A model for ordinal responses with an application to policy interest rate

JOB MARKET PAPER

Andrei Sirchenko*
European University Institute, Florence, Italy
Email: andrey.sirchenko@eui.eu
http://www.eui.eu/Personal/Researchers/Sir/Index.html

December 9, 2012

Abstract

The decisions to reduce, leave unchanged, or increase (the price, rating, policy interest rate, etc.) are often characterized by abundant no-change outcomes that are generated by different processes. Moreover, the positive and negative responses can also be driven by distinct forces. To capture the heterogeneity of the data-generating process this paper develops a two-stage cross-nested model, combining three ordered probit equations. In the policy rate setting context, the first stage, a policy inclination decision, determines a latent policy stance (loose, neutral or tight), whereas the two latent amount decisions, conditional on a loose or tight stance, fine-tune the rate at the second stage. The model allows for the possible correlation among the three latent decisions. This approach identifies the driving factors and probabilities of three types of zeros: the "neutral" zeros, generated directly by a neutral policy stance, and two kinds of "offset" zeros, the "loose" and "tight" zeros, generated by a loose or tight stance, offset at the second stage. Monte Carlo experiments show good performance in small samples. Both the simulations and empirical applications to the panel data on individual policymakers' votes for the interest rate demonstrate the superiority with respect to the conventional and two-part models.

JEL classification: C33; C35; E52.

Keywords: ordinal responses; zero-inflated outcomes; three-part model; cross-nested model; policy interest rate; MPC votes; real-time data; panel data.

---

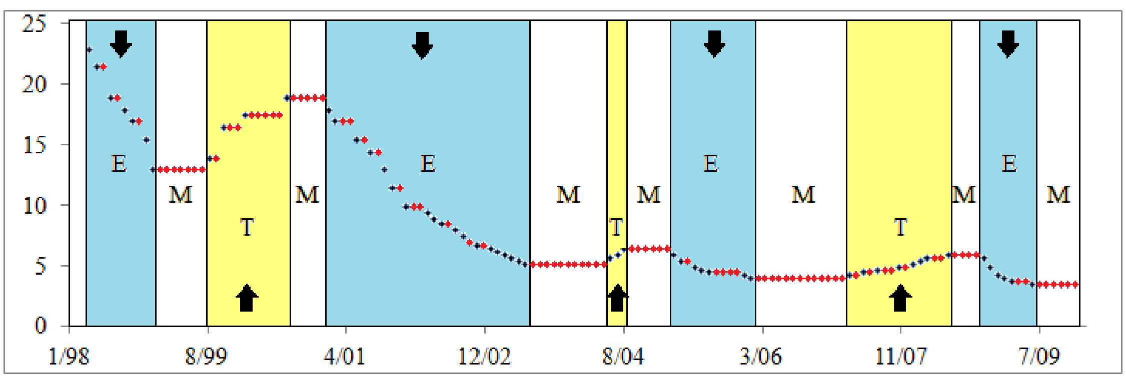
*This paper is based on research supported by a grant from the Global Development Network (via Economics Education and Research Consortium's grant competition) as well as a grant from the National Bank of Poland. I am also thankful to Jérôme Adda, Michael Beenstock, James D. Hamilton, Mark Harris, Peter R. Hansen, Helmut Lütkepohl, Massimiliano Marcellino, Chiara Monfardini, Simon van Norden, Dobromil Serwa, Grzegorz Szafrański, Francis Vella, as well as to participants of MOOD 2012 in Rome and seminars at National Bank of Poland, New Economic School in Moscow, and University of Bologna for useful discussions and comments on earlier versions of this paper.

1 Introduction

Ordinal dependent variables, taking on negative, zero and positive values, are often characterized by the abundant and potentially heterogeneous observations in the middle (neutral or zero) category. For instance, most central banks adjust policy rates by discrete increments – namely multiples of 25 basis points – and no-change decisions commonly constitute an absolute majority (e.g., 63, 66, 76 and 79 percent in the US Federal Reserve, National Bank of Poland, Bank of England and European Central Bank, respectively)[1]. As Figure 1 shows, the policy rate of the National Bank of Poland (NBP) remained unchanged during three different circumstances: namely, during policy tightening; during maintaining (between the reversals); and during periods of easing. Many of "zeros" that are clustered between the reversals during the maintaining periods, are likely to be driven by different forces than many of those that are situated between the changes in the same direction during periods of policy tightening or easing.

To illustrate this, Table 1 reports the average values of macroeconomic indicators, observed separately only during policy decisions to either increase, reduce or leave the rate unchanged. In the no-change case, these values are reported separately for periods of policy tightening, maintaining and easing. The economic conditions, observed on average when the rates were not changed during the tightening/easing periods, are much closer to those, observed when the rates were increased/reduced, than to those that prevailed on average when the rates were maintained between the reversals. On the other hand, some of the no-change decisions during the maintaining periods occurred under the economic circumstances, similar to those observed during some decisions to hike or cut the rates.

Figure 1: The reference rate of the National Bank of Poland



Notes: E/M/T denote the periods of policy easing/maintaining/tightening.

These stylized facts suggest that no-change decisions can be generated by different decision-making processes. In addition, the positive and negative changes may be also driven by distinct determinants. This definitely poses a problem for a standard discrete-choice model such as the ordered probit (OP) or logit model. In such situations, it would be a misspecification to disregard the heterogeneity of zeros, to treat all the observations as coming from the same data-generating process (d.g.p.), and to apply a conventional model,

[1]During the 6/1997-10/2012, 1/1999-10/2012, 10/1982-10/2012 and 3/1998-10/2012 periods, respectively.

based on a single equation. This paper develops a three-equation cross-nested ordered probit (CNOP) model for such types of ordinal outcomes, and illustrates the model in the context of policy interest rate decisions.

Table 1: Economic conditions observed on average at different policy rate decisions during the periods of policy easing, maintaining and tightening

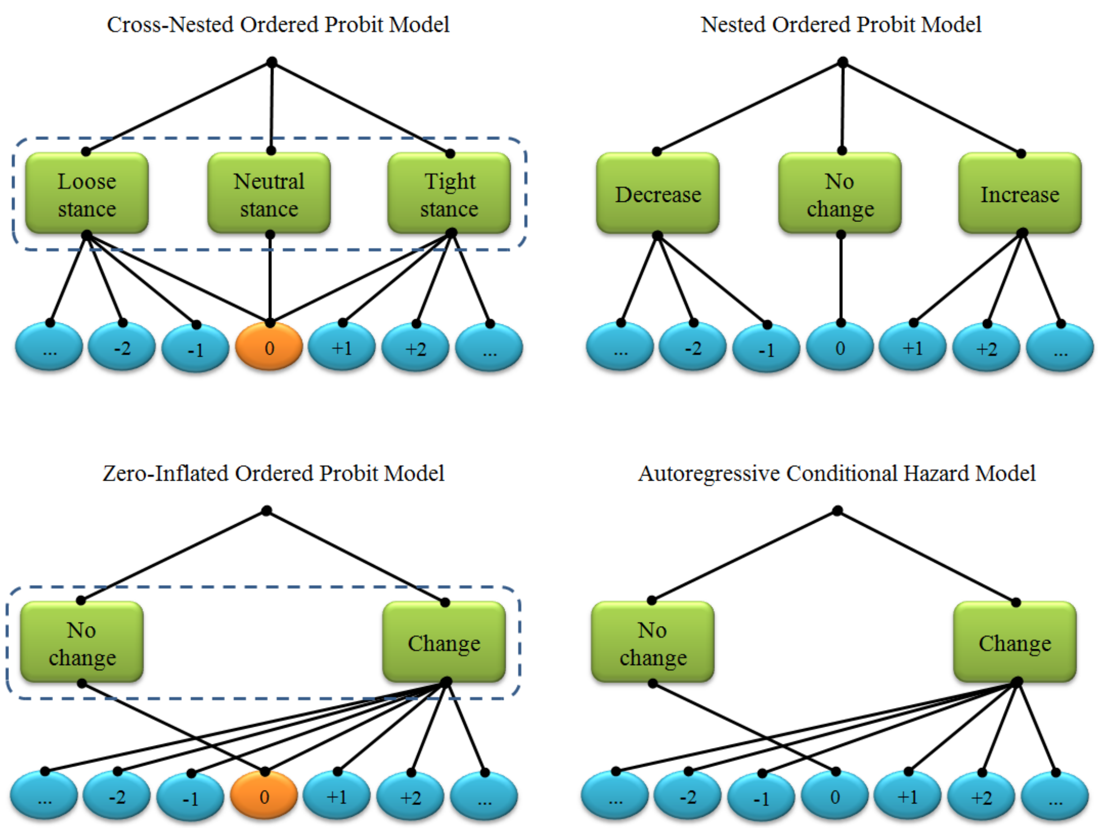
Policy period	Policy rate decision	cpi^e-tar	$\Delta(cpi^e$-$tar)$	*situation*
Easing	Reduce	-0.37	-0.23	8.93
	No change	-0.62	-0.20	8.29
Maintaining	No change	**0.34**	**-0.02**	**15.15**
Tightening	No change	2.41	0.21	19.80
	Increase	1.70	0.23	20.41

Notes: Sample period: 02/1998 – 12/2009; *situation* - index of expected general economic situation in industry from Business Tendency Survey; $cpi^e - tar$ - deviation of expected CPI over next 12 months (Ipsos-Demoskop survey of consumers) from the NBP target; Δ - recent monthly change.

Suppose that an ordinal dependent variable – for example, a discrete change to policy rate – can be in three latent regimes (loose, neutral or tight), where it can take on only nonpositive, zero or nonnegative values, respectively. The upper left panel of Figure 2 shows a decision tree. The first stage, a policy *inclination* decision, sets the regime, i.e. monetary policy stance. The inclination decision is driven by a direct reaction to the economic conditions, particularly to the developments since the last policy meeting. At the second stage, if the stance is neutral, no further policy actions are taken and the rate is maintained. If the stance is loose (tight), the policymakers can cut (hike) the rate by certain amount or may leave it unchanged. These two *amount* decisions, conditional on either loose or tight policy stances, fine-tune the rate and are more of a tactical and institutional nature. The model allows for the possible correlation among the three latent decisions. Under this interpretation, we can classify three kinds of zeros and describe how they arise: the "always" or "neutral" zeros, generated directly by neutral policy reactions to economic conditions; and two kinds of "not-always" or "offset" zeros, the "loose" and "tight" zeros, generated by loose or tight policy regimes, offset by tactical and institutional reasons.

The existence of different types of no-change decisions is justified by the very nature of monetary policymaking, which involves processing huge amounts of data, meeting different and sometimes conflicting goals, and which is often conducted by a committee composed of heterogeneous members, as well as by the discrete nature of the interest rate changes themselves. For example, despite a loose policy stance, the policymakers can maintain the rate due to the following reasons. First, the recent "policy bias" statement of the central bank, which indicates the most likely policy direction in the immediate future, was neutral or even tightening (this addresses the policymakers' concerns about the competence and credibility of the central bank's communication). Second, the dissenting policymakers at the last meeting preferred the higher rate, creating an upward pressure to the rate at the current meeting (this accounts for the fact that the monetary policy is commonly conducted by a committee, often composed of heterogeneous members)[2]. Third, the rate was already lowered at the last meeting (this reflects the general reluctance to move the rate frequently). Fourth, the cumulative changes to the economic indicators since the date of the last non-zero policy rate adjustment do not suggest the policy easing (the policymakers, who face uncertainty about the economy and incur the costs in the case of the subsequent rate reversal, prefer to wait and to react to more accumulated economic information in order to minimize the risk of the reversals). Finally, the policy rate has already reached the lower zero bound.

Figure 2: Decision trees of the CNOP, NOP, ZIOP and ACH models



As we shall observe, the proposed three-equation models are fairly easy to estimate via maximum likelihood. The Monte Carlo results suggest good performance of the proposed cross-nested ordered probit (CNOP) models in the small samples and demonstrate its superiority with respect to the conventional OP model producing biased estimates if the d.g.p. is heterogeneous. The new models are applied to explain the policy interest rate decisions of the NBP, using a panel of the individual votes of the Monetary Policy Council (MPC)

[2]See, for example, Gerlach-Kristen (2004) and Sirchenko (2010), who documented that the dissenting views of policymakers at the last policy meeting help predict the next policy decision of the Bank of England and National Bank of Poland, respectively.

members and *real-time* macroeconomic data available at MPC meetings. The individual policy preferences of the policymakers appeared to be well-modelled by a new approach. The empirical application demonstrates the advantages of the new models in separating different decision-making paths for three types of zeros, identifying the determinants of policy decisions and estimating the marginal effects of explanatory variables on the predicted probabilities.

The CNOP model is able to identify the driving factors of each decision and estimate the probabilities of the latent policy regimes and three types of zeros. As a practical matter, this allows certain variables to affect the inclination and amount decisions differently; hence, the probabilities of positive, negative and three types of zero outcomes may be driven by different sources. The model estimates how the decomposition of no-change decisions depends on the observed data, and sheds additional light on monetary policy inertia. Only about a quarter of observed zeros appeared to be generated by the neutral policy stance. This finding suggests a high degree of gradualism and deliberate interest-rate smoothing in the decision-making process of the NBP. The conventional OP models, based on a single latent equation, are shown to confuse the marginal effects of the explanatory variables that have an impact only on one decision or opposing impacts on both decisions. Besides, the marginal effects of the explanatory variables reveal the non-monotonic relationships between these variables and choice probabilities. The standard OP models overlook such non-monotonic patterns.

The proposed CNOP model is related to three strands of econometric literature. On the one hand, it can be described as a two-level cross-nested ordered probit model, an extension of a two-level nested ordered probit (NOP) model with three nests (see upper right panel of Figure 2). At the upper level of the NOP model the policymakers decide whether to increase, maintain, or decrease the rate. This trilemma is modelled by a trichotomous OP model. In case of a no-change decision, no further policy actions are taken, and the rate remains unchanged. If the policymakers decide to hike or to cut the rate, they have to choose the amount of the change. This fine-tuning lower level, conditional on the decision to increase or decrease the rate at the upper level, is modeled by two distinct OP models. Overall, the NOP model combines three equations with, in general, different sets of covariates. Therefore, in contrast to a standard single-equation OP model, in the NOP model, one set of explanatory variables may be relevant for the rate cuts, while another set may be relevant for the hikes. The third set of covariates would affect the no-change decisions. In the CNOP model the three nests overlap – they all contain the zero outcomes. It creates three distinct d.g.p's, generating zero observations.

Notice also another key difference between the NOP and CNOP models: in the former both levels' decisions are observable, whereas in the latter they are observed partially, only when the outcome is nonzero. In the CNOP model the outcomes in the inflated zero category are observationally equivalent – we never know from which of the three regimes the zeros arise, whereas in the NOP model we always know to which of the three nests the observed outcomes belong. In this sense the three regimes in the CNOP model are latent.

In case of the unordered categorical data where the choices can be grouped into nests of similar options, the nested logit model is used widely. Several kinds of multinomial logit models with overlapping nests have been also proposed. Wen and Koppelman (2001) introduced a generalized nested logit model, which contains the other cross-nested logit models as special cases. The nested and cross-nested models, specifically designed for the ordered alternatives, are not used so widely[3].

On the other hand, the CNOP model can be perceived as a three-part middle-category-inflated mixture model. The two-part mixture models, developed to deal with both the abundant zeros and unobserved heterogeneity, include the zero-inflated Poisson (Lambert 1992) and negative binomial (Greene 1994) models for count outcomes, as well as the zero-inflated ordered probit (ZIOP) model (Harris and Zhao 2007) and zero-inflated proportional odds model (Kelley and Anderson 2008) for ordinal variables. These zero-inflated models are the natural extensions of the two-part (or hurdle, or split-population) models, first proposed by Cragg (1971) for non-negative continuous data, and then developed for the count data (Mullahy 1986), survival time data (Schmidt and Witte 1989), and discrete ordered time-series data (the autoregressive conditional hazard (ACH) model of Hamilton and Jorda 2002). A two-part model basically represents a two-level model with two nests. It combines a binary outcome model for the probability of crossing the hurdle (the upper-level *participation* decision) with a truncated-at-zero model for the outcomes above the hurdle (the lower-level *amount* decision). The difference between the two-part ACH and ZIOP models (see bottom panels of Figure 2) is that in the former the two parts are estimated separately, the zero observations are excluded from the second part, and, hence, the discrimination among different kinds of zeros is not accommodated, whereas in the latter the two nests overlap, assuming two types of zeros, and, hence, the probability of zeros is "inflated". The ZIOP model is able to identify the different d.g.p's of two kinds of zeros[4]. Hamilton and Jorda (2002) applied the ACH model to the changes to the Federal funds rate target, made by the US Federal Open Market Committee; Brooks et al. (2012) applied the ZIOP model to the panel data on the changes to the policy rate, preferred by each member of the Bank of England's Monetary Policy Committee.

The three-part CNOP model is a natural generalization of the two-part ZIOP model. A trichotomous participation decision (increase versus no change versus decrease) seems to be more realistic than a binary one (change versus no change) if applied to such types of ordinal data – the policymakers, who are willing to adjust the rate, have naturally already decided in which direction they want to move it. Combining these two distinct decisions at the upper hurdle into one category, as done in the ZIOP model, may seriously distort the inference. The same explanatory variable can have different weights in the decisions to increase or reduce the rate. Besides, the CNOP model allows the probabilities and magnitudes of the positive changes to the rate to be affected by different determinants than those of the negative changes. The ZIOP model is more suitable if applied to explain such decisions as, for example, the levels of consumption, when the upper hurdle is naturally binary (to consume or not to consume).

Finally, the two-part model is similar by structure to a discrete version of the sample selection model[5]. However, in the sample selection model the first hurdle, the *selection* decision, determines whether the outcome variable is *observed*, rather than whether the activity is *undertaken*, as in the two-part model, where *all* the outcomes are actually *observed*. In many applications, in the absence of the sample selection problem, there is no need in modeling the latent potential, as opposed to the observed actual outcomes, but there is a

[3] Small (1987) proposed a model for ordered outcomes, called the ordered generalized extreme value model, that has overlapping nests.

[4] On the other hand, the ZIOP model assumes no serial correlation among the latent residuals, whereas the ACH model accounts for the serial dependence in discrete-valued time series.

[5] The early contributions are Gronau (1974) and Heckman (1976 and 1979), among others.

need to model the "corner solution" outcomes or address the heterogeneity instead[6].

The NOP and CNOP econometric frameworks are introduced in the next section (including their extended versions, the NOPC and CNOPC models, where the mechanisms determining the three decisions are dependent). Section 3 briefly reports the results of Monte Carlo simulations to assess and compare the finite sample performance of the OP, NOP(C) and CNOP(C) models, as well as the performance of the *LR* and *Vuong* tests and model selection criteria. In Section 4 the six alternative models – OP, multinomial probit, generalized OP, ZIOP, CNOP and CNOPC – are applied to explain policy interest rate decisions of the NBP, using a panel of the individual votes of the MPC members and *real-time* macroeconomic data available at policy meetings. Section 5 concludes. The Online Appendix contains three parts: Appendix A with the details of Monte Carlo design, Appendix B with the details of Monte Carlo results and Appendix C with supplemental output from empirical application.

2 The econometric framework

The CNOP model allows for any number of ordered discrete categories of the dependent variable greater than two, while the NOP model degenerates to the standard OP model in case of three outcome categories. For ease of exposition and without loss of generality, the observed dependent variable is assumed to take on a finite number of discrete values j coded as $\{-J, ..., -1, 0, 1, ..., J\}$, and the inflated neutral outcome is coded as zero[7].

The proposed models are suitable for the large survey data, both cross-sectional and longitudinal, though a sufficiently long discrete-valued time series is also applicable. Since in this paper the models are applied to the panel data, the econometric framework is presented in the panel context using double subscript, where the index i denotes one of N cross-sectional units and index t denotes one of T time periods. The application to the pure cross-sectional or time-series data is straightforward by setting N or T to one.

Each observation is treated as an independent draw from the population both along the cross-sectional and time-series dimensions. Thus, it is assumed that the cross-sectional units are independent, that the model specification is dynamically complete, hence, there is no serial correlation among the latent errors[8].

2.1 The cross-nested ordered probit (CNOP) model

Let $r_{it} = \{-1, 0, 1\}$ be a trichotomous latent variable that determines whether the individual policy stance is loose, neutral or tight, and let m_{it}^- and m_{it}^+ be the discrete nonpositive and nonnegative latent variables that set the magnitude of Δy_{it}, conditional on $r_{it} = -1$ and

[6] For a debate between the sample selection and two part-models see Leung and Yu (1996), Jones (2000), Dow and Norton (2003), Madden (2008).

[7] Of course, the inflated outcome does not have to be in the *very* middle of ordered categories. If the inflated outcome is at the end of the ordered scale, the three-part CNOP model reduces to the two-part ZIOP model.

[8] The treatments of spatial effects (that are quite reasonably expected in the panel with small N) and serial autocorrelation of the disturbance terms are among the possible extentions of the model. For example, the CNOP model can be further extended by allowing for the serial correlation among the latent residuals and employing the dynamic OP specifications (Eichengreen, Watson and Grossman, 1985) of three latent equations, estimated via the Gibbs sampler.

$r_{it} = 1$, respectively. Then assume that the observed vote for a change to policy rate Δy_{it} is generated as

$$\Delta y_{it} = \frac{|r_{it}|}{2}\left\{(1-r_{it})\, m_{it}^{-} + (1+r_{it}) m_{it}^{+}\right\} = \begin{cases} m_{it}^{-} & \text{if} \quad r_{it} = -1, \\ 0 & \text{if} \quad r_{it} = 0, \\ m_{it}^{+} & \text{if} \quad r_{it} = 1. \end{cases}$$

Notice that r_{it} is observed only if $\Delta y_{it} \neq 0$, while m_{it}^{-} and m_{it}^{+} are observed only if $\Delta y_{it} < 0$ or $\Delta y_{it} > 0$, respectively. Conditional on a set of explanatory variables, we will assume further that the mechanisms generating r_{it}, m_{it}^{-} and m_{it}^{+} are either independent or dependent.

The model assumes two stages and three regimes, and includes three OP latent equations. At the first stage (the upper level of the decision tree – see the top left panel of Figure 2) there is a continuous latent variable r_{it}^{*}, representing the magnitude of the policymaker i's policy stance and set at a meeting t in response to the observed data according to a policy *inclination* equation

$$r_{it}^{*} = \mathbf{x}_{it}'\boldsymbol{\beta} + \nu_{it}, \tag{1}$$

where $\mathbf{x}_{it}$ is the t^{th} row of an observed $T_i \times K_\beta$ data matrix $\mathbf{X}_i$, T_i is the number of observations available for the individual i, $\boldsymbol{\beta}$ is a $K_\beta \times 1$ vector of unknown coefficients, and ν_{it} is an error term, independently and identically distributed (i.i.d.) across i and t.

The regime-setting decision r_{it} is coded as -1, 0, or 1, if the policymaker i's policy stance is loose, neutral or tight, respectively. The correspondence between r_{it}^{*} and r_{it} is given by the matching rule

$$r_{it} = \begin{cases} -1 & \text{if} \qquad r_{it}^{*} \leq \alpha_1, \\ 0 & \text{if} \quad \alpha_1 < r_{it}^{*} \leq \alpha_2, \\ 1 & \text{if} \quad \alpha_2 < r_{it}^{*}, \end{cases}$$

where $-\infty < \alpha_1 \leq \alpha_2 < \infty$ are unknown threshold parameters to be estimated.

Under the assumption that the disturbance term ν_{it} is distributed with the cumulative distribution function (c.d.f.) F, the probabilities of each possible outcome of r_{it} are:

$$\begin{array}{lll} \Pr(r_{it} = -1|\mathbf{x}_{it}) & = \Pr(r_{it}^{*} \leq \alpha_1|\mathbf{x}_{it}) & = F(\alpha_1 - \mathbf{x}_{it}'\boldsymbol{\beta}), \\ \Pr(r_{it} = 0|\mathbf{x}_{it}) & = \Pr(\alpha_1 < r_{it}^{*} \leq \alpha_2|\mathbf{x}_{it}) & = F(\alpha_2 - \mathbf{x}_{it}'\boldsymbol{\beta}) - F(\alpha_1 - \mathbf{x}_{it}'\boldsymbol{\beta}), \\ \Pr(r_{it} = 1|\mathbf{x}_{it}) & = \Pr(\alpha_2 < r_{it}^{*}|\mathbf{x}_{it}) & = 1 - F(\alpha_2 - \mathbf{x}_{it}'\boldsymbol{\beta}). \end{array} \tag{2}$$

At the second stage (the lower level of the decision tree) there are three regimes and two latent *amount* equations.

- *Regime* $r_{it} = -1$ *(loose policy stance).*

Conditional on being in regime $r_{it} = -1$ the continuous latent variable m_{it}^{-*}, representing the desired change to the rate, is determined by the *amount* equation

$$m_{it}^{-*} = \mathbf{z}_{it}^{-\prime}\boldsymbol{\gamma} + \varepsilon_{it}^{-}, \tag{3}$$

where $\boldsymbol{\gamma}$ is a $K_\gamma \times 1$ vector of unknown coefficients, $\mathbf{z}_{it}^-$ is the t^{th} row of an observed $T_i \times K_\gamma$ data matrix $\mathbf{Z}_i^-$, and ε_{it}^- is an i.i.d. error term with the c.d.f. F^-.

The discrete change to the rate m_{it}^- is determined according to the rule-

$$m_{it}^- = j \text{ if } \mu_{j-1}^- < y_{it}^{-*} \leq \mu_j^- \text{ for } j = -J \text{ to } 0,$$

where $-\infty = \mu_{-J-1}^- \leq \mu_{-J}^- \leq ... \leq \mu_{-1}^- \leq \mu_0^- = \infty$ are J unknown thresholds to be estimated.

The conditional probability of a particular outcome j is given by

$$\Pr(m_{it}^- = j|\mathbf{z}_{it}^-, r_{it} = -1) = \begin{cases} F^-(\mu_{-J}^- - \mathbf{z}_{it}^{-\prime}\boldsymbol{\gamma}) & \text{for} \quad j = -J, \\ F^-(\mu_j^- - \mathbf{z}_{it}^{-\prime}\boldsymbol{\gamma}) - F^-(\mu_{j-1}^- - \mathbf{z}_{it}^{-\prime}\boldsymbol{\gamma}) & \text{for} \quad -J < j < 0, \\ 1 - F^-(\mu_{-1}^- - \mathbf{z}_{it}^{-\prime}\boldsymbol{\gamma}) & \text{for} \quad j = 0, \\ 0 & \text{for} \quad 0 < j \leq J, \end{cases}$$

which can be written more compactly, given that $-\infty = \mu_{-J-1}^-$ and $\mu_0^- = \infty$, as

$$\Pr(m_{it}^- = j|\mathbf{z}_{it}^-, r_{it} = -1) = \begin{cases} F^-(\mu_j^- - \mathbf{z}_{it}^{-\prime}\boldsymbol{\gamma}) - F^-(\mu_{j-1}^- - \mathbf{z}_{it}^{-\prime}\boldsymbol{\gamma}) & \text{for} \quad -J \leq j \leq 0, \\ 0 & \text{for} \quad 0 < j \leq J. \end{cases} \tag{4}$$

- *Regime $r_{it} = 0$ (neutral policy stance).*

Conditional on being in regime $r_{it} = 0$ no further policy actions are taken - the rate remains unchanged:

$$\Delta y_{it}|(r_{it} = 0) = 0.$$

Therefore, the conditional probability of a particular outcome j is given by

$$\Pr(\Delta y_{it} = j|\mathbf{x}_{it}, r_{it} = 0) = \begin{cases} 0 & \text{for} \quad j \neq 0, \\ 1 & \text{for} \quad j = 0. \end{cases} \tag{5}$$

- *Regime $r_{it} = 1$ (tight policy stance).*

Conditional on being in regime $r_{it} = 1$ the continuous latent variable m_{it}^{+*}, representing the desired change to the rate, is set by the other *amount* equation

$$m_{it}^{+*} = \mathbf{z}_{it}^{+\prime}\boldsymbol{\delta} + \varepsilon_{it}^+, \tag{6}$$

where $\boldsymbol{\delta}$ is a $K_\delta \times 1$ vector of unknown coefficients, $\mathbf{z}_{it}^+$ is the t^{th} row of an observed $T_i \times K_\delta$ data matrix $\mathbf{Z}_i^+$, and ε_{it}^+ is an i.i.d. error term with the c.d.f. F^+.

The discrete change to the rate m_{it}^+ is determined by

$$m_{it}^+ = j \text{ if } \mu_{j-1}^+ < y_{it}^{+*} \leq \mu_j^+ \text{ for } j = 0 \text{ to } J,$$

where $-\infty = \mu_{-1}^+ \leq \mu_0^+ \leq ... \leq \mu_{J-1}^+ \leq \mu_J^+ = \infty$ are J unknown thresholds to be estimated.

The conditional probability of a particular outcome j is given by

$$\Pr(m_{it}^{+}=j|\mathbf{z}_{it}^{+},r_{it}=1)=\begin{cases}0 & \text{for} \quad -J\leq j<0,\\ F^{+}(\mu_{j}^{+}-\mathbf{z}_{it}^{+\prime}\boldsymbol{\delta})-F^{+}(\mu_{j-1}^{+}-\mathbf{z}_{it}^{+\prime}\boldsymbol{\delta}) & \text{for} \quad 0\leq j\leq J.\end{cases} \tag{7}$$

Assuming that ν_{it}, ε_{it}^{-} and ε_{it}^{+} are independent, the full unconditional probabilities to observe the outcome j are given by combining the probabilities in (2), (4), (5) and (7):

$$\Pr(\Delta y_{it}=j|\mathbf{x}_{it},\mathbf{z}_{it}^{-},\mathbf{z}_{it}^{+})=\begin{cases}I_{j=0}\Pr(r_{it}=0|\mathbf{x}_{it})\Pr(\Delta y_{it}=j|\mathbf{x}_{it},r_{it}=0)\\ +I_{j\geq 0}\Pr(r_{it}=1|\mathbf{x}_{it})\Pr(m_{it}^{+}=j|\mathbf{z}_{it}^{+},r_{it}=1)\\ +I_{j\leq 0}\Pr(r_{it}=-1|\mathbf{x}_{it})\Pr(m_{it}^{-}=j|\mathbf{z}_{it}^{-},r_{it}=-1)\end{cases}$$

$$=\begin{cases}I_{j=0}[F(\alpha_2-\mathbf{x}_{it}^{\prime}\boldsymbol{\beta})-F(\alpha_1-\mathbf{x}_{it}^{\prime}\boldsymbol{\beta})]\\ +I_{j\geq 0}[1-F(\alpha_2-\mathbf{x}_{it}^{\prime}\boldsymbol{\beta})][F^{+}(\mu_{j}^{+}-\mathbf{z}_{it}^{+\prime}\boldsymbol{\delta})-F^{+}(\mu_{j-1}^{+}-\mathbf{z}_{it}^{+\prime}\boldsymbol{\delta})]\\ +I_{j\leq 0}F(\alpha_1-\mathbf{x}_{it}^{\prime}\boldsymbol{\beta})[F^{-}(\mu_{j}^{-}-\mathbf{z}_{it}^{-\prime}\boldsymbol{\delta})-F^{-}(\mu_{j-1}^{-}-\mathbf{z}_{it}^{-\prime}\boldsymbol{\delta})],\end{cases} \tag{8}$$

where $I_{j\geq 0}$ is an indicator function such that $I_{j\geq 0}=1$ if $j\geq 0$ and $I_{j\geq 0}=0$ otherwise (analogously for $I_{j=0}$ and $I_{j\leq 0}$).

The proposed model, as any model with a latent variable, is not identified without some (arbitrary) assumptions. Let us assume the standard normal form of the error distributions F, F^{-} and F^{+}, and also that the intercept components of $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$ and $\boldsymbol{\delta}$ are all equal to zero[9]. However, the above probabilities are absolutely *estimable* functions, i.e. they are invariant to the identifying assumptions. These probabilities can be estimated by using the partial (pooled) ML estimator of the vector of parameters $\boldsymbol{\theta}=(\boldsymbol{\alpha}^{\prime},\boldsymbol{\beta}^{\prime},\boldsymbol{\mu}^{-\prime},\boldsymbol{\gamma}^{\prime},\boldsymbol{\mu}^{+\prime},\boldsymbol{\delta}^{\prime})^{\prime}$ that solves

$$\max_{\boldsymbol{\theta}\epsilon\boldsymbol{\Theta}}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{j=-J}^{J}q_{itj}\ln[\Pr(\Delta y_{it}=j|\mathbf{x}_{it},\mathbf{z}_{it}^{-},\mathbf{z}_{it}^{+},\boldsymbol{\theta})], \tag{9}$$

where q_{itj} is an indicator function such that $q_{itj}=1$ if $\Delta y_{it}=j$ and 0 otherwise. All the parameters in all three equations are separately identified (up to scale) through functional form.

The typical panels contain data covering a short timespan for each individual. In this case, the asymptotic arguments rely on N tending to infinity. With T fixed and $N\to\infty$, this estimator is consistent and $\sqrt{N}$-asymptotically normal without any additional assumptions other than the standard identification assumptions and regularity conditions (see Wooldridge 2010, pp. 489-490). However, the usual asymptotic standard errors and test statistics obtained from pooled estimation are valid only under the assumption of no serial correlation among error terms ν_{it}, ε_{it}^{-} and ε_{it}^{+}. Without dynamic completeness, the standard errors must be adjusted for serial dependence, for example, by using a robust to density misspecification sandwich estimator of asymptotic variance of $\boldsymbol{\theta}$

[9]Employing the ordered logit or complementary log-log counterparts are among the possible alternative versions of the proposed model.

$$\widehat{Avar(\widehat{\boldsymbol{\theta}})} = \left(-\sum_{i=1}^{N}\sum_{t=1}^{T}\mathbf{H}_{it}(\widehat{\boldsymbol{\theta}})\right)^{-1}\left(\sum_{i=1}^{N}\left[\sum_{t=1}^{T}\mathbf{s}_{it}(\widehat{\boldsymbol{\theta}})\sum_{t=1}^{T}\mathbf{s}_{it}(\widehat{\boldsymbol{\theta}})'\right]\right)\left(-\sum_{i=1}^{N}\sum_{t=1}^{T}\mathbf{H}_{it}(\widehat{\boldsymbol{\theta}})\right)^{-1}, \quad (10)$$

where $\mathbf{s}_{it}(\widehat{\boldsymbol{\theta}})$ is the score vector and $\mathbf{H}_{it}(\widehat{\boldsymbol{\theta}})$ is the expected Hessian (see Wooldridge 2010, pp. 490-493). The asymptotic standard errors of $\widehat{\boldsymbol{\theta}}$ are the square roots of the diagonal elements of (10).

In the application to panel data with small N and relatively large T, we are basically in the realm of time-series analysis, and the asymptotic arguments rely on T tending towards infinity, standard identification and stationarity assumptions. Using either fixed T and $N \to \infty$ or fixed N and $T \to \infty$ asymptotics, the above pooled ML estimator in (9) is consistent and asymptotically normal even if the error terms are arbitrarily serially correlated, the dynamics are not correctly specified, and $\mathbf{X}_i$, $\mathbf{Z}_i^-$ and $\mathbf{Z}_i^+$ contain not strictly exogenous covariates, lags of covariates and lagged Δy_{it}[10].

2.2 The nested ordered probit (NOP) model

The only difference between the NOP and CNOP models is that all three nests of the NOP model do not overlap, i.e. regimes $r_{it} = -1$ and $r_{it} = 1$ do not allow for a no-change response (see the top panels of Figure 2). Therefore, in the NOP model the full unconditional probabilities to observe an outcome j (again, assuming that the disturbance terms of three latent equations are independent) are given by

$$\Pr(\Delta y_{it} = j|\mathbf{z}_{it}^-, \mathbf{z}_{it}^+, \mathbf{x}_{it}) = \begin{cases} I_{j=0}\Pr(r_{it} = 0|\mathbf{x}_{it}) + \\ I_{j>0}\Pr(r_{it} = 1|\mathbf{x}_{it})\Pr(w_{it}^+ = j|\mathbf{z}_{it}^+, r_{it} = 1) \\ +I_{j<0}\Pr(r_{it} = -1|\mathbf{x}_{it})\Pr(w_{it}^- = j|\mathbf{z}_{it}^-, r_{it} = -1) \end{cases}$$

$$= \begin{cases} I_{j=0}[F(\alpha_2 - \mathbf{x}_{it}'\boldsymbol{\beta}) - F(\alpha_1 - \mathbf{x}_{it}'\boldsymbol{\beta})] \\ +I_{j>0}[1 - F(\alpha_2 - \mathbf{x}_{it}'\boldsymbol{\beta})][F^+(\mu_j^+ - \mathbf{z}_{it}^{+\prime}\boldsymbol{\delta}) - F^+(\mu_{j-1}^+ - \mathbf{z}_{it}^{+\prime}\boldsymbol{\delta})] \\ +I_{j<0}F(\alpha_1 - \mathbf{x}_{it}'\boldsymbol{\beta})[F^-(\mu_j^- - \mathbf{z}_{it}^{-\prime}\boldsymbol{\gamma}) - F^-(\mu_{j-1}^- - \mathbf{z}_{it}^{-\prime}\boldsymbol{\gamma})], \end{cases} \quad (11)$$

where now $-\infty = \mu_{-J-1}^- \leq \mu_{-J}^- \leq ... \leq \mu_{-1}^- = \infty$ and $-\infty = \mu_0^+ \leq ... \leq \mu_{J-1}^+ \leq \mu_J^+ = \infty$ are $2(J-1)$ unknown thresholds to be estimated at the lower level (instead of $2J$ in the CNOP model), and the other parameters and assumptions are analogous to those in the CNOP model.

To estimate the NOP model one can employ the ML estimator from (9), using the probabilities from (11). The log of the likelihood function of the NOP model, in contrast to that of the CNOP one, is separable with respect to the parameters in three latent equations. Thus, solving (9) is equivalent to maximizing separately the likelihoods of three OP models, corresponding to the above three latent equations (1), (3) and (6), where the data matrices $\mathbf{Z}_i^+$ and $\mathbf{Z}_i^-$ are truncated to contain only the rows with $\Delta y_{it} > 0$ and $\Delta y_{it} < 0$, respectively.

[10]This result is analogous to employing pooled OLS estimation in linear panel models.

2.3 Relaxing assumption of independent disturbances

The NOP and CNOP models can be further extended by relaxing the assumption that the error terms $\boldsymbol{\nu}$, $\boldsymbol{\varepsilon}^-$ and $\boldsymbol{\varepsilon}^+$ are uncorrelated, and introducing the correlated versions of the models, NOPC and CNOPC ones. I now assume that $(\boldsymbol{\nu}, \boldsymbol{\varepsilon}^-)$ and $(\boldsymbol{\nu}, \boldsymbol{\varepsilon}^+)$ follow the standardized bivariate normal distributions with the correlation coefficients ρ^- and ρ^+, respectively. The full unconditional probabilities to observe an outcome j for the CNOPC model can be written now as

$$\begin{aligned}
&\Pr(\Delta y_{it} = j) = I_{j=0}[F(\alpha_2 - \mathbf{x}'_{it}\boldsymbol{\beta}) - F(\alpha_1 - \mathbf{x}'_{it}\boldsymbol{\beta})] \\
&+ I_{j\geq 0}[F_2(\mathbf{x}'_{it}\boldsymbol{\beta} - \alpha_2; \mu_j^+ - \mathbf{z}_{it}^{+\prime}\boldsymbol{\delta}; -\rho^+) - F_2(\mathbf{x}'_{it}\boldsymbol{\beta} - \alpha_2; \mu_{j-1}^+ - \mathbf{z}_{it}^{+\prime}\boldsymbol{\delta}; -\rho^+)] \\
&+ I_{j\leq 0}[F_2(\alpha_1 - \mathbf{x}'_{it}\boldsymbol{\beta}; \mu_j^- - \mathbf{z}_{it}^{-\prime}\boldsymbol{\gamma}; \rho^-) - F_2(\alpha_1 - \mathbf{x}'_{it}\boldsymbol{\beta}; \mu_{j-1}^- - \mathbf{z}_{it}^{-\prime}\boldsymbol{\gamma}; \rho^-)],
\end{aligned} \tag{12}$$

where $F_2(\phi_1;\phi_2;\xi)$ is the c.d.f. of the standardized bivariate normal distribution with the correlation coefficient ξ between the two random variables ϕ_1 and ϕ_2.

The full unconditional probabilities to observe an outcome j for the NOPC model are given by

$$\begin{aligned}
&\Pr(\Delta y_{it} = j) = I_{j=0}[F(\alpha_2 - \mathbf{x}'_{it}\boldsymbol{\beta}) - F(\alpha_1 - \mathbf{x}'_{it}\boldsymbol{\beta})] \\
&+ I_{j>0}[F_2(\mathbf{x}'_{it}\boldsymbol{\beta} - \alpha_2; \mu_j^+ - \mathbf{z}_{it}^{+\prime}\boldsymbol{\delta}; -\rho^+) - F_2(\mathbf{x}'_{it}\boldsymbol{\beta} - \alpha_2; \mu_{j-1}^+ - \mathbf{z}_{it}^{+\prime}\boldsymbol{\delta}; -\rho^+)] \\
&+ I_{j<0}[F_2(\alpha_1 - \mathbf{x}'_{it}\boldsymbol{\beta}; \mu_j^- - \mathbf{z}_{it}^{-\prime}\boldsymbol{\gamma}; \rho^-) - F_2(\alpha_1 - \mathbf{x}'_{it}\boldsymbol{\beta}; \mu_{j-1}^- - \mathbf{z}_{it}^{-\prime}\boldsymbol{\gamma}; \rho^-)].
\end{aligned} \tag{13}$$

To estimate the CNOPC and NOPC models by ML, we have to solve (9), replacing the probabilities in (8) and (11) with those in (12) and (13), respectively, and re-defining the vector of parameters $\boldsymbol{\theta}$ as $\boldsymbol{\theta} = (\boldsymbol{\alpha}', \boldsymbol{\beta}', \boldsymbol{\mu}^{-\prime}, \boldsymbol{\gamma}', \boldsymbol{\mu}^{+\prime}, \boldsymbol{\delta}', \rho^-, \rho^+)'$.

2.4 Partial effects

The partial (or marginal) effect (PE) of each continuous covariate on the probability of each discrete choice is computed as the partial derivative with respect to this covariate, holding all the others fixed at their sample median values. For the discrete-valued covariates the PE is computed as the change in the probabilities, when this covariate changes by one increment and all the others are fixed. To facilitate the derivation of the PEs, the matrices of covariates and corresponding vectors of parameters can be partitioned as follows:

$$\begin{aligned}
&\mathbf{X} = (\mathbf{W}, \mathbf{P}, \mathbf{M}, \widetilde{\mathbf{X}}), \quad &&\mathbf{Z}^+ = (\mathbf{W}, \mathbf{P}, \mathbf{V}, \widetilde{\mathbf{Z}}^+), \quad &&\mathbf{Z}^- = (\mathbf{W}, \mathbf{M}, \mathbf{V}, \widetilde{\mathbf{Z}}^-), \\
&\boldsymbol{\beta} = (\boldsymbol{\beta}'_w, \boldsymbol{\beta}'_p, \boldsymbol{\beta}'_m, \widetilde{\boldsymbol{\beta}}')', \quad &&\boldsymbol{\delta} = (\boldsymbol{\delta}'_w, \boldsymbol{\delta}'_p, \boldsymbol{\delta}'_v, \widetilde{\boldsymbol{\delta}}')', \quad &&\boldsymbol{\gamma} = (\boldsymbol{\gamma}'_w, \boldsymbol{\gamma}'_m, \boldsymbol{\gamma}'_v, \widetilde{\boldsymbol{\gamma}}')',
\end{aligned}$$

where $\mathbf{W}$ includes only the variables common for $\mathbf{X}$, $\mathbf{Z}^+$ and $\mathbf{Z}^-$; $\mathbf{P}$ includes only the variables common for both $\mathbf{X}$ and $\mathbf{Z}^+$, but which are not in $\mathbf{Z}^-$; $\mathbf{M}$ includes only the variables common for both $\mathbf{X}$ and $\mathbf{Z}^-$, but not in $\mathbf{Z}^+$; $\mathbf{V}$ includes only the variables common for both $\mathbf{Z}^-$ and $\mathbf{Z}^+$, but not in $\mathbf{X}$; whereas $\widetilde{\mathbf{X}}$, $\widetilde{\mathbf{Z}}^+$ and $\widetilde{\mathbf{Z}}^-$ include only those unique variables that appear only in one of the latent equations.

A matrix of covariates $\mathbf{X}^*$ and the vectors of parameters for $\mathbf{X}^*$ can be written down as

$$\mathbf{X}^* = (\mathbf{W}, \mathbf{P}, \mathbf{M}, \widetilde{\mathbf{X}}, \mathbf{V}, \widetilde{\mathbf{Z}}^+, \widetilde{\mathbf{Z}}^-),\ \boldsymbol{\beta}^* = (\boldsymbol{\beta}'_w, \boldsymbol{\beta}'_p, \boldsymbol{\beta}'_m, \widetilde{\boldsymbol{\beta}}', \mathbf{0}', \mathbf{0}', \mathbf{0}')',$$

$\boldsymbol{\delta}^* = (\boldsymbol{\delta}'_w, \boldsymbol{\delta}'_p, \mathbf{0}', \mathbf{0}', \boldsymbol{\delta}'_v, \widetilde{\boldsymbol{\delta}}', \mathbf{0}')'$, $\boldsymbol{\gamma}^* = (\boldsymbol{\gamma}'_w, \mathbf{0}', \boldsymbol{\gamma}'_m, \mathbf{0}', \boldsymbol{\gamma}'_v, \mathbf{0}', \widetilde{\boldsymbol{\gamma}}')'$.

The partial effects of the row vector $\mathbf{x}^*_{it}$ on the overall probabilities in (12) can be now computed for the CNOPC model as

$$\begin{aligned}
\underset{\Pr(\Delta y_{it}=j)}{\mathbf{PE}} &= -I_{j=0}[f(\alpha_2 - \mathbf{x}'_{it}\boldsymbol{\beta}) - f(\alpha_1 - \mathbf{x}'_{it}\boldsymbol{\beta})]\boldsymbol{\beta}^* \\
&+ I_{j\geq 0}\left\{\left[F\left(\frac{\mathbf{x}'_{it}\boldsymbol{\beta}-\alpha_2+\rho^+(\mu^+_{j-1}-\mathbf{z}^{+\prime}_{it}\boldsymbol{\delta})}{\sqrt{1-(\rho^+)^2}}\right) f(\mu^+_{j-1} - \mathbf{z}^{+\prime}_{it}\boldsymbol{\delta})\right.\right. \\
&\left. - F\left(\frac{\mathbf{x}'_{it}\boldsymbol{\beta}-\alpha_2+\rho^+(\mu^+_{j}-\mathbf{z}^{+\prime}_{it}\boldsymbol{\delta})}{\sqrt{1-(\rho^+)^2}}\right) f(\mu^+_{j} - \mathbf{z}^{+\prime}_{it}\boldsymbol{\delta})\right]\boldsymbol{\delta}^* \\
&\left. + \left[F\left(\frac{\mu^+_j-\mathbf{z}^{+\prime}_{it}\boldsymbol{\delta}+\rho^+(\mathbf{x}'_{it}\boldsymbol{\beta}-\alpha_2)}{\sqrt{1-(\rho^+)^2}}\right) - F\left(\frac{\mu^+_{j-1}-\mathbf{z}^{+\prime}_{it}\boldsymbol{\delta}+\rho^+(\mathbf{x}'_{it}\boldsymbol{\beta}-\alpha_2)}{\sqrt{1-(\rho^+)^2}}\right)\right] f(\mathbf{x}'_{it}\boldsymbol{\beta}-\alpha_2)\boldsymbol{\beta}^*\right\} \\
&+ I_{j\leq 0}\left\{\left[F\left(\frac{\alpha_1-\mathbf{x}'_{it}\boldsymbol{\beta}-\rho^-(\mu^-_{j-1}-\mathbf{z}^{-\prime}_{it}\boldsymbol{\gamma})}{\sqrt{1-(\rho^-)^2}}\right) f(\mu^-_{j-1} - \mathbf{z}^{-\prime}_{it}\boldsymbol{\gamma})\right.\right. \\
&\left. - F\left(\frac{\alpha_1-\mathbf{x}'_{it}\boldsymbol{\beta}-\rho^-(\mu^-_{j}-\mathbf{z}^{-\prime}_{it}\boldsymbol{\gamma})}{\sqrt{1-(\rho^-)^2}}\right) f(\mu^-_{j} - \mathbf{z}^{-\prime}_{it}\boldsymbol{\gamma})\right]\boldsymbol{\gamma}^* \\
&\left. - \left[F\left(\frac{\mu^-_j-\mathbf{z}^{-\prime}_{it}\boldsymbol{\gamma}-\rho^-(\alpha_1-\mathbf{x}'_{it}\boldsymbol{\beta})}{\sqrt{1-(\rho^-)^2}}\right) - F\left(\frac{\mu^-_{j-1}-\mathbf{z}^{-\prime}_{it}\boldsymbol{\gamma}-\rho^-(\alpha_1-\mathbf{x}'_{it}\boldsymbol{\beta})}{\sqrt{1-(\rho^-)^2}}\right)\right] f(\alpha_1-\mathbf{x}'_{it}\boldsymbol{\beta})\boldsymbol{\beta}^*\right\},
\end{aligned} \tag{14}$$

where f is the probability density function (p.d.f.) of the standard normal distribution F. The PEs for the NOPC model are given by replacing $I_{j\geq 0}$ with $I_{j>0}$ and $I_{j\leq 0}$ with $I_{j<0}$. The PEs for the NOP and CNOP models are obtained as above by setting $\rho^- = \rho^+ = 0$. The asymptotic standard errors of the PEs are computed using the Delta method as the square roots of the diagonal elements of

$$\widehat{Avar(\underset{\Pr(\Delta y_{it}=j)}{\widehat{\mathbf{PE}(\boldsymbol{\theta})}})} = \nabla_\theta \underset{\Pr(\Delta y_{it}=j)}{\widehat{\mathbf{PE}(\boldsymbol{\theta})}})\widehat{Avar}(\widehat{\boldsymbol{\theta}})\nabla_\theta \underset{\Pr(\Delta y_{it}=j)}{\widehat{\mathbf{PE}(\boldsymbol{\theta})}})'.$$

2.5 Model comparison

The performance of competing models can be compared by using the model selection tests and informational criteria.

The NOP and CNOP models are nested in the NOPC and CNOPC models, respectively, as their uncorrelated special cases. The NOP model is nested in the CNOP model. The latter becomes a NOP model with the same value of the likelihood function if $\mu^-_{-1} \to \infty$ and $\mu^+_0 \to -\infty$, and hence, $\Pr(y^+_{it} = 0|\mathbf{z}^+_{it}, r_{it} = 1) \to 0$ and $\Pr(y^-_{it} = 0|\mathbf{z}^-_{it}, r_{it} = -1) \to 0$, which can be implemented by letting μ^-_{-1} and μ^+_0 to be equal to the largest and smallest numbers available for the estimation software. Testing the NOP versus NOPC, NOP versus CNOP, NOP versus CNOPC, NOPC versus CNOPC, and CNOP versus CNOPC model can be performed with the likelihood ratio (LR) test.

The OP models is not nested in either of the two-level models, and vice versa. However, the OP model is not strictly non-nested with them. All five models overlap if all their slope coefficients are restricted to being zero (i.e. if $\boldsymbol{\beta} = \mathbf{0}$, $\boldsymbol{\gamma} = \mathbf{0}$, $\boldsymbol{\delta} = \mathbf{0}$, and the vector of slope parameters in the OP latent equation is also fixed to zero), and only the thresholds are estimated. Therefore, testing the OP versus any of the two-level models, as well as the NOPC versus CNOP model (which overlap if both reduce to the NOP model) can be conducted with a test for non-nested overlapping models, such as the *Vuong* test (due to Vuong 1989) that utilizes the statistical significance between the difference in the log likelihoods. The testing procedure is sequential. First, we need to verify that the two models are not equivalent, i.e. separately perform t- or F-tests to check whether the parameters of interest violate the overlapping constraints. Second, if the overlapping restrictions can be rejected, we have to conduct the *Vuong* test for strictly non-nested models. The null hypothesis of this test is that both models are misspecified, but equally close to the unknown true d.g.p. The test statistic is very simple to compute: it is equal to the average difference of the individual likelihoods divided by the estimated standard error of those individual differences. Under the null hypothesis, the *Vuong* test statistic converges in distribution to a standard normal one. If the absolute value of the test statistic is less than the critical value, say 1.96, we cannot discriminate between the two models given the data. If the test statistic exceeds 1.96, we reject the equivalence in favor of one of the models; if the test statistic smaller than -1.96, we reject the equivalence in favor of the other.

The following model-selection information criteria are computed: $AIC = -2l(\boldsymbol{\theta}) + 2k$, $BIC = -2l(\boldsymbol{\theta}) + ln(N)k$, $cAIC = -2l(\boldsymbol{\theta}) + (1 + \ln(N))k$ (consistent AIC), $AICc = AIC + 2k(k+1)/(N-k-1)$ (corrected AIC), and $HQIC = -2l(\boldsymbol{\theta}) + 2ln(ln(N))k$, where k is the total number of the estimated parameters. The adjusted *McFadden* pseudo-R^2 measure of fit (given by $1 - (l(\boldsymbol{\theta}) - k)/l_0(\boldsymbol{\theta})$, where $l_0(\boldsymbol{\theta})$ is the value of the restricted likelihood function, maximized with all the slope parameters in $\boldsymbol{\theta}$ fixed to zero) can also be used for the model selection, but its selection results are equivalent to those of the AIC, because the value of the $l_0(\boldsymbol{\theta})$ is identical in all the above models. Another measure of fit, the *Hit rate*, is computed as the percentage of correct predictions, where the predicted discrete outcome is that with the highest estimated probability.

3 Finite sample performance

I conducted massive Monte Carlo experiments to illustrate and compare the finite sample performance of the ML estimators in the single- and three-equation models, namely to assess the bias and uncertainty of the estimates of parameters and partial effects, as well as their asymptotic standard errors, the performance of the *LR* and *Vuong* tests and model selection criteria as discussed in the previous section, and the effect of exclusion restrictions. The simulations were performed using GAUSS programming language (version 10) with CML module (version 2) for the constrained ML estimation. The details of Monte Carlo design and the results of these simulations are reported and discussed in Appendix A. Here I provide a brief summary of Monte Carlo design and main findings.

The observations in the repeated samples were drawn independently. This corresponds to either the cross-sectional model with uncorrelated units or to the time series model without serial dependence. Therefore, the results are applicable to assess the finite-sample performance of the ML estimator with: (i) i.i.d. cross-sectional data and $N \to \infty$ asymptotics; (ii) dynamically complete model for time-series data and $T \to \infty$ asymptotics; (iii) i.i.d. panel data with fixed T and $N \to \infty$ asymptotics; and (iv) dynamically complete model for panel data with fixed N and $T \to \infty$ asymptotics[11].

Five different d.g.p's were simulated: OP, NOP, NOPC, CNOP, and CNOPC. For each d.g.p., 3000 repeated samples with 250, 500 and 1000 observations were generated. Under

[11] The Monte Carlo simulations for the panel data with small N and relatively large T, where latent errors are either not autocorrelated or autocorrelated, will be added soon.

each d.g.p. and for each sample size, several competing models were estimated, always including the OP and NOP models as the benchmarks. I found that: (i) it requires two to three times more observations for the three-part models to achieve the same accuracy of the estimated parameters as that of the OP model; (ii) each of the five models under its own d.g.p., not surprisingly, estimates the quantities of interest better than the other models; however, the three-part models under the true OP d.g.p. perform much better than the OP model under the NOP(C) and CNOP(C) d.g.p's; (iii) as the sample size increases, the performance of the three-part models under the OP d.g.p. improves drastically, whereas the performance of the OP model under the NOP(C) and CNOP(C) d.g.p's improves only slightly.

Under any three-part d.g.p, the *Vuong* tests tend to correctly favor the true model versus the OP model in almost 100% of replications, as sample size increases. However, under the OP d.g.p. the *Vuong* tests of the NOP and CNOP models versus the OP model fail to discriminate between the two models, and are never in favor of the true OP model. The *LR* tests of the NOP versus NOPC and the CNOP versus CNOPC model (when the true d.g.p. is correlated) both have an empirical size between 4 and 5 percent, very close to the nominal size of 5 percent. Regarding the information criteria, while the *AIC* and *AICc* under the OP, NOP and CNOP d.g.p's select the true model slightly less frequently than the *BIC* and *cAIC*, under the NOPC and CNOPC d.g.p's they clearly outperform the *HQIC* and especially the *BIC* and *cAIC*.

In addition, in order to assess the effect of exclusion restrictions, three different scenarios of the overlap among the covariates in the specifications of three latent equations were simulated: "no overlap" (each covariate belongs only to one equation), "partial overlap" (each covariate belongs to two equations) and "complete overlap" (all three equations have the same set of covariates). I found that the more exclusion restrictions the more accurate the estimates of the *PE*s, and the fewer the problems with estimation. The simulation results suggest that the asymptotic estimator might not perform well without the exclusion restrictions, that is with the complete overlap among the covariates, in the small samples (fewer than 35 observations per parameter). In case of the NOPC and CNOPC models under the partial overlap scenario in the small samples there might be the problems with the convergence and invertibility of the Hessian.

4 An application to policy interest rate

> "It is highly desirable that policy practice be formalized to the maximum possible extent."
>
> – W. Poole, then-President of the Federal Reserve Bank of St. Louis[12]

The policy rate is a key determinant of the other short-term market interest rates, and of sharp interest for market participants: "What the market needs to know is the policy response function by which the central bank acts in a consistent way over time" (Poole, 2003). Furthermore, "if practitioners in financial markets gain a better understanding of how policy is likely to respond to incoming information, asset prices and bond yields will tend to respond to economic data in ways that further the central bank's policy objectives" (Bernanke, 2007). Another important reason to model policy rate is a search for better

[12] See Poole (2006).

policy. In order to improve it, we have to obtain a clear empirical description of what is going to be improved. It is really hard to evaluate the monetary policy without describing it, using an econometric model.

In this section I let the real-world data speak, and apply the proposed and conventional OP models to explain the systematic components of policy interest rate decisions of the NBP, employing a panel of the individual votes of the MPC members and *real-time* macroeconomic data available at each MPC meeting during the 1998-2009 period[13].

4.1 Data

Since the adoption of direct inflation targeting in 1998 the NBP policy rate – the reference rate – may be undoubtedly treated as a principal instrument of Polish monetary policy[14]. The reference rate has been always set administratively by the MPC of the NBP, and is not an outcome of the interaction between the market supply and demand. The MPC consists of ten members and makes policy rate decisions once per month by means of formal voting. The Council members are appointed for a non-renewable term of six years, but the Chair may serve for two consecutive terms. The first term lasted from February 1998 through January 2004[15]. The second term lasted from February 2004 through January 2010.

Table 2: Frequency distribution of the individual votes of the MPC members on policy rate

Δy_{it} (preferred rate change by member i)	Decrease	No change	Increase	All
Number of observations	309	889	187	1385
Percentage	22%	64%	14%	100%

Notes: The sample period is from 04/1998 through 12/2009. The first meeting of the second MPC in February 2004 is omitted.

The MPC has always altered the levels of policy rates in discrete adjustments – the multiples of 25 basis points (bp) - made in the range from 25 to 250 bp. Table 2 shows the frequency distribution of the dependent variable – the individual MPC members' votes (reported in Tables 23, 24 and 25 of Appendix C) for the changes to the rate in the period 1998/04 - 2009/12. To provide a reliable inference, the individual policy preferences are consolidated for analysis into three categories: increase, no change and decrease. At a monthly policy meeting, each member can express his or her preferred policy rate change, and can make a proposition to be voted on. If no proposition is made, there is no voting at all, and the rate remains unchanged; otherwise, the Chair selects the largest proposed move and the members vote on it. If the first voted proposition commands a majority, then the others are not voted on; otherwise, the members vote on the alternative proposal. As a matter of fact, the second voted proposal has always been passed. In case of two rounds

[13] The data are taken from Sirchenko (2008) and updated till the end of 2009.

[14] See Sirchenko (2008) and references therein for the background of monetary policy in Poland.

[15] However, one member was replaced before the policy meeting in January 2004, and another passed away, so his seat was filled midterm in August 2003. Because the first MPC Chair had resigned in December 2000, the Chair since then has been appointed with a three-year lag with respect to the other members.

of voting, the desired interest rate changes during the first round are used in estimations. The first two meetings (in February and March 1998) of the newly-established MPC are dropped from the sample to account for a transition to a new policy regime of inflation targeting. The first meeting of the second MPC in February 2004 is also omitted. The policymakers have been absent from the meetings 15 times. Among the 1385 observations used in estimations, the policymakers preferred to leave the rate unchanged 889 times (in 64 percent of cases), as Table 2 reports.

The policy inclination decision is assumed to be driven by a direct response to new economic information, such as inflation developments, the prospects for real economy, the spread between long- and short-term market interest rates, and recent change to the ECB policy rate. The amount decisions, fine-tuning and smoothing the rate, are expected to be driven by the tactical institutional factors such as: (i) recent ”policy bias” or ”balance of risks” statements (addressing the policymakers' concerns about the competence and credibility of the central bank's communication); (ii) the dissent among the policymakers at the previous policy meeting (if dissenters preferred a lower policy rate, it creates a downward pressure to the rate at the current meeting, and, hence, the probability of rate cut increases); and (iii) the change to the rate, made by the MPC at the previous policy meeting (reflecting the inertia of monetary policy and deliberate interest-rate smoothing behavior of the central bank).

The indicator of policy bias ($bias_t$) at the meeting t is defined as -1 if it is ”easing”, 0 if ”neutral”, and 1 if ”restrictive”. The measure of dissent among the policymakers is calculated as follows. Consider a committee with M members. For each member i and each policy meeting t define the individual dissent indicator

$$d_{it} = \begin{cases} 1 & \text{if } \Delta y_{it} > \Delta nbpr_t, \\ 0 & \text{if } \Delta y_{it} = \Delta nbpr_t, \\ \text{-1} & \text{if } \Delta y_{it} < \Delta nbpr_t, \end{cases} \tag{15}$$

where Δy_{it} is the change to the reference rate preferred by member i, and $\Delta nbpr_t$ is the change made by the MPC. The measure of dissent at the meeting t ($dissent_t$) is then defined as the average of individual dissents across all MPC members:

$$dissent_t = \frac{1}{M}\sum_{i=1}^{M} d_{it}. \tag{16}$$

Table 26 of Appendix C reports for each MPC meeting the values of policy bias indicator ($bias_t$), overall dissent at the meeting ($dissent_t$), and policy rate decision of the Council ($\Delta nbpr_t$). There was at least one dissenting member in 44 percent of MPC meetings. As Table 27 of Appendix C shows for each MPC member, the average values of individual dissents (d_{it}) across all meetings are between -0.232 and 0.400, i.e. the most "dovish" member (Ziółkowska) preferred a lower interest rate than the majority in 23.2% of meetings, while the most "hawkish" member (Filar) was in favor of a higher interest than the majority in 40% of meetings.

A dummy variable for the expected inflation above the official inflation target ($I(cpi_t^e > tar_t)$) is included into $\mathbf{Z}^+$only. The change to the rate at the last policy meeting ($\Delta nbpr_{t-1}$) is allowed to enter all three equations. The detailed definitions of all variables used in the study are given in Table 3. The sample descriptive statistics is shown in Table 22 of Appendix C.

To account for the unobserved individual heterogeneity of policy preferences, I allow for intercept variation[16]. Slope heterogeneity is not of a concern, since our interest is in estimating the average effects of explanatory variables. Under an assumption that the slope coefficients differ randomly across individuals, the pooled estimator gives unbiased estimates of these average effects.

Table 3: Definitions of variables

Mnemonics	Variable description (source of data)
	Dependent variable
Δy_i	Change to NBP reference rate, preferred by i MPC member: 1 if an increase, 0 if no change, -1 if a decrease (NBP).
	Variables in **X** only
Δcpi	Last monthly change to consumer price index (CPI), annual rate in percent (GUS - Central Statistical Office of Poland).
situation	Index of expected general economic situation in industry from Business Tendency Survey, divided by 100
spread	Difference between 12- and 1-month Poland interbank offer rate, 5-business-day moving average, annualized percent (Thompson Reuters).
$\Delta ecbr$	Change to the ECB policy rate (since 02/1999, in 1998 - to Bundesbank policy rate, set equal to zero in 01/1999), announced at the last policy meeting, annualized percent (ECB and Bundesbank).
	Variables in **X**, $\mathbf{Z}^-$ and $\mathbf{Z}^+$
$\Delta nbpr$	Change to the NBP reference rate, announced at an MPC meeting, annualized percent (NBP).
$I(h)_i$	1 if the average $Dissent_i$ (from the first up to the last MPC meeting) is greater than 0.1, 0 - otherwise; see Eq.(17).
$I(d)_i$	1 if the average $Dissent_i$ (from the first up to the last MPC meeting) is less than -0.1, 0 - otherwise; see Eq.(17).
$I(Bal)_i$	1 if i MPC member is Balcerowicz, and 0 otherwise. The other MPC members are coded as: *Cze* - Czekaj, *Dab* - Dąbrowski, *Fil* - Filar, *Gra* - Grabowski, *Gro* - Gronkiewicz-Waltz, *Joz* - Józefiak, *Krz* - Krzyżewski, *Lac* - Łączkowski, *Nie* - Nieckarz, *Nog* - Noga, *Ows* - Owsiak, *Pie* - Pietrewicz, *Pru* - Pruski, *Ros* - Rosati, *Skr* - Skrzypek, *Sla* - Sławiński, *Was* - Wasilewska-Trenkner, *Woj* - Wojtyna, *Woz* - Wójtowicz, *Zio* - Ziółkowska.
	Variables in $\mathbf{Z}^-$ and/or $\mathbf{Z}^+$ only
dissent	Measure of dissent at an MPC meeting, defined by Eq. (15) (NBP).
bias	Indicator of "policy bias" or "balance of risks" statements (available since 02/2000, set equal to zero before): -1 if "easing", 0 if "neutral", and 1 if "restrictive" (NBP).
$I(cpi^e > tar)$	1 if $cpi^e > tar$, and 0 otherwise; tar is the official inflation target; cpi^e is the expected CPI over next 12 months, annual rate in percent (Ipsos-Demoskop survey of consumers and NBP).

I consider two alternative specifications of the CNOP(C) models. Both include the following common covariates: in $\mathbf{X}$ – Δcpi_t, $situation_t$, $spread_t$, $\Delta ecbr_t$, and $\Delta nbpr_{t-1}$; in $\mathbf{Z}^-$ – $\Delta nbpr_{t-1}$, $dissent_{t-1}$, and $bias_{t-1}$; and in $\mathbf{Z}^+$ – $\Delta nbpr_{t-1}$, $dissent_{t-1}$, $bias_{t-1}$, and $I(cpi^e_t > tar_t)$. In addition to the above, the fixed effects (FE) specification includes

[16]Since the model is highly non-linear, failure to address the heterogeneity can lead to a bias, not just inefficiency, even if all covariates are truly exogenous, whereas no bias emerges in the linear case.

twenty dummy variables for individual MPC members, allowing each individual to have a different intercept in all three latent equations[17]. The FE specification is an appropriate approach here, because we don't have a sample of individuals drawn randomly from a large population, but instead possess a full set of all twenty-one MPC members. Given that the cross-sectional dimension ($N = 21$) is small relative to the observed numbers of time periods (T_i are about 67 on average, ranging from 36 to 76), we don't have the "incidental parameters problem" (see Neyman and Scott 1948, Lancaster 2000). Nor should we expect any significant fixed T asymptotic bias of our estimator with such a large temporal size[18].

An alternative specification with dummies for hawkish and dovish members (the HD specification) is more parsimonious, and includes only two dummy variables, $I(h)_{it}$ and $I(d)_{it}$, defined for $t \geq 2$ as 1 if the average individual dissent (from the first up to the previous MPC meeting) is, respectively, above 0.1 or below -0.1, and 0 - otherwise:

$$I(h)_{it} = \begin{cases} 1 & \text{if } \frac{1}{t-1}\sum_{j=1}^{t-1} d_{i,t-1} > 0.1, \\ 0 & \text{otherwise,} \end{cases} \quad \text{and } I(d)_{it} = \begin{cases} 1 & \text{if } \frac{1}{t-1}\sum_{j=1}^{t-1} d_{i,t-1} < -0.1, \\ 0 & \text{otherwise.} \end{cases} \tag{17}$$

The HD specification, which includes only two instead of twenty dummies in each of the three latent equations, saves 54 degrees of freedom compared to the FE specification. Thus, it can produce more efficient estimates of the common slopes.

5 Estimation results

The following eight competing models were estimated by pooled ML, using the same set of explanatory variables (i.e. all the covariates in $\mathbf{X}$, $\mathbf{Z}^-$ and $\mathbf{Z}^+$ of the CNOP model): (i) the standard OP model; (ii) the OP model with random effects (REOP); (iii) the generalized OP (GOP) model that relaxes the parallel regression assumption of the standard OP model, and that allows the slope coefficients to differ by outcome category; (iv) the multinomial probit (MNP) model, simultaneously estimating binary probits for all possible comparisons (in our case two) among the outcome categories; (v) the two-equation ZIOP model that allows zero observations to come from two different processes; (vi) the ZIOP(a) model, which is identical to the ZIOP model, except that all the covariates in the participation equation are taken by their absolute values to take into account the binary (change versus no change) nature of the first-stage decision; and (vii-viii) the three-equation CNOP and CNOPC models with different sets of covariates in each equation[19]. To give the ZIOP and ZIOP(a) models better chances, all of the CNOP covariates are included into both parts, contrary to the three-part models.

Table 4 reports the summary statistics from six alternative models with FE specification and REOP model[20]. The two- and three-equation models demonstrate a sharp increase in

[17] The individual dummy for Gronkiewicz-Waltz, the first MPC Chair (in 1998-2000) and the only MPC member in the sample, who has never dissented, is omitted.

[18] For example, using Monte Carlo methods, Greene (2004) studied the incidental parameters problem for discrete-choice panel models, including the OP model. As T increases from 2 to 20, the 160 percent bias of the estimated coefficients reduces to 6 percent.

[19] In addition, the ordered logit and multinomial logit counterparts were also estimated with similar, but slightly worse likelihoods than those of the OP and MNP models.

[20] The loglikelihood (and hit rate) of the ordered logit and multinomial logit models (not reported) are

the likelihood and hit rate compared to the single-equation ones. The CNOP model is superior to the others according to *AIC* and *HQIC*, while the ZIOP(a) model is favored by the *BIC* and *AICc*. However, all the *Vuong* tests of the CNOP model versus the ZIOP(a), ZIOP and OP models are in favor of the CNOP model at the 1 percent significance level. The CNOPC model with heavily parameterized FE specification has experienced problems with the invertibility of the Hessian (likely, due to the multicollinearity problems).

Table 4: Changes to policy rate: comparison of alternative models with fixed effects specification including twenty individual dummies

Model	REOP	OP	GOP	MNP	ZIOP	ZIOPa	CNOP
ln l(θ)	-728.5	-696.1	-640.9	-639.1	-580.7	-559.3	-502.6
# of parameters	11	30	58	58	59	59	78
AIC	1479.1	1452.1	1397.7	1394.2	1279.4	1236.6	**1161.2**
BIC	1536.6	1609.1	1701.3	1697.7	1588.2	**1545.3**	1569.4
Corrected *AIC*	1547.7	1639.1	1759.7	1755.7	1647.2	**1604.3**	1647.4
HQIC	1500.6	1510.9	1511.4	1507.7	1394.9	1352.1	**1313.9**
Hit rate		0.745		0.760	0.804	0.816	0.833
Vuong vs OP					-6.98**	-8.31**	-11.41**
Vuong vs ZIOP						-1.92	-4.81**
Vuong vs ZIOPa							-3.94**

Notes: **/* denote statistical significance at 1/5 percent level, respectively. For computations of information criteria and *Vuong* test statistics see Section 2.5.

The estimations of seven models with HD specification are reported in Table 5[21]. The far more parsimonious HD specification demonstrates a rather good fit, and is preferred over the FE specification by *BIC*, *AICc* and *HQIC* for all the models, and by *AIC* for the CNOP and ZIOP(a) models. Again, the two- and three-equation models have far better fits than the single-equation ones. The CNOP model is now overwhelmingly superior to all of the others (including now also the CNOPC model) according to all information criteria. The *Vuong* tests of the CNOP model versus the ZIOP and OP models are in favor of the CNOP model at the 1% significance level, and versus the ZIOP(a) model at the 5% level. The CNOPC model exhibits insignificant increase in the likelihood according to the *LR* test (p-value is 0.65). The estimated correlation coefficients ρ^- and ρ^+ (and their standard errors in parentheses) are -0.48(0.64) and 0.25(0.29), respectively.

The details for the specifications and estimated coefficients of the OP, ZIOP, CNOP and CNOPC models are presented in Tables 28 and 29 of Appendix C for the FE specification, and Table 30 of Appendix C for the HD specification. I only briefly discuss the estimated coefficients focusing instead on the marginal effects of the explanatory variables on the choice probabilities. The policy inclination decisions of the CNOP(C) models indeed appear to be driven by reaction to the economic situation. All the coefficients on the changes to

-702.4 (0.744) and -643.6 (0.764), respectively – very to those of the OP and MNP counterparts.

[21] The loglikelihood (and hit rate) of the ordered logit and multinomial logit models (not reported) are -721.9 (0.759) and -685.6 (0.762), respectively – very to those of the OP and MNP counterparts.

inflation (Δcpi_t), expected economic situation ($situation_t$), interest rate spread ($spread_t$), and recent change to the ECB policy rate ($\Delta ecbr_t$) are statistically significant at the 1% level, and have the expected positive signs. The coefficients on both policy bias ($bias_{t-1}$) and dissent among the policymakers ($dissent_{t-1}$) are both insignificant if included into $\mathbf{X}$ (p-values are larger than 0.2), but are significant at the 1% level if included in $\mathbf{Z}^-$ (for both $bias_{t-1}$ and $dissent_{t-1}$) and $\mathbf{Z}^+$(for $bias_{t-1}$ only) in the FE specification of the CNOP model. In two equations of the ZIOP model, the coefficients on both $bias_{t-1}$ and $dissent_{t-1}$ have the opposite signs.

Table 5: Changes to policy rate: comparison of alternative models with specification including only two dummies, for hawkish and dovish policymakers

Model	OP	GOP	MNP	ZIOP	ZIOPa	CNOP	CNOPC
ln l(θ)	-715.1	-684.6	-682.4	-631.4	-586.6	-557.1	-556.7
# of parameters	12	22	22	23	23	22	24
AIC	1454.2	1413.2	1408.9	1308.8	1219.1	**1158.2**	1161.3
BIC	1517.0	1528.4	1524.0	1429.2	1339.5	**1273.3**	1286.9
Corrected *AIC*	1529.0	1550.5	1546.0	1452.2	1362.5	**1295.3**	1310.9
HQIC	1477.7	1456.3	1451.9	1353.9	1264.2	**1201.3**	1208.3
Hit rate	0.749		0.773	0.783	0.793	0.828	0.829
Vuong vs OP				-5.92**	-7.57**	-8.48**	-8.69**
Vuong vs ZIOP					-4.19**	-4.50**	-4.65**
Vuong vs ZIOPa						-2.06*	-2.15*
LR vs MIOP							0.87

Notes: **/* denote statistical significance at 1/5 percent level, respectively. For computations of information criteria and *Vuong* test statistics see Section 2.5.

The coefficient on the last change to the NBP policy rate ($\Delta nbpr_{t-1}$) is statistically significant at the 1% level in all equations of the OP, ZIOP and CNOP(C) models. This variable represents the endogenous partial adjustment of policy rate, due to the intentional interest-rate smoothing and intrinsic gradualism of central bank behavior. We expect a positive coefficient in a single-equation OP model; therefore, for example, in the case of a hike to the rate, the probability of a hike/cut at the next meeting should be larger/smaller, *ceteris paribus*, than in the case of a cut. However, the coefficient on $\Delta nbpr_{t-1}$ has a negative sign in the OP model. Using the OP models, one would conclude that the larger the hike to the rate at the last meeting, the more likely is a cut at the next meeting. Arguably, this nonsensical result is due to an assumption of the single-equation OP model that the effect of a particular variable is homogenous.

In the CNOP(C) models we assume that the observed changes to the rate are the result of three distinct decisions, on which a given variable may have the opposite effects. We expect a high level of persistency in the latent policy stance, because of the slow cyclical fluctuations of macroeconomic indicators that exogenously drive the policy stance. Moreover, the central bank is conservative and dislikes frequent reverses in the direction of interest rate changes. Therefore, we expect a positive coefficient on $\Delta nbpr_{t-1}$ in the inclination equation.

However, we expect the negative coefficients in the amount equations for the same reasons of policy gradualism and inertia. The amount decisions are conditional on the policy stance, and are unidirectional: nonpositive or nonnegative, if the policy stance is loose or tight, respectively. The policymakers are cautious and tend to "wait and see", once they have moved the rate. Therefore, we expect a negative sign of the coefficient on $\Delta nbpr_{t-1}$ in the amount equations. As a matter of fact, while the rate changes are positively correlated during the whole sample (the first-order autocorrelation coefficient is 0.22), they are actually negatively correlated during the tightening and easing sub-periods (the autocorrelation coefficients are -0.22 and -0.05, respectively).

Table 6: Changes to the policy rate: partial effects of covariates on probabilities in the OP, ZIOP and CNOP models

	Pr(Δy_{it} = "decrease")			Pr(Δy_{it} = "no change")			Pr(Δy_{it} = "increase")		
	OP	ZIOP	CNOP	OP	ZIOP	CNOP	OP	ZIOP	CNOP
$spread_t$	-0.137***	-0.004	-0.287***	0.122***	0.014	0.269***	0.015***	-0.010	0.018**
	(0.015)	(0.015)	(0.038)	(0.016)	(0.055)	(0.038)	(0.004)	(0.040)	(0.009)
$\Delta ecbr_t$	-0.057***	0.051	-0.061***	0.044***	-0.197*	0.042***	0.013***	0.146*	0.018***
	(0.009)	(0.044)	(0.014)	(0.008)	(0.103)	(0.015)	(0.004)	(0.076)	(0.006)
$situation_t$	-0.136*	-0.164	-0.613***	0.121*	0.635	0.574***	0.015	-0.471	0.039**
	(0.076)	(0.166)	(0.129)	(0.067)	(0.424)	(0.125)	(0.010)	(0.303)	(0.019)
Δcpi_t	-0.172***	0.179	-0.458***	0.153***	-0.695**	0.429***	0.019***	0.515**	0.029**
	(0.022)	(0.146)	(0.078)	(0.023)	(0.329)	(0.078)	(0.005)	(0.252)	(0.013)
$\Delta nbpr_{t-1}$	0.021***	0.102	-0.065***	-0.019***	-0.190***	0.058***	-0.002***	0.088*	0.007*
	(0.004)	(0.062)	(0.015)	(0.004)	(0.069)	(0.016)	(0.001)	(0.046)	(0.004)
$I(h)_{it}$	-0.089***	-0.037	-0.067***	0.043***	0.168	-0.037	0.046***	-0.131	0.103***
	(0.012)	(0.036)	(0.018)	(0.014)	(0.113)	(0.024)	(0.010)	(0.096)	(0.028)
$I(d)_{it}$	0.174***	-0.044	0.162***	-0.167***	0.124	-0.159***	-0.006***	-0.081	-0.003*
	(0.038)	(0.039)	(0.042)	(0.038)	(0.101)	(0.042)	(0.002)	(0.079)	(0.002)
$bias_{t-1}$	-0.101***	-0.035	-0.062***	0.005	0.015	0.049***	0.096***	0.019	0.012*
	(0.013)	(0.028)	(0.017)	(0.019)	(0.012)	(0.016)	(0.013)	(0.018)	(0.007)
$dissent_{t-1}$	-0.231***	-0.119	-0.042***	0.206***	0.039	0.036***	0.025***	0.080	0.006
	(0.044)	(0.093)	(0.014)	(0.042)	(0.037)	(0.013)	(0.007)	(0.071)	(0.004)
$I(cpi^e > tar)_t$	-0.017	-0.090		0.016	-0.003	-0.003*	0.002	0.093	0.003*
	(0.017)	(0.073)		(0.016)	(0.010)	(0.002)	(0.002)	(0.071)	(0.002)

Notes: For definitions of variables see Table 3. ***/**/* denote statistical significance at 1/5/10 percent level, respectively. Robust to serial dependence asymptotic standard errors are in parentheses. Partial effects of a given covariate are computed for the HD specification holding the other covariates at their sample median values.

Indeed, the coefficient on $\Delta nbpr_{t-1}$ has a positive sign in the policy inclination equation, but the negative signs in the amount equations of the CNOP(C) models. It means that the rate hike/cut at the last meeting increases the probability of the tight/loose policy stance at the next meeting (compared to the cut/hike), but reduces the probability of the rate hike/cut conditional on the tight/loose policy regime. Obviously, the OP model fails to disentangle the opposite directions of the effect of $\Delta nbpr_{t-1}$ on the inclination and amount decisions. In the ZIOP model, this effect has a positive sign in both equations – a nonsensical result. The ZIOP(a) model is able to produce more sensible results: the sign is positive in the participation equation, but negative in the amount equation.

The partial (marginal) effects on the probabilities in the OP, ZIOP and CNOP models are compared in Table 6. In contrast to the OP and CNOP models, the estimated effects in the ZIOP model are mostly insignificant, and, if they are significant, often have the opposite sign than those in the CNOP model. For example, results based on the ZIOP model suggest that a half percent month-over-month increase in the annual rate of inflation – holding the rest of the explanatory variables at their sample median values – implies a 0.347 fall in the probability of no change to the rate at the next meeting. This contrasts with the OP and CNOP model results where we can conclude that it would actually lead to, respectively, a 0.077 and a 0.214 rise in the probability of no change.

Table 7: Changes to the policy rate: decomposition of partial effects of covariates on Pr(Δy=0) into three components

Covariates	Pr(Δy_{it} = "no change")		
	Loose stance	Neutral stance	Tight stance
$spread_t$	-0.101 (0.017)***	0.274 (0.068)***	0.095 (0.042)**
$\Delta ecbr_t$	-0.021 (0.006)***	-0.033 (0.033)	0.097 (0.025)***
$situation_t$	-0.215 (0.049)***	0.586 (0.182)***	0.204 (0.091)**
Δcpi_t	-0.161 (0.031)***	0.437 (0.128)***	0.152 (0.063)**
$\Delta nbpr_{t-1}$	-0.023 (0.006)***	-0.090 (0.060)	0.171 (0.050)***
$I(h)_{it}$	-0.020 (0.006)***	-0.084 (0.048)*	0.067 (0.030)**
$I(d)_{it}$	-0.006 (0.007)	-0.139 (0.044)***	-0.014 (0.007)*
$bias_{t-1}$	0.062 (0.017)***		-0.012 (0.007)*
$dissent_{t-1}$	0.042 (0.014)***		-0.006 (0.004)
$I(cpi^e>tar)_t$			-0.003 (0.002)*

Notes: For definitions of variables see Table 3. ***/**/* denote statistical significance at 1/5/10 percent level, respectively. Robust to serial dependence asymptotic standard errors are in parentheses. Partial effects of a given covariate are computed for the CNOP model with HD specification holding the other covariates at their sample median values.

The key differences are in the effects of the recent change to the policy rate ($\Delta nbpr_{t-1}$). The OP and CNOP models have the opposite signs of the *PE* of $\Delta nbpr_{t-1}$ on the probabilities of all three alternatives, and the ZIOP and CNOP models – on the probabilities of two choices. For example, the *PE* of $\Delta nbpr_{t-1}$ on the probability of rate hike/cut is positive/negative in the CNOP model, but negative/positive in the OP model. Results, based upon the OP and ZIOP models, suggest that a 25 bp increase in the recent change to the policy rate results, respectively, in a 0.019 and a 0.19 fall in the probability of no change to the rate at the next meeting, holding the rest of the explanatory variables at their sample median values. However, basing our forecast on the CNOP model, we would conclude that it would actually lead to a 0.058 rise in the probability of no change at the next meeting. In contrast to the CNOP model, the PEs of $\Delta nbpr_{t-1}$ in the OP and ZIOP models are not consistent with the observed gradualism of monetary policy decisions.

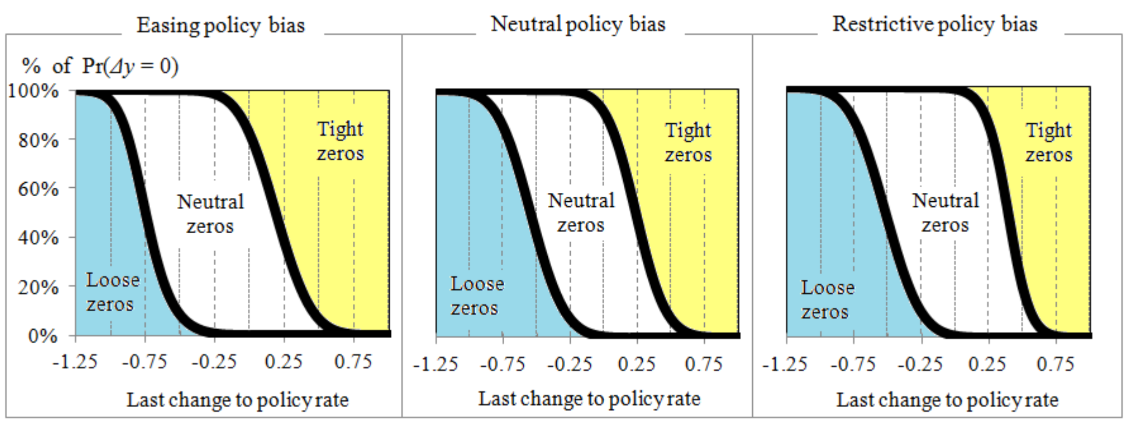
Figure 3: Changes to policy rate: predicted probabilities as functions of rate change and policy bias at the last MPC meeting from the OP and CNOP models

Notes: Predicted probabilities from the OP and CNOP models with HD specification are computed for the range of $\Delta nbpr_{t-1}$ and three possible values of policy bias at the previous MPC meeting, holding $I(h)_{it}$ and $I(d)_{it}$ at 1 and 0, respectively, and the rest of the explanatory variables at their sample median values.

The partial effects on the unconditional probability of no change to the rate $\Pr(\Delta y_{it} = 0)$ in the CNOP(C) models can be decomposed into three components – $\Pr(\Delta y_{it} = 0|r_{it} = -1)$, $\Pr(\Delta y_{it} = 0|r_{it} = 0)$ and $\Pr(\Delta y_{it} = 0|r_{it} = 1)$ – generated by the loose, neutral and tight policy regimes. For example, as Table 7 reports, the 0.43(0.08) PE of the change to inflation Δcpi_t on $\Pr(\Delta y_{it} = 0)$ in the CNOP model is a combined result of three opposing components: the -0.16(0.03), 0.44(0.13) and 0.15(0.06) PEs in the loose, neutral and tight policy regimes, respectively. Or, the 0.058(0.016) PE of the recent change to the policy rate $\Delta nbpr_{t-1}$on $\Pr(\Delta y_{it} = 0)$ is a sum of the -0.023(0.006), -0.090(0.060) and 0.171(0.050) PEs in the loose, neutral and tight policy stances, respectively[22].

Figure 3 contrasts the predicted probabilities from the OP and CNOP models for the range of $\Delta nbpr_{t-1}$ and three possible values of policy bias at the previous MPC meeting (easing, neutral or restrictive), holding $I(h)_{it}$ and $I(d)_{it}$ at 1 and 0, respectively, and the rest of the explanatory variables at their sample median values. The predicted probabilities from the OP and CNOP models exhibit striking differences. In the OP model, the probabilities of all three choices change monotonically through almost the entire range of $\Delta nbpr_{t-1}$. In contrast, the predicted probabilities in the CNOP model reveal non-monotonic patterns. For example, in the OP model under the easing policy bias $\Pr(\Delta y_{it} = 0)$ monotonically decreases if $\Delta nbpr_{t-1}$ increases, whereas in the CNOP model $\Pr(\Delta y_{it} = 0)$ is decreasing if $\Delta nbpr_{t-1}$ is less than 100 bp, but is increasing sharply otherwise and becoming closer and closer to one for values of $\Delta nbpr_{t-1}$ greater than 0 bp. Clearly, such a non-monotonic relationship is overlooked in the marginal effects from the OP model.

Figure 4: Changes to policy rate: decomposition of Pr(Δy=0) into three components as function of rate change and policy bias at the last MPC meeting



Notes: Predicted probabilities from the CNOP model with HD specification are computed for the range of $\Delta nbpr_{t-1}$ and three possible values of policy bias at the previous MPC meeting, holding $I(h)_{it}$ and $I(d)_{it}$ at 1 and 0, respectively, and the rest of the explanatory variables at their sample median values.

The decomposition of $\Pr(\Delta y_{it} = 0)$ into three components (the loose, neutral and tight zeros) is illustrated in Figure 4 for the range of $\Delta nbpr_{t-1}$ and three values of $bias_{t-1}$. The graphs show, for example, that if the rate was increased at the last meeting by 25 bp and if the policy bias was easing, then $\Pr(\Delta y_{it} = 0)$ is composed, on average, by 37% of the neutral zeros and 63% of the tight zeros. If the policy bias was neutral, then $\Pr(\Delta y_{it} = 0)$ is composed of 46% of the neutral zeros and 54% of the tight zeros. If the policy bias was

[22] Standard errors are in parentheses.

tight, then $\Pr(\Delta y_{it} = 0)$ is composed of 84% of the neutral zeros and 16% of the tight zeros.

The classification tables for the OP and CNOP models with both specifications are contrasted in Table 8. Compared to the OP model, the CNOP model with both specifications demonstrates the drastic improvement in the correct predictions of cuts (from 49% in the OP to 79% in the CNOP model with FE specifications) and hikes to the rate (from 65% to 77%). The CNOP model predicts fewer zeros (876) than the OP model (980), but predicts more zeros *correctly* (767 or 86%) than the OP model (only 758 or 85%). The simple OP model, as typical, tends to over-predict the most observed outcome, i.e. no-change decision in our case. The "adjusted noise-to-signal ratio" is clearly lower in the CNOP than in the OP model for the cuts (9% versus 19%) and no-change outcomes (25% versus 52%), and is rather low (4%-5%) for the hikes in both models.

Table 8: Changes the policy rate: classification tables of observed and predicted outcomes for the OP and CNOP models

Actual outcomes	Predicted outcomes: Cut	Predicted outcomes: No change	Predicted outcomes: Hike	Hit rate	Adj. noise to signal ratio	Predicted outcomes: Cut	Predicted outcomes: No change	Predicted outcomes: Hike	Total	Hit rate	Adj. noise to signal ratio
			OP model					**CNOP model**			
			Specification with fixed effects (twenty individual dummies)								
Cut	**152**	157	0	0.49	0.19	**243**	66	0	309	0.79	0.09
No change	103	**758**	28	0.85	0.52	74	**767**	48	889	0.86	0.25
Hike	0	65	**122**	0.65	0.04	0	43	**144**	187	0.77	0.05
Total	255	980	150	0.75		317	876	192	1385	0.83	
			Specification with two dummies for hawkish and dovish members								
Cut	**167**	142	0	0.56	0.16	**237**	72	0	309	0.79	0.09
No change	94	**746**	49	0.83	0.49	72	**771**	46	889	0.86	0.28
Hike	0	62	**125**	0.67	0.06	0	48	**139**	187	0.74	0.05
Total	261	950	174	0.75		309	891	185	1385	0.83	

Notes: A particular choice is predicted if its predicted probability exceeds the predicted probabilities of the alternatives. An "adjusted noise-to-signal ratio", introduced by Kaminsky and Reinhart (1999), is defined as follows. Let A denote the event that the decision is predicted and occurred; let B denote the event that the decision is predicted but not occurred; let C denote the event that the decision is not predicted but occurred; let D denote the event that the decision is not predicted and not occurred. The desirable outcomes fall into categories A and D, while noisy ones fall into categories B and C. A perfect prediction would have no entries in B and C, while a noisy prediction would have many entries in B and C, but few in A and D. The "adjusted noise-to-signal" ratio is defined as $[B/(B+D)]/[A/(A+C)]$.

The estimated predicted probabilities of three latent policy regimes, averaged across all MPC members, are shown for each policy meeting in Figure 5 of Appendix C together with the policy rate decisions made by the MPC. Averaged over entire sample, they are 0.51, 0.18 and 0.31 for the loose, neutral and tight policy stances, respectively, as Table 9 reports. These probabilities are also computed separately for the periods of policy easing, maintaining and tightening, as well as separately for the MPC decisions to change the rate or to leave it unchanged during the easing and tightening periods. The computations reveal that the average probability of the neutral policy stance during the maintaining periods is only 0.11. In spite of the 0.54 probability of the loose stance and 0.35 probability of the tight stance, the MPC maintained the rate for long periods between the rate reversals. The no-change decisions of the MPC during the easing periods were generated with the 0.52/0.43/0.05 probabilities of the loose/neutral/tight policy stance, while during the tightening periods these probabilities were 0.31/0.01/0.67, respectively. The MPC decisions to reduce the rate were generated with the 0.72/0.28/0.00 probabilities of the loose/neutral/tight policy stance. This closely mimics the observed 0.76/0.24/0.00 frequencies of the individual policy preferences to reduce/maintain the rate during the MPC decisions to cut it. The MPC decisions to increase the rate were generated with the 0.10/0.00/0.90 probabilities of the loose/neutral/tight policy stance. The observed frequencies of the individual policy preferences to reduce/maintain/increase the rate during the MPC decisions to hike it were 0.00/0.13/0.87, respectively, suggesting that the policy stances of the dissenters were actually loose during the MPC decisions to hike the rate.

Table 9: The individual policy decisions and predicted probabilities of individual policy stances during the periods of policy easing, maintaining and tightening

Policy period	MPC decision	Average predicted probabilities of individual latent policy stances			Observed frequencies of individual voted policy decisions		
		loose $\Pr(r_{it}=-1)$	neutral $\Pr(r_{it}=0)$	tight $\Pr(r_{it}=1)$	decrease	no change	increase
Easing	decrease	0.72	0.28	0.00	0.76	0.24	0.00
	no change	0.52	0.43	0.05	0.06	0.94	0.00
Maintaining	no change	0.54	0.11	0.35	0.01	0.93	0.06
Tightening	no change	0.32	0.01	0.67	0.00	0.83	0.17
	increase	0.10	0.00	0.90	0.00	0.13	0.87
Whole	all	0.51	0.18	0.31	0.22	0.64	0.14

Notes: The policy easing, maintaining and tightening periods are shown in Figure 1. The predicted probabilities are from the CNOP model with HD specification.

These findings are further confirmed and refined in Table 10, which reports the average predicted probabilities of individual policy stances separately for the individual policy decisions of MPC members to reduce, maintain or increase the rate. The average probability of the neutral policy stance during the individual no-change decisions is only 0.20. Table 10 also reports the decomposition of $\Pr(\Delta y_{it} = 0)$ into three parts – $\Pr(\Delta y_{it} = 0|r_{it} = -1)$,$\Pr(\Delta y_{it} = 0|r_{it} = 0)$ and $\Pr(\Delta y_{it} = 0|r_{it} = 1)$ – corresponding to the "loose", "neutral" and "tight" zeros. The average predicted probability of no change during the observed no-change decisions is generated by the loose/neutral/tight policy regimes with the 0.48/0.25/0.27 probabilities, respectively. For the entire sample the average predicted probability of no change is decomposed as 0.46/0.28/0.26, respectively. The vast majority (about 75%) of observed zeros appeared to be generated by the tight or loose policy stances, offset by the amount decisions at the second stage. These findings suggest a high degree of the purposeful inertia in the policy-making process of the NBP: only a quarter (at most) of observed no-change decisions appears to be generated by neutral policy reaction to key macroeconomic indicators such as inflation, real activity and ECB policy rate. The OP and ZIOP models have failed to detect this, and produced the biased estimates of the *PE*s of $\Delta nbpr_{t-1}$.

Table 10: Predicted probabilities of individual policy stances and decomposition of Pr(Δy=0) into three policy regimes

Individual voted policy decision	Average predicted probabilities of individual policy stances			Decomposition of $\Pr(\Delta y_{it} = 0)$		
	loose $\Pr(r_{it} = -1)$	neutral $\Pr(r_{it} = 0)$	tight $\Pr(r_{it} = 1)$	"loose zeros"	"neutral zeros"	"tight zeros"
Cut	0.77	0.23	0.00	0.33	0.67	0.00
No change	0.51	0.20	0.30	0.48	0.25	0.27
Hike	0.14	0.00	0.86	0.39	0.01	0.61
All	0.51	0.18	0.31	0.46	0.28	0.26

Notes: The predicted probabilities are from the CNOP model with HD specification.

6 Conclusions

"The model is often smarter than you are. ...(T)he act of putting your thoughts together into a coherent model often forces you into conclusions you never intended..."

-Paul Krugman[23]

Ordinal responses, when the decisionmakers face the choices to reduce, to leave unchanged or to increase (e.g., prices, consumption, ratings or policy interest rates) or when they must indicate the negative, neutral or positive attitudes or opinions, are often characterized by abundant observations in the middle neutral or zero category (no change or indifferent attitude). Such excessive "zeros" can be generated by different groups of population or by separate decision-making processes. Besides, the positive and negative outcomes can be driven by distinct sources too. In such situations, it would be a misspecification to treat all the observations as coming from the same d.g.p., and to apply a standard single-equation model. This paper develops a more flexible cross-nested ordered probit model for such types of ordinal outcomes, combining three OP latent equations with different sets of covariates.

The proposed CNOP(C) models allow the separate mechanisms to determine what I call the *inclination* decision ($y \leq 0$ versus $y = 0$ versus $y \geq 0$, interpreted as a loose,

[23] From the essay "Delusions of Growth" in Krugman (1999).

neutral or tight policy stance) and two *amount* decisions, conditional on the loose or tight policy stance (the magnitude of y when it is nonpositive or nonnegative, respectively). The inclination decision is driven by the direct reaction to the changes in the macroeconomic environment, whereas the amount decisions allow policy stance to be offset by the tactical and institutional features of policymaking process. This three-regime approach is able to discriminate among the following three types of zeros: the "always" or "neutral" zeros, generated directly by the neutral policy reaction to the economic developments; and two kinds of "not-always" or "offset" zeros, the "loose" and "tight" zeros, generated by the loose or tight policy inclinations offset by the tactical reasons. The model also allows for the possible correlation among three latent decisions.

The Monte Carlo results suggest good performance of the model in the small samples and demonstrate its superiority with respect to the conventional OP model, which produces biased estimates of the discrete-choice probabilities and of the marginal effects of the covariates on those probabilities, if the underlying d.g.p. is heterogeneous. Although the proposed approach indeed tends to require larger sample sizes than the usual OP model, due to the heavier parameterization involved, the simulations suggest that the CNOP(C) models provide accurate and reliable inference even in small samples (about 200 observations) with a mixture of different d.g.p's.

The CNOP(C) models are applied to explain policy rate decisions of the National Bank of Poland, using the panel of the individual votes of the MPC members and real-time macroeconomic data available at the policy-making meetings. The voting preferences appeared to be well modelled by proposed two-step three-regime approach. The real-world data favor the CNOP model. Not only does it fit the data much better, but it also has some important advantages over the single- and two-equation models, such as the standard and generalized OP, multinomial probit and zero-inflated OP models.

In particular, the CNOP(C) models are able to identify the driving factors of each decision. For example, the rate change, made at the previous MPC meeting, has the opposing impacts on the inclination and amount decisions. The conventional OP models are shown to confuse the marginal effects of the explanatory variables that only have an impact on one decision or opposing impacts on two decisions. In addition, the proper estimation of the marginal effects of the explanatory variables is shown to exhibit the presence of a non-monotonic relationship between these variables and outcome probabilities. It might have the important implications for the statistical inference, since the OP model fails to detect such non-monotonic patterns.

The CNOP(C) models are able to estimate the probabilities of three types of zeros and how this decomposition depends on the observed data. The vast majority (about 75%) of observed zeros appeared to be generated by the tight or loose policy stances, offset by the inertial amount decisions. These findings suggest a high degree of intentional interest-rate smoothing in the decision-making process of the NBP: only a quarter of observed no-change decisions appears to be explained by neutral policy reaction to economic conditions.

It is quite plausible that the small changes to the rate can be also inflated and characterized by two types of observations, coming either from the loose (tight) or neutral policy stance. By adding the third amount equation, conditional on the neutral policy regime, with three outcome categories (no change, small cut and small hike), the resulting four-part CNOP(C) models will allow for heterogeneity in the three middle categories.

References

Bernanke, B. S. (2007). "Federal Reserve communications." Speech at the Cato Institute 25th Annual Monetary Conference, Washington DC, November 14.

Brooks, R., M. Harris, and C. Spencer (2012). "Inflated ordered outcomes." *Economics Letters* 117, 683-686.

Dow, W. and E. Norton (2003). "Choosing between and interpreting the Heckit and two-part models for corner solutions." *Health Services and Outcomes Research Methodology* 4(1), 5–18.

Eichengreen, B., M. Watson, and R. Grossman (1985). "Bank rate policy under the interwar gold standard: a dynamic probit model", *Economic Journal* 95, September, 725-745.

Gerlach-Kristen, P. (2004). "Is the MPC's voting record informative about future UK monetary policy?" *Scandinavian Journal of Economics* 106(2), 299-313.

Cragg, J. G. (1971). "Some statistical models for limited dependent variables with application to the demand for durable goods." *Econometrica* 39(5), September, 829-844.

Greene, W. (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. (2004). "Convenient estimators for the panel probit model." *Empirical Economics* 29(1), 21-47.

Gronau, R. (1974). "Wage comparisons — a selectivity bias." *The Journal of Political Economy* 82(6), 1119-1143.

Hamilton, J. D. and O. Jorda (2002). "A model for the federal funds rate target." *Journal of Political Economy* 110(5), October, 1135-1167.

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

Heckman, 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 Social Measurement* 5, 475–492.

Heckman, J. (1979). "Sample selection bias as a specification error." *Econometrica* 47, 53–161.

Jones, A. M. (2000). "Health econometrics" in *Handbook of Health Economics*, Vol. 1A, Culyer, A., Newhouse J. (Eds.), North Holland.

Kaminsky, G. L. and C. M. Reinhart (1999). "The twin crises: the causes of banking and balance-of-payments problems." *The American Economic Review* 89(3), June, 473-500.

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.

Krugman, P. (1999). "*The accidental theorist and other dispatches from the dismal science.*" W.W. Norton & Company, Inc., New-York.

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

Lancaster T.(2000) "The incidental parameters problem since 1948." *Journal of Econometrics* 95(2), 391-414.

Leung, S. F. and S. Yu (1996). "On the choice between sample selection and two-part models." *Journal of Econometrics* 72, 197-229.

Madden, D. (2008). "Sample selection versus two-part models revisited: The case of female smoking and drinking." *Journal of Health Economics* 27, 300–307.

Mullahy, J. (1986). "Specification and testing of some modified count data models." *Journal of Econometrics* 33, 341-365.

Neyman, J. and E. Scott (1948). "Consistent estimates based on partially consistent observations." *Econometrica* 49, 1417-1426.

Poole, W. (2003). "Fed transparency: how, not whether", Federal Reserve Bank of St. Louis *Review* 85(6), November/December, pp. 1-8

Poole, W. (2006). "The Fed's monetary policy rule", Federal Reserve Bank of St. Louis *Review* 88(1), January/February, pp. 1-11

Schmidt, P. and A. D. Witte (1989). "Predicting criminal recidivism using "split population" survival time models." *Journal of Econometrics* 40(1), 141-159.

Sirchenko, A. (2008). "Modelling monetary policy in real time: Does discreteness matter?" Economics Education and Research Consortium working paper No. 08-07, July.

Sirchenko, A. (2010). "Policymakers' votes and predictability of monetary policy." UCSD Economics working paper No. 1672194, December.

Small, K. (1987). "A discrete choice model for ordered alternatives." *Econometrica* 55, 409-424.

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.

Wooldridge, J. M. (2010). "*Econometric analysis of cross section and panel data.*" 2^{nd} edition, MIT Press.