Estimation of nested and cross-nested zero-inflated ordered probit models

David Dale[*†]
Yandex, Moscow

Andrei Sirchenko
Higher School of Economics, Moscow

Abstract

We develop the maximum likelihood estimators and provide the STATA commands, `nop`, `ziop-2` and `ziop-3`, which estimate the three-part nested ordered probit model, the two-part cross-nested zero-inflated ordered probit models of Harris and Zhao (2007, *Journal of Econometrics* 141(2): 1073–1099) and Brooks, Harris and Spencer (2012, *Economics Letters* 117(3): 683–686), and the three-part cross-nested zero-inflated ordered probit model, with both exogenous and endogenous switching. The zero-inflated models address the inflation of neutral (zero) observations and allows zeros to emerge in two or three latent regimes. The three-part models allow the probabilities of positive and negative outcomes to be generated by distinct processes. We investigate the finite-sample performance of proposed estimators by Monte Carlo simulations, and illustrate the models with an empirical application to federal funds rate target.

Keywords: ordinal responses, nested ordered probit, zero inflation, endogenous switching, Monte Carlo simulation, federal funds rate target.

[*] We gratefully acknowledge support from the Basic Research Program of the Higher School of Economics.
[†] Corresponding author. Email:

1 Introduction

We introduce the STATA commands, `nop`, `ziop-2` and `ziop-3`, which estimate the two-level nested and cross-nested ordered-probit models including the zero- and middle-inflated models of Harris and Zhao (2007), Bagozzi and Mukherjee (2012) and Brooks, Harris and Spencer (2012). The rationale behind the two-level nested decision process is standard in the discrete-choice modeling when the set of alternatives faced by a decision-maker can be partitioned into subsets (or nests) with similar correlated alternatives. A choice among the nests and a choice among the alternatives within each nest can be driven by different sets of observed and unobserved factors (and common factors can have different weights). In the case of unordered categorical data, in which choices can be grouped into the nests of similar options, the nested logit model is a popular method.

The nested models for ordinal data are far more rare although the rationale behind such approach is also straigtforward: choosing among a decrease, no change or an increase is quite different from choosing between a small decrease or a large decrease; and choosing between a small decrease or a large decrease is quite different from choosing between a small increase or a large increase. This leads to three implicit decisions: a regime decision — a choice among the nests, and two outcome decisions — the choices of the magnitude of decreases and increases (see the top left panel of Figure 1).

Figure 1. Nested and cross-nested zero-inflated ordered probit models

TBA

Furthermore, it would be useful for the no-change (zero) alternative to be in three nests: its own one, another with decreases and another with increases; so some no-change decisions can be driven by similar factors as increases or decreases. This leads to a three-part cross-nested model with the nests overlapping at a zero response; hence, the probability of zeros is 'inflated'. Since the regime decision is not observable, the zeros are observationally equivalent — it is never known to which of the three nests the observed zero belongs. While several types of models with overlapping nests for unordered categorical responses are developed (Vovsha 1997; Wen and Koppelman 2001), the cross-nested models for ordinal outcomes are very scarce.[1]

The prevalence of status quo, neutral or zero outcomes is observed in many fields, including economics, sociology, technometrics, psychology and biology. The heterogeneity of

[1] Small (1987) introduced an ordered-choice model with overlapping nests, which contain two adjacent choices.

zeros is well recognized — see Winkelmann (2008) and Greene and Hensher (2010) for a review. Studies discriminate among different types of zeros such as: no visits to doctor due to good health, iatrophobia, or medical costs; no children due to infertility or choice; no illness due to strong resistance or lack of infection. In the studies of survey responses using an odd-point Likert-type scale, where the respondents must indicate the negative, neutral or positive attitude or opinion, the heterogeneity of indifferent responses (a true neutral option versus an undecided, or ambivalent, or uninformed one, commonly reported as neutral) is also well-recognized and sometimes labeled as the middle category endorsement or inflation (Bagozzi and Mukhetjee 2012, Hernández, Drasgow and Gonzáles-Romá 2004, Kulas and Stachowski 2009). In the decision-making experiments and micro-level studies of consumer choices, election votes and other repeated choices, the prevalence of no-change decisions is often attributed to the status quo bias – a tendency to do nothing or maintain one's previous decision, even though it is not always objectively superior to the available options (Hartman, Doane and Woo 1991, Kahneman, Knetsch, and Thaler 1991). It is a cognitive bias, explained by both the rational causes (informational or cognitive limitations, transition or analysis costs) and irrational ones (mental illusions and various psychological inclinations such as convenience, habit, inertia, fear, innate conservatism, loss aversion, and reputation concern) — see Samuelson and Zeckhauser (1988) for an excellent exposition.

The two-part zero-inflated models, developed to address the unobserved heterogeneity of zeros, combines a binary choice model for the probability of crossing the hurdle (to participate or not to participate; to consume or not to consume) with a count or ordered-choice model for nonnegative outcomes above the hurdle: the two parts are jointly estimated and the zero observations can emerge in both parts. The two-part zero-inflated models include the zero-inflated Poisson (Lambert 1992), negative binomial (Greene 1994), binomial (Hall 2002) and generalized Poisson (Famoye and Singh 2003) models for count outcomes, and the zero-inflated ordered probit model (Harris and Zhao 2007) and zero-inflated proportional odds model (Kelley and Anderson 2008) for non-negative ordinal responses.

The model of Harris and Zhao (2007) is suitable for explaining decisions such as the levels of consumption, when the upper hurdle is naturally binary (to smoke or not to smoke) and the ordinal responses are typically non-negative (see the top left panel of Figure 1). Thus, the abundant zeros are situated at one end of the ordered scale. Bagozzi and Mukherjee (2012) and Brooks, Harris and Spencer (2012) extended the model of Harris and Zhao (2007) and proposed the middle-inflated ordered probit model for an ordinal outcome, which ranges from negative to positive responses, and where an abundant outcome is situated in the middle of the choice spectrum (see the bottom left panel of Figure 1). The three-part cross-nested zero-inflated ordered probit model (see the bottom right panel of Figure 1), introduced in Sirchenko (2013), is a natural generalization of the models of Harris and Zhao (2007), Bagozzi and Mukherjee (2012) and Brooks, Harris and Spencer (2012). A trichotomous regime decision is more realistic and flexible than a binary decision (change or no change) if applied to ordinal data with negative, zero and positive values.

2 Models

The observed dependent variable y_t, $t = 1, 2, ..., T$ is assumed to take on a finite number of ordinal values j coded as $\{-J^-, ..., -1, 0, 1, ..., J^+\}$, where a typically predominant (and potentially heterogeneous) response is coded as zero. The latent unobserved (or only partially observed) variables are denoted by $*$. Each model assumes an ordered-choice regime decision and the ordered-choice outcome decisions conditional on each regime. The regime decision is allowed to be correlated with each outcome decision. Each decision is modeled by an ordered probit approach. We denote by $\mathbf{x}_t, \mathbf{x}_t^-, \mathbf{x}_t^+$ and $\mathbf{z}_t$ the t^{th} rows of the observed data matrices (which in addition to predetermined explanatory variables may also include the lags of y_t), by $\boldsymbol{\beta}, \boldsymbol{\beta}^-, \boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ the vectors of unknown slope parameters, by $\boldsymbol{\alpha}, \boldsymbol{\alpha}^-, \boldsymbol{\alpha}^+$ and $\boldsymbol{\mu}$ the vectors of unknown threshold parameters, by $\varepsilon_t, \varepsilon_t^-, \varepsilon_t^+$ and ν_t the error terms that are independently and identically distributed (iid) across t with normal cumulative distribution function (CDF) Φ and with variances $\sigma^2, \sigma_-^2.\sigma_+^2$, and σ_ν^2, respectively, and by $\Phi_2(g_1;g_2;\lambda)$ the CDF of the bivariate normal distribution of the two random variables g_1 and g_2 with the correlation coefficient λ and variances σ_1^2 and σ_2^2:

$$\Phi_2(g_1;g_2;\lambda) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\lambda^2}} \int_{-\infty}^{g_1} \int_{-\infty}^{g_2} \exp\left(-\frac{u^2/\sigma_1^2 - 2\lambda uw/\sigma_1\sigma_2 + w^2/\sigma_2^2}{2(1-\lambda^2)}\right) dudw.$$

Three-part nested ordered probit (NOP) model

Despite the wide-spread use of nested logit models for unordered categorical responses we are not aware of any example of the nested ordered probit/logit model for ordinal responses in the literature. The NOP model can be described as

$$
\begin{array}{lll}
\text{Regime decision:} & r_t^* = \mathbf{z}_t\boldsymbol{\gamma} + \nu_t, \quad s_t = \begin{cases} 1 & \text{if} \quad \mu_2 < r_t^*, \\ 0 & \text{if} \quad \mu_1 < r_t^* \le \mu_2, \\ -1 & \text{if} \quad r_t^* \le \mu_1. \end{cases} \\
\text{Outcome decisions:} & y_t^{-*} = \mathbf{x}_t^{-}\boldsymbol{\beta}^{-} + \varepsilon_t^{-}, \quad y_t^{+*} = \mathbf{x}_t^{+}\boldsymbol{\beta}^{+} + \varepsilon_t^{+}, \\
& y_t = \begin{cases} j(j>0) & \text{if} \quad s_t = 1 \text{ and } \alpha_{j-1}^{+} < y_t^{+*} \le \alpha_j^{+}, \\ 0 & \text{if} \quad s_t = 0, \\ j(j<0) & \text{if} \quad s_t = -1 \quad \text{and } \alpha_j^{-} < y_t^{-*} \le \alpha_{j+1}^{-}, \end{cases} \\
& \text{where } -\infty = \alpha_0^{+} \le \alpha_1^{+} \le ... \le \alpha_{J^+}^{+} = \infty \\
& \text{and } -\infty = \alpha_{-J^-}^{-} \le \alpha_{-J+1}^{-} \le ... \le \alpha_0^{-} = \infty. \\
\text{Correlation among decisions:} & \begin{bmatrix} \nu_t \\ \varepsilon_t^i \end{bmatrix} \overset{iid}{\sim} \mathcal{N}\left(\begin{matrix} 0 \\ 0 \end{matrix}, \begin{bmatrix} \sigma_\nu^2 & \rho_i\sigma_\nu\sigma_i \\ \rho_i\sigma_\nu\sigma_i & \sigma_i^2 \end{bmatrix} \right), i \in \{-,+\}.
\end{array}
$$

The probabilities of the outcome j in the NOP model are given by

$$
\begin{aligned}
&\Pr(y_t = j|\mathbf{z}_t, \mathbf{x}_t^{-}, \mathbf{x}_t^{+}) = I_{j<0}\Pr(r_t^* \le \mu_1 \text{ and } \alpha_j^{-} < y_t^{-*} \le \alpha_{j+1}^{-}|\mathbf{z}_t, \mathbf{x}_t^{-}) \\
&+I_{j=0}\Pr(\mu_1 < r_t^* \le \mu_2|\mathbf{z}_t) + I_{j>0}\Pr(\mu_2 < r_t^* \text{ and } \alpha_{j-1}^{+} < y_t^{+*} \le \alpha_j^{+}|\mathbf{z}_t, \mathbf{x}_t^{+}) \\
&= I_{j<0}\Pr(\nu_t \le \mu_1 - \mathbf{z}_t\boldsymbol{\gamma} \text{ and } \alpha_j^{-} - \mathbf{x}_t^{-}\boldsymbol{\beta}^{-} < \varepsilon_t^{-} \le \alpha_{j+1}^{-} - \mathbf{x}_t^{-}\boldsymbol{\beta}^{-}) \\
&+I_{j=0}\Pr(\mu_1 - \mathbf{z}_t\boldsymbol{\gamma} < \nu_t \le \mu_2 - \mathbf{z}_t\boldsymbol{\gamma}) \\
&+I_{j>0}\Pr(\mu_2 - \mathbf{z}_t\boldsymbol{\gamma} < \nu_t \text{ and } \alpha_{j-1}^{+} - \mathbf{x}_t^{+}\boldsymbol{\beta}^{+} < \varepsilon_t^{+} \le \alpha_j^{+} - \mathbf{x}_t^{+}\boldsymbol{\beta}^{+}) \\
&= I_{j<0}[\Phi_2(\mu_1 - \mathbf{z}_t\boldsymbol{\gamma}; \alpha_{j+1}^{-} - \mathbf{x}_t^{-}\boldsymbol{\beta}^{-}; \rho_-) - \Phi_2(\mu_1 - \mathbf{z}_t\boldsymbol{\gamma}; \alpha_j^{-} - \mathbf{x}_t^{-}\boldsymbol{\beta}^{-}; \rho_-)] \\
&+I_{j=0}[\Phi(\mu_2 - \mathbf{z}_t\boldsymbol{\gamma}) - \Phi(\mu_1 - \mathbf{z}_t\boldsymbol{\gamma})] \\
&+I_{j>0}[\Phi_2(-\mu_2 + \mathbf{z}_t\boldsymbol{\gamma}; \alpha_j^{+} - \mathbf{x}_t^{+}\boldsymbol{\beta}^{+}; -\rho_+) - \Phi_2(-\mu_2 + \mathbf{z}_t\boldsymbol{\gamma}; \alpha_{j-1}^{+} - \mathbf{x}_t^{+}\boldsymbol{\beta}^{+}; -\rho_+)],
\end{aligned} \tag{1}
$$

where $I_{j\le 0}$ is an indicator function such that $I_{j\le 0} = 1$ if $j \le 0$, and $I_{j\le 0} = 0$ if $j > 0$ (analogously for $I_{j=0}$ and $I_{j\le 0}$).

In the case of exogenous switching (when $\rho_- = \rho_+ = 0$), the probabilities of the outcome j in the NOP can be computed as

$$
\begin{aligned}
&\Pr(y_t = j|\mathbf{z}_t, \mathbf{x}_t^{-}, \mathbf{x}_t^{+}, \rho_- = \rho_+ = 0) \\
&= I_{j<0}\Phi(\mu_1 - \mathbf{z}_t\boldsymbol{\gamma})[\Phi(\alpha_{j+1}^{-} - \mathbf{x}_t^{-}\boldsymbol{\beta}^{-}) - \Phi(\alpha_j^{-} - \mathbf{x}_t^{-}\boldsymbol{\beta}^{-})] \\
&+I_{j=0}[\Phi(\mu_2 - \mathbf{z}_t\boldsymbol{\gamma}) - \Phi(\mu_1 - \mathbf{z}_t\boldsymbol{\gamma})] \\
&+I_{j>0}[1 - \Phi(\mu_2 - \mathbf{z}_t\boldsymbol{\gamma})][\Phi(\alpha_j^{+} - \mathbf{x}_t^{+}\boldsymbol{\beta}^{+}) - \Phi(\alpha_{j-1}^{+} - \mathbf{x}_t^{+}\boldsymbol{\beta}^{+})].
\end{aligned}
$$

In the case of three outcome categories the NOP model degenerates to the conventional ordered probit model.

Two-part cross-nested zero-inflated ordered probit (ZIOP-2) model

The ZIOP-2 model, which represents the middle-inflated ordered probit model of Bagozzi and Mukherjee (2012) and Brooks, Harris and Spencer (2012), can be described by the following system

$$
\begin{array}{ll}
\text{Regime decision:} & r_t^* = \mathbf{z}_t\boldsymbol{\gamma} + \nu_t, \quad s_t^* = \begin{cases} 1 & \text{if} \quad \mu < r_t^*, \\ 0 & \text{if} \quad r_t^* \leq \mu. \end{cases} \\
\text{Outcome decision:} & y_t^* = \mathbf{x}_t\boldsymbol{\beta} + \varepsilon_t, \\
& y_t = \begin{cases} j & \text{if} \quad s_t^* = 1 \text{ and } \alpha_{j-1} < y_t^* \leq \alpha_j, \\ 0 & \text{if} \quad s_t^* = 0, \end{cases} \\
& \text{where } -\infty = \alpha_{-J^- -1} \leq \alpha_{-J^-} \leq ... \leq \alpha_{J^+} = \infty. \\
\begin{array}{l}\text{Correlation among} \\ \text{decisions:}\end{array} & \begin{bmatrix} \nu_t \\ \varepsilon_t \end{bmatrix} \overset{iid}{\sim} \mathcal{N}\left(\begin{matrix} 0 \\ 0 \end{matrix}, \begin{bmatrix} \sigma_\nu^2 & \rho\sigma_\nu\sigma_\varepsilon \\ \rho\sigma_\nu\sigma_\varepsilon & \sigma_\varepsilon^2 \end{bmatrix} \right).
\end{array}
$$

The probabilities of the outcome j in the ZIOP-2 model are given by

$$
\begin{aligned}
&\Pr(y_t = j|\mathbf{z}_t, \mathbf{x}_t) = I_{j=0}\Pr(r_t^* \leq \mu|\mathbf{z}_t) \\
&+ I_{j\geq 0}\Pr(\mu < r_t^* \text{ and } \alpha_{j-1} < y_t^* \leq \alpha_j|\mathbf{z}_t, \mathbf{x}_t) \\
&= I_{j=0}\Pr(\nu_t \leq \mu - \mathbf{z}_t\boldsymbol{\gamma}) + I_{j\geq 0}\Pr(\mu - \mathbf{z}_t\boldsymbol{\gamma} < \nu_t \text{ and } \alpha_{j-1} - \mathbf{x}_t\boldsymbol{\beta} < \varepsilon_t \leq \alpha_j - \mathbf{x}_t\boldsymbol{\beta}) \\
&= I_{j=0}\Phi(\mu - \mathbf{z}_t\boldsymbol{\gamma}) + \Phi_2(-\mu + \mathbf{z}_t\boldsymbol{\gamma}; \alpha_j - \mathbf{x}_t\boldsymbol{\beta}; -\rho) - \Phi_2(-\mu + \mathbf{z}_t\boldsymbol{\gamma}; \alpha_{j-1} - \mathbf{x}_t\boldsymbol{\beta}; -\rho).
\end{aligned} \tag{2}
$$

In the case of exogenous switching (when $\rho = 0$), the probabilities of the outcome j in the ZIOP-2 model can be computed as

$$
\begin{aligned}
&\Pr(y_t = j|\mathbf{z}_t, \mathbf{x}_t, \rho = 0) = I_{j=0}\Phi(\mu - \mathbf{z}_t\boldsymbol{\gamma}) \\
&+ [1 - \Phi(\mu - \mathbf{z}_t\boldsymbol{\gamma})][\Phi(\alpha_j - \mathbf{x}_t\boldsymbol{\beta}) - \Phi(\alpha_{j-1} - \mathbf{x}_t\boldsymbol{\beta})].
\end{aligned}
$$

If $y_t \geq 0$ for $\forall t$, the ZIOP-2 model reduces to the model in Harris and Zhao (2007).

Three-part cross-nested zero-inflated ordered probit (ZIOP-3) model

The ZIOP-3 model, developed by Sirchenko (2013), generalizes the ZIOP-2 model and can be described by the following system

$$
\begin{array}{ll}
\text{Regime decision:} & r_t^* = \mathbf{z}_t\boldsymbol{\gamma} + \nu_t, \quad s_t^* = \begin{cases} 1 & \text{if} \quad \mu_2 < r_t^*, \\ 0 & \text{if} \quad \mu_1 < r_t^* \leq \mu_2, \\ -1 & \text{if} \quad r_t^* \leq \mu_1. \end{cases} \\
\text{Outcome decisions:} & y_t^{-*} = \mathbf{x}_t^-\boldsymbol{\beta}^- + \varepsilon_t^-, \quad y_t^{+*} = \mathbf{x}_t^+\boldsymbol{\beta}^+ + \varepsilon_t^+, \\
& y_t = \begin{cases} j (j \geq 0) & \text{if} \quad s_t^* = 1 \quad \text{and } \alpha_{j-1}^+ < y_t^{+*} \leq \alpha_j^+, \\ 0 & \text{if} \quad s_t^* = 0, \\ j (j \leq 0) & \text{if} \quad s_t^* = -1 \text{ and } \alpha_j^- < y_t^{-*} \leq \alpha_{j+1}^-, \end{cases} \\
& \text{where } -\infty = \alpha_{-1}^+ \leq \alpha_0^+ \leq ... \leq \alpha_{J^+}^+ = \infty \\
& \text{and } -\infty = \alpha_{-J^-}^- \leq \alpha_{-J+1}^- \leq ... \leq \alpha_1^- = \infty. \\
\begin{array}{l}\text{Correlation among} \\ \text{decisions:}\end{array} & \begin{bmatrix} \nu_t \\ \varepsilon_t^i \end{bmatrix} \overset{iid}{\sim} \mathcal{N}\left(\begin{matrix} 0 \\ 0 \end{matrix}, \begin{bmatrix} \sigma_\nu^2 & \rho_i\sigma_\nu\sigma_i \\ \rho_i\sigma_\nu\sigma_i & \sigma_i^2 \end{bmatrix} \right), i \in \{-, +\}.
\end{array}
$$

The probabilities of the outcome j in the ZIOP-3 model are given by

$$\begin{aligned}
&\Pr(y_t = j|\mathbf{z}_t, \mathbf{x}_t^-, \mathbf{x}_t^+) = I_{j\le 0}\Pr(r_t^* \le \mu_1 \text{ and } \alpha_j^- < y_t^{-*} \le \alpha_{j+1}^-|\mathbf{z}_t, \mathbf{x}_t^-) \\
&+I_{j=0}\Pr(\mu_1 < r_t^* \le \mu_2|\mathbf{z}_t) + I_{j\ge 0}\Pr(\mu_2 < r_t^* \text{ and } \alpha_{j-1}^+ < y_t^{+*} \le \alpha_j^+|\mathbf{z}_t, \mathbf{x}_t^+) \\
&= I_{j\le 0}\Pr(\nu_t \le \mu_1 - \mathbf{z}_t\boldsymbol{\gamma} \text{ and } \alpha_j^- - \mathbf{x}_t^-\boldsymbol{\beta}^- < \varepsilon_t^- \le \alpha_{j+1}^- - \mathbf{x}_t^-\boldsymbol{\beta}^-) \\
&+I_{j=0}\Pr(\mu_1 - \mathbf{z}_t\boldsymbol{\gamma} < \nu_t \le \mu_2 - \mathbf{z}_t\boldsymbol{\gamma}) \\
&+I_{j\ge 0}\Pr(\mu_2 - \mathbf{z}_t\boldsymbol{\gamma} < \nu_t \text{ and } \alpha_{j-1}^+ - \mathbf{x}_t^+\boldsymbol{\beta}^+ < \varepsilon_t^+ \le \alpha_j^+ - \mathbf{x}_t^+\boldsymbol{\beta}^+) \\
&= I_{j\le 0}[\Phi_2(\mu_1 - \mathbf{z}_t\boldsymbol{\gamma}; \alpha_{j+1}^- - \mathbf{x}_t^-\boldsymbol{\beta}^-;\rho_-) - \Phi_2(\mu_1 - \mathbf{z}_t\boldsymbol{\gamma}; \alpha_j^- - \mathbf{x}_t^-\boldsymbol{\beta}^-;\rho_-)] \\
&+I_{j=0}[\Phi(\mu_2 - \mathbf{z}_t\boldsymbol{\gamma}) - \Phi(\mu_1 - \mathbf{z}_t\boldsymbol{\gamma})] \\
&+I_{j\ge 0}[\Phi_2(-\mu_2 + \mathbf{z}_t\boldsymbol{\gamma}; \alpha_j^+ - \mathbf{x}_t^+\boldsymbol{\beta}^+; -\rho_+) - \Phi_2(-\mu_2 + \mathbf{z}_t\boldsymbol{\gamma}; \alpha_{j-1}^+ - \mathbf{x}_t^+\boldsymbol{\beta}^+; -\rho_+)],
\end{aligned} \tag{3}$$

where $I_{j\le 0}$ is an indicator function such that $I_{j\le 0} = 1$ if $j \le 0$, and $I_{j\le 0} = 0$ if $j > 0$ (analogously for $I_{j=0}$ and $I_{j\le 0}$).

In the case of exogenous switching (when $\rho_- = \rho_+ = 0$), the probabilities of the outcome j in the NOP can be computed as

$$\begin{aligned}
&\Pr(y_t = j|\mathbf{z}_t, \mathbf{x}_t^-, \mathbf{x}_t^+, \rho_- = \rho_+ = 0) \\
&= I_{j\le 0}\Phi(\mu_1 - \mathbf{z}_t\boldsymbol{\gamma})[\Phi(\alpha_{j+1}^- - \mathbf{x}_t^-\boldsymbol{\beta}^-) - \Phi(\alpha_j^- - \mathbf{x}_t^-\boldsymbol{\beta}^-)] \\
&+I_{j=0}[\Phi(\mu_2 - \mathbf{z}_t\boldsymbol{\gamma}) - \Phi(\mu_1 - \mathbf{z}_t\boldsymbol{\gamma})] \\
&+I_{j\ge 0}[1 - \Phi(\mu_2 - \mathbf{z}_t\boldsymbol{\gamma})][\Phi(\alpha_j^+ - \mathbf{x}_t^+\boldsymbol{\beta}^+) - \Phi(\alpha_{j-1}^+ - \mathbf{x}_t^+\boldsymbol{\beta}^+)].
\end{aligned}$$

The inflated outcome does not have to be in the *very* middle of ordered categories. If it is located at the *end* of the ordered scale, i.e. if $y_t \ge 0$ for $\forall t$, the ZIOP-3 model reduces to the ZIOP-2 model of Harris and Zhao (2007).

Maximum likelihood (ML) estimation

The probabilities in each ordered probit model representing the regime and outcome decisions can be consistently estimated under fairly general conditions by an asymptotically normal ML estimator (Basu and de Jong 2007). The simalteneous estimation of the ordered probit equations in the NOP, ZIOP-2 and ZIOP-3 models can be performed using an ML estimator of the vector of the parameters $\boldsymbol{\theta}$ that solves

$$\max_{\boldsymbol{\theta}\epsilon\boldsymbol{\Theta}} \sum_{t=1}^{T} \sum_{j=-J^-}^{J^+} q_{tj} \ln[\Pr(y_t = j|\mathbf{x}_t^{all}, \boldsymbol{\theta})], \tag{4}$$

where q_{tj} is an indicator function such that $q_{tj} = 1$ if $y_t = j$ and $q_{itj} = 0$ otherwise; $\boldsymbol{\theta}$ includes $\boldsymbol{\gamma}$, $\boldsymbol{\mu}$, $\boldsymbol{\beta}^-$, $\boldsymbol{\beta}^+$, $\boldsymbol{\alpha}^-$ and $\boldsymbol{\alpha}^+$ for the NOP model, $\boldsymbol{\gamma}$, $\boldsymbol{\mu}$, $\boldsymbol{\beta}$, $\boldsymbol{\alpha}$ and $\boldsymbol{\rho}$ for the ZIOP-2 model, and $\boldsymbol{\gamma}$, $\boldsymbol{\mu}$, $\boldsymbol{\beta}^-$, $\boldsymbol{\beta}^+$, $\boldsymbol{\alpha}^-$, $\boldsymbol{\alpha}^+$, $\boldsymbol{\rho}^-$ and $\boldsymbol{\rho}^+$ for the ZIOP-3 model; $\boldsymbol{\Theta}$ is a parameters' space; $\mathbf{x}_t^{all}$ is a vector that contains the values of all unique covariates from all equations of the model at observation t; and $\Pr(y_t = j|\mathbf{x}_t^{all}, \boldsymbol{\theta})$ are probabilities in either (1), or (3), or (2) for, respectively, the NOP, ZIOP-2 and ZIOP-3 models.

The intercept components of $\boldsymbol{\beta}$, $\boldsymbol{\beta}^-$, $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ are identified up to scale and location, that is only jointly with the corresponding threshold parameters $\boldsymbol{\alpha}$, $\boldsymbol{\alpha}^-$, $\boldsymbol{\alpha}^+$ and $\boldsymbol{\mu}$ and variances σ^2, σ^2_-, σ^2_+, and σ^2_ν. As is common in the identification of discrete choice models, the variances σ^2, σ^2_-, σ^2_+, and σ^2_ν are fixed to unit, and the intercept components of $\boldsymbol{\beta}$, $\boldsymbol{\beta}^-$, $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ are fixed to zero. The probabilities in (1), (2) and (3) are invariant to these (arbitrary) identifying assumptions. Up to scale and location, we can identify all parameters in $\boldsymbol{\theta}$ because of the nonlinearity of ordered probit equations, i.e. via the functional form (Heckman 1978; Wilde 2000). However, since the normal CDF is approximately linear in the middle of its support, the simultaneous estimation of two or three equations may experience a weak identification problem if all decision and outcome equations contain the same set of covariates. To enhance the precision of parameter estimates we may impose exclusion restrictions on the specification of covariates in each equation. The asymptotic standard errors of $\widehat{\boldsymbol{\theta}}$ can be computed from the Hessian matrix.

The three regimes (nests) in the NOP model are fully observable, contrary to the latent (only partially observed) regimes in the ZIOP-2 and ZIOP-3 models. The log of the likelihood function of the NOP model — again in contrast with the ZIOP-2 and ZIOP-3 models — is separable with respect to the parameters in the three equations. Thus, solving (4) for the NOP model is equivalent to maximizing separately the likelihoods of the three ordered probit models, representing the regime and outcome decisions in (**??**), if the data matrices in the outcome decisions are truncated to contain only those rows $\mathbf{x}_t^-$ or $\mathbf{x}_t^-$ for which $y_t < 0$ or $y_t > 0$, respectively.

Marginal effects

The marginal effect (ME) of a continuous covariate k (the k^{th} element of $\mathbf{x}_t^{all}$) on the probability of each discrete outcome j are computed for the ZIOP-3 model as

$$
\begin{aligned}
ME_{k,j,t} &= \tfrac{\partial \Pr(y_t=j|\boldsymbol{\theta})}{\partial \mathbf{x}_{t,k}^{all}} = I_{j\le 0}\left\{\left[\Phi\left(\tfrac{\mu_1-\mathbf{z}_t\boldsymbol{\gamma}-\rho_-(\alpha_j^- -\mathbf{x}_t^-\boldsymbol{\beta}^-)}{\sqrt{1-(\rho_-)^2}}\right) f(\alpha_j^- - \mathbf{x}_t^-\boldsymbol{\beta}^-)\right.\right.\\
&-\Phi\left(\tfrac{\mu_1-\mathbf{z}_t\boldsymbol{\gamma}-\rho_-(\alpha_{j+1}^- -\mathbf{x}_t^-\boldsymbol{\beta}^-)}{\sqrt{1-(\rho_-)^2}}\right) f(\alpha_{j+1}^- - \mathbf{x}_t^-\boldsymbol{\beta}^-)\Bigg]\boldsymbol{\beta}_k^{-all}\\
&-\left[\Phi\left(\tfrac{\alpha_{j+1}^- -\mathbf{x}_t^-\boldsymbol{\beta}^- -\rho_-(\mu_1-\mathbf{z}_t\boldsymbol{\gamma})}{\sqrt{1-(\rho_-)^2}}\right)-\Phi\left(\tfrac{\alpha_j^- -\mathbf{x}_t^-\boldsymbol{\beta}^- -\rho_-(\mu_1-\mathbf{z}_t\boldsymbol{\gamma})}{\sqrt{1-(\rho_-)^2}}\right)\right] f(\mu_1-\mathbf{z}_t\boldsymbol{\gamma})\boldsymbol{\gamma}_k^{all}\Bigg\}\\
&-I_{j=0}[f(\mu_2-\mathbf{z}_t\boldsymbol{\gamma})-f(\mu_1-\mathbf{z}_t\boldsymbol{\gamma})]\boldsymbol{\gamma}_k^{all}\\
&+I_{j\ge 0}\left\{\left[\Phi\left(\tfrac{\mathbf{z}_t\boldsymbol{\gamma}-\mu_2+\rho_+(\alpha_{j-1}^+ -\mathbf{x}_t^+\boldsymbol{\beta}^+)}{\sqrt{1-(\rho_+)^2}}\right) f(\alpha_{j-1}^+ - \mathbf{x}_t^+\boldsymbol{\beta}^+)\right.\right.\\
&-\Phi\left(\tfrac{\mathbf{z}_t\boldsymbol{\gamma}-\mu_2+\rho^+(\alpha_j^+ -\mathbf{x}_t^+\boldsymbol{\beta}^+)}{\sqrt{1-(\rho_+)^2}}\right) f(\alpha_j^+ - \mathbf{x}_t^+\boldsymbol{\beta}^+)\Bigg]\boldsymbol{\beta}_k^{+all}\\
&+\left[\Phi\left(\tfrac{\alpha_j^+ -\mathbf{x}_t^+\boldsymbol{\beta}^+ +\rho_+(\mathbf{z}_t\boldsymbol{\gamma}-\mu_2)}{\sqrt{1-(\rho_+)^2}}\right)-\Phi\left(\tfrac{\alpha_{j-1}^+ -\mathbf{x}_t^+\boldsymbol{\beta}^+ +\rho_+(\mathbf{z}_t\boldsymbol{\gamma}-\mu_2)}{\sqrt{1-(\rho_+)^2}}\right)\right] f(\mathbf{z}_t\boldsymbol{\gamma}-\mu_2)\boldsymbol{\gamma}_k^{all}\Bigg\},
\end{aligned}
$$

where f is the probability density function of the standard normal distribution, and $\boldsymbol{\gamma}_k^{all}$, $\boldsymbol{\beta}_k^{-all}$ and $\boldsymbol{\beta}_k^{+all}$ are the parameters on the k^{th} covariate in $\mathbf{x}_t^{all}$ in each equation ($\boldsymbol{\gamma}_k^{all}$, $\boldsymbol{\beta}_k^{-all}$ or $\boldsymbol{\beta}_k^{+all}$ is zero if the k^{th} covariate in $\mathbf{x}_t^{all}$ is not included into the corresponding equation). For a discrete-valued covariate, the ME can be computed as the change in the probabilities when this covariate changes by one increment and all other covariates are fixed.

The MEs for the NOP model are given by replacing $I_{j\ge 0}$ with $I_{j>0}$ and $I_{j\le 0}$ with $I_{j<0}$. The MEs for the ZIOP-2 model are computed as

$$
\begin{aligned}
ME_{k,j,t} &= \tfrac{\partial \Pr(y_t=j|\boldsymbol{\theta})}{\partial \mathbf{x}_{t,k}^{all}} = -I_{j=0}[f(\mu-\mathbf{z}_t\boldsymbol{\gamma})]\boldsymbol{\gamma}_k^{all}\\
&+\left[\Phi\left(\tfrac{\mathbf{z}_t\boldsymbol{\gamma}-\mu+\rho(\alpha_{j-1}-\mathbf{x}_t\boldsymbol{\beta})}{\sqrt{1-\rho^2}}\right) f(\alpha_{j-1}-\mathbf{x}_t\boldsymbol{\beta})-\Phi\left(\tfrac{\mathbf{z}_t\boldsymbol{\gamma}-\mu+\rho(\alpha_j-\mathbf{x}_t\boldsymbol{\beta})}{\sqrt{1-\rho^2}}\right) f(\alpha_j-\mathbf{x}_t\boldsymbol{\beta})\right]\boldsymbol{\beta}_k^{all}\\
&+\left[\Phi\left(\tfrac{\alpha_j-\mathbf{x}_t\boldsymbol{\beta}+\rho(\mathbf{z}_t\boldsymbol{\gamma}-\mu)}{\sqrt{1-\rho^2}}\right)-\Phi\left(\tfrac{\alpha_{j-1}-\mathbf{x}_t\boldsymbol{\beta}+\rho(\mathbf{z}_t\boldsymbol{\gamma}-\mu)}{\sqrt{1-\rho^2}}\right)\right] f(\mathbf{z}_t\boldsymbol{\gamma}-\mu)\boldsymbol{\gamma}_k^{all}.
\end{aligned}
$$

The asymptotic standard errors of the MEs are computed using the Delta method as the square roots of the diagonal elements of

$$
\widehat{Var(\widehat{\mathbf{ME}}_{k,j,t})} = \nabla_\theta \widehat{\mathbf{ME}}_{k,j,t}\widehat{Var(\widehat{\boldsymbol{\theta}})}\nabla_\theta \widehat{\mathbf{ME}}_{k,j,t})'.
$$

The relations among the models and their comparison

In this section I discuss the relations among the NOP, ZIOP-2 and ZIOP-3 models and the choice of a formal model-selection test, which depends on whether the models are nested in each other.

The exogenous-switching version of each model is nested in the endogenous-switching version of this model as its uncorrelated special case; their comparison can be performed using any classical lilkelihood-based test for nested hypotheses, such as the likelihood ratio (LR) test.

The NOP model is nested in the ZIOP-3 model. The latter becomes a NOP model if $\alpha_{-1}^{-} \to \infty$ and $\alpha_{1}^{+} \to -\infty$; therefore, $\Pr(y_t = 0|\mathbf{x}_t^{+}, s_t = 1) \to 0$ and $\Pr(y_t = 0|\mathbf{x}_t^{-}, s_t = -1) \to 0$. Thus, the comparison of the NOP and ZIOP-3 models can also be performed with the LR test; however, the critical values of the classical LR test are invalid since some standard regularity conditions of the classical LR test fail to hold. In particular, the values of α_{-1}^{-} and α_{1}^{+} in the null hypothesis are not the interior points of the parameter space; hence, the asymptotic distribution of the LR statistics is not standard. Instead, we may use the simulated critical values provided in Andrews (2001).

Generally, the ZIOP-2 model is not a special case of the ZIOP-3 model, and vice versa. However, they are not strictly non-nested and overlap if all their slope parameters are fixed to zeros. We can compare them using a likelihood-based test for non-nested overlapping models, such as the Vuong test (Vuong 1989). A special case when the ZIOP-3 model nests the ZIOP-2 model emerges under some restrictions on the parameters provided: (i) y_t only has three choices, (ii) the regressors in $\mathbf{x}_t^{-}$ and $\mathbf{x}_t^{+}$ in the outcome equations of the ZIOP-3 model contain all regressors in $\mathbf{z}_t$ in the ZIOP-2 regime equation, and (iii) the regressors in $\mathbf{z}_t$ in the regime equation of the ZIOP-3 model include all regressors in the $\mathbf{x}_t$ in the ZIOP-2 amount equation (see Appendix for the details). In this case, the selection between the ZIOP-3 and ZIOP-2 models can be performed using any classical likelihood-based test for nested hypotheses, which can be interpreted as a misspecification test for the latter.

3 The nop, ziop-2 and ziop-3 commands

Syntax

`ziop-3` *depvar indepvars* [*if*] [*in*] [, `zp(`*varlist*`)` `zn(`*varlist*`)` `infcat(`*integer* 0`)` `correlated` `cluster(`*varname*`)` `robust` `initial(`*string*`)`]

This command estimates by ML the three-part cross-nested zero-inflated ordered probit model with possibly different sets of covariates in the regime and outcome equations and possibly endogenous switching among three latent regimes.

`ziop-2` *depvar indepvars* [*if*] [*in*] [, `z (`*varlist*`)` `infcat(`*integer* 0`)` `correlated` `cluster(`*varna*
`robust` `initial(`*string*`)`]

This command estimates by ML the two-part cross-nested zero-inflated ordered probit model with possibly different sets of covariates in the regime and outcome equations and possibly endogenous switching among two latent regimes.

`nop` *depvar indepvars* [*if*] [*in*] [, `zp(`*varlist*`)` `zn(`*varlist*`)` `infcat(`*integer* 0`)` `correlated` `cluster(`*varname*`)` `robust` `initial(`*string*`)`]

This command estimates by ML the three-part nested ordered probit model with possibly different sets of covariates in the regime and outcome equations and possibly endogenous switching among three latent regimes..

Options

options	Description
`zp(`*varlist*`)`	list of covariates for positive response in NOP and CNOP models; by default, it equals *indepvars*, the list of covariates for initial stage
`zn(`*varlist*`)`	list of covariates for negative response in NOP and CNOP models; by default, it equals *indepvars*, the list of covariates for initial stage
`z(`*varlist*`)`	list of covariates for non-zero response in ZIOP models; by default, it equals *indepvars*, the list of covariates for initial stage
`infcat(`*integer*`)`	value of the response variable that should be modeled as inflated; by default, it equals 0
`correlated`	flag that errors in the first and second stages may be correlated, forcing estimation of CNOPc, NOPc or ZIOPc model
`robust`	flag that variance-covariance estimator must be robust (based on "sandwich") estimate
`cluster(`*varname*`)`	clustering variable for robust variance estimator
`initial(`*string*`)`	whitespace-delimited list of initial parameter values for estimation, in the following order: β, α, γ^+, μ^+, γ^-, μ^-, ρ^-, ρ^+

Examples

TBD

Stored results

`nop`, `ziop-2`, and `ziop-3` store the following in `e()`:

`e(N)`	number of observations
`e(cmd)`	`cnop`, `nop`, or `miop`, respectively
`e(depvar)`	dependent variable of regression
`e(b)`	parameters vector
`e(V)`	variance-covariance matrix
`e(sample)`	marks estimation sample

Postestimation commands

The predict command

The `predict` command after the nop, ziop-2 and ziop-3 estimation commands produces either predicted probabilities or expected values of the responses.

`predict` *varname* [*if*] [*in*] `[, zeroes regime output(`*string*`) at(`*string*`)]`

`name` is the name of predicted variable, if it is single, or prefix for names, if there are several predicted variables

`zeroes` indicates that different types of zeroes (i.e. “intrinsic zeroes“, or “positive zeroes“, or “negative zeroes“) must be predicted instead of different response values.

`regime` indicates that different groups of response (negative, positive or zero) must be predicted instead of different response values. This option is ignored if `zeroes` option is on.

`output(string)` specifies type of aggregating predicted probabilities of different response. Possible values are `mode` and `mean`, for predicting average or most probable outcome, and `cum` for predicting cumulative response probabilities (i.e. $\Pr(y_t <= -2)$, $\Pr(y_t <= -1)$, $\Pr(y_t <= 0)$ etc.). If not specified, raw response probabilities are predicted ($\Pr(y_t = -2)$, $\Pr(y_t = -1)$, $\Pr(y_t = 0)$ etc.).

The cnopmargins command

`cnopmargins [, at(`*string*`) nominal(`*varlist*`) zeroes regime]`

This command prints marginal effects for the last estimated model (either NOP, or ZIOP-2, or ZIOP-3), calculated at the specified point, along with confidence intervals.

`at(string)` specifies at which point predictions must be calculated. If at is specified, (as a list of `varname=value` expressions, separated by comma), prediction is calculated at this point and posted on the screen without saving to the dataset. If some covariate names are not specified, their mean value is taken instead.

`nominal` is a space-separated list of covariates which should be considered as nominal; marginal effect is then calculated as difference between values at 0 and at 1.

`zeroes` and `regime` indicate that marginal effects should be calculated for different zeroes or for groups of response variable, as in `predict` command.

The cnopprobabilities command

`cnopprobabilities [, at(`*string*`) zeroes regime]`

This command prints predicted probabilities for the last estimated model (either NOP, or ZIOP-2, or ZIOP-3) , calculated at the specified point, along with confidence intervals. The point `at` is specified like in `cnopmargins`.

The cnopcontrasts command

`cnopcontrasts [, at(`*string*`) to(`*string*`) zeroes regime]`

This command prints differences in predicted probabilities for the last estimated model (either NOP, or ZIOP-2, or ZIOP-3), calculated between the specified points, along with confidence intervals. The points `at` and `to` are specified like `at` in `cnopmargins`.

Examples

TBD

4 Monte Carlo simulations

We conducted extensive Monte Carlo experiments to illustrate the finite sample performance of the ML estimators of each model.

Monte Carlo design

We simulated six processes generated by the NOP, ZIOP-2 and ZIOP-3 models with both exogenous and endogenous switching. The repeated samples with 200, 500 and 1000 observations were independently generated and then estimated by the true model. The number of replications was 10,000 in each experiment.

Three covariates $\mathbf{w_1}$, $\mathbf{w_2}$ and $\mathbf{w_3}$ were drawn in each replication as $\mathbf{w_1} \overset{iid}{\sim} \mathcal{N}(0,1)+2$, $\mathbf{w_2} \overset{iid}{\sim} \mathcal{N}(0,1)$, and $\mathbf{w_3} = -1$ if $\mathbf{u} \leq 0.3$, 0 if $0.3 < \mathbf{u} \leq 0.7$, or 1 if $\mathbf{u} > 0.7$, where $\mathbf{u} \overset{iid}{\sim} \mathcal{U}[0,1]$. The repeated samples were generated for the NOP and ZIOP-2 DGPs with $\mathbf{Z} = (\mathbf{w_1}, \mathbf{w_2})$, $\mathbf{X}^- = (\mathbf{w_1}, \mathbf{w_3})$, $\mathbf{X}^+ = (\mathbf{w_2}, \mathbf{w_3})$, and for the ZIOP-2 DGP with $\mathbf{Z} = (\mathbf{w}_1, \mathbf{w}_3)$, $\mathbf{X} = (\mathbf{w}_2, \mathbf{w}_3)$. The dependent variable y was generated with five outcome categories: -2, -1, 0, 1 and 2. The parameters were calibrated to yield on average the following frequencies of the above outcomes: 7%, 14%, 58%, 14% and 7%, respectively. To avoid the divergence of ML estimates due to the problem of complete separation (perfect prediction), which could happen if actual number of observations in any outcome category is very low, the samples with any outcome category frequency lower than 6% were re-generated. The matrix of the MEs has $3 \times 5 = 15$ elements; their values, which depend on the values of the regressors, are computed at the population medians of the covariates. The variances of the errors in the regime and outcome equations were fixed to one. The true values of all other parameters in the simulations are shown in Table 1. The simulations and estimations were performed using the MATA programming language. The starting values for slope and threshold parameters were obtained using the independent ordered-probit estimations of each equation. The starting values for ρ, ρ_- and ρ_+ were obtained by maximizing the logarithms of the likelihood functions of the endogenous-switching models holding the other parameters fixed at their estimates in the corresponding exogenous-switching model.

Table 1. Monte Carlo simulations: The true values of parameters

TBA

Monte Carlo results

Table 2 reports the measures of accuracy for the estimates of the probabilities and MEs. For each model, the bias and RMSE decrease as sample size increases. RMSE decreases in most cases faster than asymptotic rate $\sqrt{n}$. This may be caused by a small number of large deviations in parameter estimation in small samples. For most of models and sample sizes, the bias and RMSE are slightly higher for the endogenous-switching version. This is expected from a more complex model, estimated with the same sample size.

Standard error estimates for parameters on average correspond to the actual standard errors. Large deviations make standard errors estimates biased, especially on small samples, but this problem rapidly decreases as sample size grows. Anyway, rare large deviations do not prevent asymptotic coverage probabilities of 95% confidence intervals from being consistent. In general, results of Monte Carlo simulations show that estimators of the proposed nested and cross-nested ordered probit models are consistent, but should be used carefully in small samples. As a rule of thumb, we would advise using at least 10 observations per variable in each outcome class, which corresponds to 1000 observations in our case (make it more optimistic!!!).

Table 2. Monte Carlo results: The accuracy of ML estimators

Sample size	True and estimated model:	NOP (exog)	NOP	ZIOP-2 (exog)	ZIOP-2	ZIOP-3 (exog)	ZIOP-3
				Probabilities			
200	Bias, %	2.3	1.5	4.4	5.1	3.3	3.1
500		1.1	0.9	2.3	3.0	1.6	1.5
1000		0.4	0.4	1.3	1.7	0.8	1.0
200	RMSE, ×100	2.4	2.6	2.8	2.9	2.7	2.9
500		1.5	1.6	1.7	1.8	1.6	1.8
1000		1.1	1.1	1.2	1.2	1.1	1.3
200	Coverage probability (at 95% level), %	94.4	94.4	95.3	95.3	95.1	94.8
500		95.4	95.2	95.6	95.6	95.9	95.7
1000		95.5	95.5	95.7	95.7	95.6	95.6
200	Bias of standard error estimates, %	4.2	4.2	6.9	6.4	5.5	15.1
500		3.9	4.6	6.9	6.1	5.3	16.6
1000		2.6	3.4	5.7	5.9	3.7	13.9
				Marginal effects on probabilities			
200	Bias, %	4.5	4.1	10.5	16.9	11.5	23.0
500		1.7	2.2	4.9	7.2	5.5	9.7
1000		0.8	1.3	2.5	3.7	2.6	5.3
200	RMSE, ×100	1.8	2.1	2.5	3.6	2.8	3.4
500		1.1	1.3	1.5	2.3	1.7	2.0
1000		0.8	0.9	1.0	1.5	1.2	1.4
200	Coverage probability (at 95% level), %	95.8	93.9	91.7	87.9	94.6	91.8
500		95.9	94.6	94.8	91.5	95.0	93.0
1000		95.6	95.0	95.3	93.9	95.1	93.9
200	Bias of standard error estimates, %	4.7	5.7	8.0	6.1	21.4	39.1
500		4.0	5.0	5.8	6.0	27.0	8.1
1000		2.4	3.4	4.2	5.7	11.6	7.4

Notes: (exog) – exogenous switching; Bias – the absolute difference between the estimated and true values, devided by the true value; RMSE – the absolute root mean square error of the estimates; Coverage probability – the percentage of times the estimated asymptotic 95% confidence intervals cover the true values; Bias of standard error estimates – the absolute difference between the average of the estimated asymptotic standard errors of the estimates and the standard deviation of the estimates in all replications. The above measures of accuracy for the estimates of the probabilities are averaged across five outcome categories, and for the estimates of the MEs are averaged across five outcome categories and across all covariates.

5 Empirical application

The existing applications of discrete-choice approach to monetary policy rules (e.g., Hamilton and Jorda, 2002; Hu and Phillips, 2004; Dolado et al., 2005; Piazzesi, 2005; Basu and de Jong, 2007; Kauppi, 2012; Van den Hauwe et al., 2013) do not allow for a regime-switching behavior of central bank.

TBA

Concluding remarks

References

Andrews, D. W. K. 2001. Testing when a parameter is on the boundary of the maintained hypothesis. *Econometrica* 69 (3): 683–734.

Bagozzi, B. E., and B. Mukherjee. 2012. A mixture model for middle category inflation in ordered survey responses. *Political Analysis* 20: 369–386.

Basu, D., and R. M. de Jong. 2007. Dynamic multinomial ordered choice with an application to the estimation of monetary policy rules. *Studies in Nonlinear Dynamics and Econometrics* 11 (4): 1–35.

Brooks, R., M. N. Harris, and C. Spencer. 2012. Inflated ordered outcomes. *Economics Letters* 117 (3): 683–686.

Famoye, F., and K. P. Singh. 2003. On inflated generalized Poisson regression models. *Advanced Applied Statistics* 3 (2): 145–158.

Greene, W. H. 1994. Accounting for excess zeros and sample selection in Poisson and negative binomial regression models. Working Paper No. 94-10, Department of Economics, Stern School of Business, New York University.

Greene, W. H., and D. A. Hensher. 2010. *Modeling ordered choices: A primer*. Cambridge University Press.

Harris, M. N., and X. Zhao. 2007. A zero-inflated ordered probit model, with an application to modelling tobacco consumption. *Journal of Econometrics* 141 (2): 1073–1099.

Hartman, R. S., M. Doane, and C.-K. Woo. 1991. Consumer rationality and the status quo. *Quarterly Journal of Economics* 106: 141–162.

Heckman, J. J. 1978. Dummy endogenous variables in a simultaneous equation system. *Econometrica* 46: 931–959.

Hernández, A., F. Drasgow and V. Gonzáles-Romá. 2004. Investigating the functioning of a middle category by means of a mixed-measurement model. *Journal of Applied Psychology* 89 (4): 687-699.

Kahneman, D., J. L. Knetsch, and R. H. Thaler. 1991. Anomalies: the endowment effect, loss aversion, and status quo bias. *Journal of Economic Perspectives* 5 (1): 193–206.

Kelley, M. E., and S. J. Anderson. 2008. Zero inflation in ordinal data: incorporating susceptibility to response through the use of a mixture model. *Statistics in Medicine* 27: 3674–3688.

Kulas, J. T. and A. A. Stachowski. 2009. Middle category endorsement in odd-numbered Likert response scales: Associated item characteristics, cognitive demands, and preferred meanings. *Journal of Research in Personality* 43: 489-493.

Lambert, D. 1992. Zero-inflated Poisson regression with an application to defects in manufacturing. *Technometrics* 34 (1): 1–14.

MacKinnon, J. G. 1996. Numerical distribution functions for unit root and cointegration tests. *Journal of Applied Econometrics* 11: 601–618.

Samuelson, W., and R. Zeckhauser. 1988. Status quo bias in decision making. *Journal of Risk and Uncertainty* 1: 7–59.

Sirchenko, A. 2013. A model for ordinal responses with an application to policy interest rate. National Bank of Poland Working Paper No. 148.

Small, K. 1987. A discrete choice model for ordered alternatives. *Econometrica* 55: 409–424.

Vovsha, P. 1997. Application of cross-nested logit model to mode choice in Tel Aviv, Israel, Metropolitan Area. *Transportation Research Record* 1607: 6–15.

Vuong, Q. 1989. Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica* 57 (2): 307–333.

Wen, C.-H., and F. Koppelman. 2001. The generalized nested logit model. *Transportation Research B* 35: 627–641.

Wilde, J. 2000. Identification of multiple equation probit models with endogenous dummy regressors. *Economics Letters* 69 (3): 309–312.

Winkelmann, R. 2008. *Econometric analysis of count data.* 5$^{\text{th}}$ edition, Springer.