^* on $\mathbf{x}$. To use ML one has first to assume a specific form of the error distribution. The ordered probit model flows from the assumption that ε is distributed normally with mean 0 and variance 1, while the ordered logit model results from the assumption that ε has a logistic distribution with mean 0 and variance $\pi^2/3$.

Since y^* is not observable its interpretation is of no interest. The main focus in ordered data is on the conditional cell probabilities given by:

$$\Pr\left(y_i = m \mid \mathbf{x}_i\right) = F\left(\tau_m - \mathbf{b'x}_i\right) - F\left(\tau_{m-1} - \mathbf{b'x}_i\right) \tag{4}$$

where F represents either the standard normal distribution function or the logistic distribution. The parameter estimates from ordered response models, such as ordered probit and ordered logit, must be transformed to yield estimates of the marginal changes, that is, to determine how a marginal change in one regressor changes the distribution of all the outcome probabilities.

Taking the partial derivative of (4) with respect to x_k yields the marginal effect,

$$ME_{mk} = \frac{\partial \Pr\left(y_i = m \mid \mathbf{x}_i\right)}{\partial x_k} = \frac{\partial F\left(\tau_m - \mathbf{b'x}_i\right)}{\partial x_k} - \frac{\partial F\left(\tau_{m-1} - \mathbf{b'x}_i\right)}{\partial x_k} = \\ = b_k\left[f\left(\tau_{m-1} - \mathbf{b'x}_i\right) - f\left(\tau_m - \mathbf{b'x}_i\right)\right] \tag{5}$$

where $f\left(\beta\right) = dF\left(\beta\right)/d\beta$. Interpretation using the marginal effects can be misleading when an independent variable is a dummy variable. Hence, it is more appropriate to calculate the discrete change which is the change in the predicted probability for a change in x_k from the start value 0 to the end value 1.

$$DS_{mk} = \frac{\Delta \Pr\left(y_i = m \mid \mathbf{x}_i\right)}{\Delta x_k} = \Pr\left(y_i = m \mid \mathbf{x}_i, x_k = 1\right) - \Pr\left(y_i = m \mid \mathbf{x}_i, x_k = 0\right) \tag{6}$$

It is clear from (5) that marginal effects are first positive (negative) and then negative (positive) depending on the sign of b_k. In the case of discrete changes, the single crossing property is not mathematically clear but is intuitive, considering that equation (6) involves the difference between two probabilities that follow the bell-shaped distribution functions of the standard normal and logistic distribution.

3. Ordered response models with heteroskedasticity

In the case where the form of the heteroskedasticity adds no new parameters i.e. $Var\left(\varepsilon_i\right) = w_i^2$, then the procedure is very simple since (4) will be,

$$\Pr\left(y_i = m \mid \mathbf{x}_i\right) = F\left(\frac{\tau_m - \mathbf{b'x}_i}{w_i}\right) - F\left(\frac{\tau_{m-1} - \mathbf{b'x}_i}{w_i}\right) \tag{7}$$

And consequently (5) and (6) will yield:

$$ME_{mk} = \frac{\partial \Pr\left(y_i = m \mid \mathbf{x}_i\right)}{\partial x_k} = \frac{b_k}{w_i}\left[f\left(\frac{\tau_{m-1} - \mathbf{b'x}_i}{w_i}\right) - f\left(\frac{\tau_m - \mathbf{b'x}_i}{w_i}\right)\right] \tag{8}$$

$$DS_{mk} = \left[F\left(\frac{\tau_m - \mathbf{b'x}_i}{w_i}\right) - F\left(\frac{\tau_{m-1} - \mathbf{b'x}_i}{w_i}\right)\right]_{x_k = 1} - \left[F\left(\frac{\tau_m - \mathbf{b'x}_i}{w_i}\right) - F\left(\frac{\tau_{m-1} - \mathbf{b'x}_i}{w_i}\right)\right]_{x_k = 0} \tag{9}$$

In the case where heteroskedasticity adds an additional parameter vector (multiplicative heteroskedasticity), i.e. $Var\left(\varepsilon_i\right) = \left[\exp\left(\boldsymbol{\gamma}'\mathbf{z}_i\right)\right]^2$, things might be trickier. If the **z** vector contains no common variables with the **x** vector, then marginal effects and discrete changes can be calculated by (8) and (9) since $w_i = \exp\left(\boldsymbol{\gamma}'\mathbf{z}_i\right)$. However, if one or more variables in the **z** vector are common to the **x** vector, then the calculation of the marginal effects and discrete changes have to account for the fact that the variables of interest appear both in the nominator and the denominator of equations (8) and (9). Formally, assume that we can break the **z** vector into two vectors $\mathbf{z}_1$ and $\mathbf{x}_2$ where the $\mathbf{x}_2$ vector contains parameters common to **x**. We can then write $Var\left(\varepsilon_i\right) = \left[\exp\left(\boldsymbol{\gamma}_1'\mathbf{z}_{1i} + \boldsymbol{\gamma}_2'\mathbf{x}_{2i}\right)\right]^2$ and (4) will take the form:

$$\Pr\left(y_i = m \mid \mathbf{x}_i\right) = F\left(\frac{\tau_m - \mathbf{b'x}_i}{\exp\left(\boldsymbol{\gamma}_1\mathbf{'z}_{1i} + \boldsymbol{\gamma}_2\mathbf{'x}_{2i}\right)}\right) - F\left(\frac{\tau_{m-1} - \mathbf{b'x}_i}{\exp\left(\boldsymbol{\gamma}_1\mathbf{'z}_{1i} + \boldsymbol{\gamma}_2\mathbf{'x}_{2i}\right)}\right) \tag{10}$$

Taking the partial derivative of (10) with respect to x_k will yield the marginal effect,

$$\begin{aligned}
ME_{mk} &= \frac{\partial \Pr\left(y_i = m \mid \mathbf{x}_i\right)}{\partial x_k} = \\
&= f\left(\frac{\tau_m - \mathbf{b'x}_i}{\exp\left(\boldsymbol{\gamma}_1\mathbf{'z}_{1i} + \boldsymbol{\gamma}_2\mathbf{'x}_{2i}\right)}\right)\frac{-b_k \exp\left(\boldsymbol{\gamma}_1\mathbf{'z}_{1i} + \boldsymbol{\gamma}_2\mathbf{'x}_{2i}\right) + \left(\tau_m - \mathbf{b'x}_i\right)\gamma_{2k}\exp\left(\boldsymbol{\gamma}_1\mathbf{'z}_{1i} + \boldsymbol{\gamma}_2\mathbf{'x}_{2i}\right)}{\left(\exp\left(\boldsymbol{\gamma}_1\mathbf{'z}_{1i} + \boldsymbol{\gamma}_2\mathbf{'x}_{2i}\right)\right)^2} - \\
&- f\left(\frac{\tau_{m-1} - \mathbf{b'x}_i}{\exp\left(\boldsymbol{\gamma}_1\mathbf{'z}_{1i} + \boldsymbol{\gamma}_2\mathbf{'x}_{2i}\right)}\right)\frac{-b_k \exp\left(\boldsymbol{\gamma}_1\mathbf{'z}_{1i} + \boldsymbol{\gamma}_2\mathbf{'x}_{2i}\right) + \left(\tau_{m-1} - \mathbf{b'x}_i\right)\gamma_{2k}\exp\left(\boldsymbol{\gamma}_1\mathbf{'z}_{1i} + \boldsymbol{\gamma}_2\mathbf{'x}_{2i}\right)}{\left(\exp\left(\boldsymbol{\gamma}_1\mathbf{'z}_{1i} + \boldsymbol{\gamma}_2\mathbf{'x}_{2i}\right)\right)^2} = \\
&= \frac{-b_k}{\exp\left(\boldsymbol{\gamma}_1\mathbf{'z}_{1i} + \boldsymbol{\gamma}_2\mathbf{'x}_{2i}\right)}\left(f\left(\frac{\tau_m - \mathbf{b'x}_i}{\exp\left(\boldsymbol{\gamma}_1\mathbf{'z}_{1i} + \boldsymbol{\gamma}_2\mathbf{'x}_{2i}\right)}\right) - f\left(\frac{\tau_{m-1} - \mathbf{b'x}_i}{\exp\left(\boldsymbol{\gamma}_1\mathbf{'z}_{1i} + \boldsymbol{\gamma}_2\mathbf{'x}_{2i}\right)}\right)\right) + \\
&+ \frac{\gamma_{2k}}{\exp\left(\boldsymbol{\gamma}_1\mathbf{'z}_{1i} + \boldsymbol{\gamma}_2\mathbf{'x}_{2i}\right)}\left(\left(\tau_m - \mathbf{b'x}_i\right) f\left(\frac{\tau_m - \mathbf{b'x}_i}{\exp\left(\boldsymbol{\gamma}_1\mathbf{'z}_{1i} + \boldsymbol{\gamma}_2\mathbf{'x}_{2i}\right)}\right) - \left(\tau_{m-1} - \mathbf{b'x}_i\right) f\left(\frac{\tau_{m-1} - \mathbf{b'x}_i}{\exp\left(\boldsymbol{\gamma}_1\mathbf{'z}_{1i} + \boldsymbol{\gamma}_2\mathbf{'x}_{2i}\right)}\right)\right)
\end{aligned} \tag{11}$$

The formulas are much simpler in the case of discrete changes, since (6) will then be:

$$\begin{aligned}
DS_{mk} &= \left[F\left(\frac{\tau_m - \mathbf{b'x}_i}{\exp\left(\boldsymbol{\gamma}_1\mathbf{'z}_{1i} + \boldsymbol{\gamma}_2\mathbf{'x}_{2i}\right)}\right) - F\left(\frac{\tau_{m-1} - \mathbf{b'x}_i}{\exp\left(\boldsymbol{\gamma}_1\mathbf{'z}_{1i} + \boldsymbol{\gamma}_2\mathbf{'x}_{2i}\right)}\right)\right]_{x_k=1} - \\
&- \left[F\left(\frac{\tau_m - \mathbf{b'x}_i}{\exp\left(\boldsymbol{\gamma}_1\mathbf{'z}_{1i} + \boldsymbol{\gamma}_2\mathbf{'x}_{2i}\right)}\right) - F\left(\frac{\tau_{m-1} - \mathbf{b'x}_i}{\exp\left(\boldsymbol{\gamma}_1\mathbf{'z}_{1i} + \boldsymbol{\gamma}_2\mathbf{'x}_{2i}\right)}\right)\right]_{x_k=0}
\end{aligned} \tag{12}$$

Note that the difference between (12) and (9) is that the numerator will vary simultaneously with the denominator. Clearly, (11) and (12) no longer rely on a single crossing property. The sign in (11) is indeterminate. While (12) still involves the difference between cell probabilities, these are scaled up or down depending on the heteroskedastic terms. In all, if x_k is subset of **x** then one should use equations (8) and (9), but if x_k is subset of $\mathbf{x}_2$ then equations (11) and (12) are appropriate.

4. Empirical illustration

To illustrate how the presence of heteroskedasticity changes marginal effects and discrete changes, we use the data from Drichoutis *et al.* (2005). The purpose of the paper was to assess the effect of several variables on nutritional label use. Specifically, their paper investigated which factors may have an effect on how often consumers tend to read the on-pack nutrition information of food products when grocery shopping. Label use was measured on a four likert scale ranging from never, not often, often and always.

We define only a parsimonious form of the model since our purpose is to illustrate the estimation process with and without heteroskedasticity. Therefore, we will only assume that label use is affected by household size (*Hsize*), age (*Age*40, *Age*55, *Age*56), education ($Educ_2$) and income (Inc_2, Inc_3),

Label Use = *f* (*Hsize*, *Age*40, *Age*55, *Age*56, $Educ_2$, Inc_2, Inc_3) (13)

In (13), only household size is a continuous variable. All the other variables are dummies, indicating various age groups, educational levels and income levels. In Table 1, we present the results of the above estimation using an ordered probit model. The first half of the table presents the results where no heteroskedasticity is assumed. The second half of the table shows results where we assumed that household size and income are responsible for heteroskedasticity.

Table 1 shows that ignoring the presence of heteroskedasticity can be misleading in terms of magnitude but also in terms of the direction of the effect. Notice also that the single crossing property vanishes since the sign of some variables in Table 1 changes twice when moving from the smallest to the largest outcome category providing a more flexible way to analyze the data.

5. Conclusion

The need to account for heteroskedasticity in ordered response models is a problem, if present, since it can lead to erroneous results. That is, it can either overestimate or underestimate the true variance and, hence the standard errors may therefore be either understated or overstated. This issue becomes more relevant considering that known econometric software packages do not handle this problem satisfactorily. We provide formulas to calculate heteroskedasticity corrected marginal effects and discrete changes. We also show that our approach deals with a very restrictive assumption of the ordered response models (i.e., single crossing property).

Table 1. Marginal effects and discrete changes of demographic variables on frequency of reading nutritional labels

Variables	**No heteroskedasticity**				**Groupwise heteroskedasticity**			
	Never	Not often	Often	Always	Never	Not often	Often	Always
Hsize	0.014	0.004	-0.005	-0.013	-0.005	0.031	0.017	-0.042
*Age*40	-0.066	-0.021	0.020	0.067	-0.079	-0.047	0.045	0.081
*Age*55	-0.120	-0.041	0.033	0.128	-0.149	-0.096	0.076	0.169
*Age*56	-0.035	-0.011	0.010	0.036	-0.051	-0.032	0.028	0.054
$Educ_2$	-0.112	-0.032	0.036	0.108	-0.120	-0.062	0.072	0.111
INC_2	0.034	0.009	-0.012	-0.031	-0.003	0.072	0.025	-0.093
INC_3	0.103	0.025	-0.037	-0.091	0.104	0.068	-0.078	-0.123

References

Boes, S. and R. Winkelmann (2005) "Ordered response models", *Working paper no. 0507, Socioeconomic Institute, University of Zurich*.

Drichoutis, A. C., P. Lazaridis and R. M. Nayga, Jr. (2005) "Nutrition knowledge and consumer use of nutritional food labels" *European Review of Agricultural Economics* **32**, 1, 93-118.

McKelvey, R., and W. Zavoina (1975) :A statistical model for the analysis of ordered level dependent variables" *Journal of Mathematical Sociology* **4**, 103-120.