Modeling status quo decisions: A cross-nested ordered probit model

Andrei Sirchenko*
Higher School of Economics, Moscow
European University Institute, Florence
E-mail: asirchenko@hse.ru

February 11, 2015

Abstract

The decisions to reduce, leave unchanged, or increase a choice variable (such as policy interest rates) are often characterized by abundant status quo outcomes that can be generated by different processes. The decreases and increases may also be driven by distinct decision-making paths. Neither standard nor zero-inflated models for ordinal responses adequately address these issues. This paper develops a flexible mixture model with endogenously switching regimes. Three latent regimes, which are interpreted in the interest rate setting context as loose, neutral and tight policy stances, create separate processes for rate hikes and cuts and overlap at a status quo outcome, generating three different types of zeros. The new model exhibits acceptable small-sample performances in Monte Carlo experiments, whereas traditional models deliver biased estimates. In the empirical application, the new model is not only highly favored by the statistical tests but also produces economically more meaningful inference with respect to existing models. More than one-third of the status quo decisions are generated by the loose or tight policy stances, suggesting a high degree of intentional interest rate smoothing.

JEL classification: C33; C35; E52.

Keywords: ordinal responses, zero-inflated outcomes, three-part mixture model, endogenous regime switching, policy interest rate, MPC votes, real-time data.

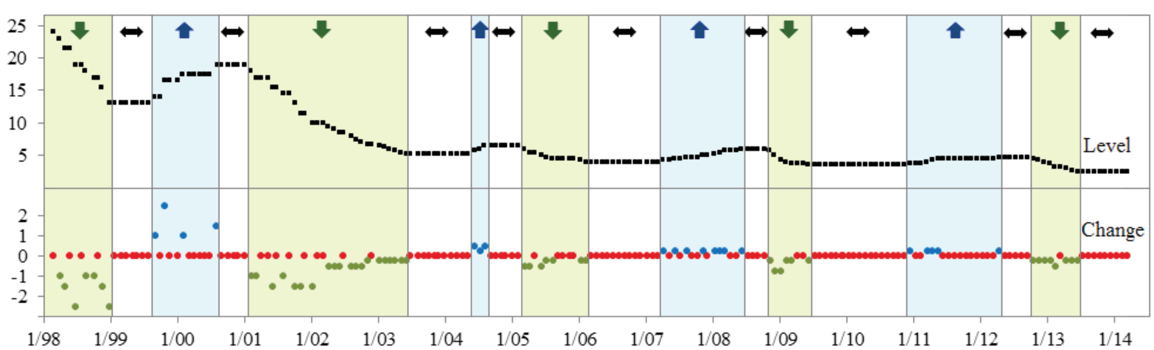
*Previous versions of this paper have been circulated under the title "A model for ordinal responses with an application to policy interest rate" as the EERC working paper No. E12/13, and as the NBP working paper No. 148.

1 Introduction

"To do nothing is sometimes a good remedy." – Hippocrates

Ordinal responses, when decision-makers face a choice to reduce, leave unchanged or increase a price (consumption, rating, or policy interest rate) are often characterized by abundant observations in the middle no-change category. Most central banks adjust policy rates by discrete increments (multiples of a quarter of a percent), and no-change decisions commonly constitute an absolute majority.[1] The preponderance of status quo responses (zeros) in many data sets suggests that zeros may emerge from fundamentally different behavioral mechanisms. For instance, the policy rates of many central banks typically remain unchanged in three different circumstances, namely: in periods of policy tightening; in periods of maintaining between rate reversals; and in periods of easing (see Figure 1).[2] Many of the zeros, situated between rate hikes during policy tightening, are likely to be driven by different economic conditions compared with many of those that are situated between cuts during policy easing. Many of the zeros, clustered between rate reversals during maintaining periods, are also likely to differ from status quo decisions during periods of easing or tightening.

Figure 1. The policy rate remains unchanged in different circumstances: during the periods of policy easing, maintaining and tightening

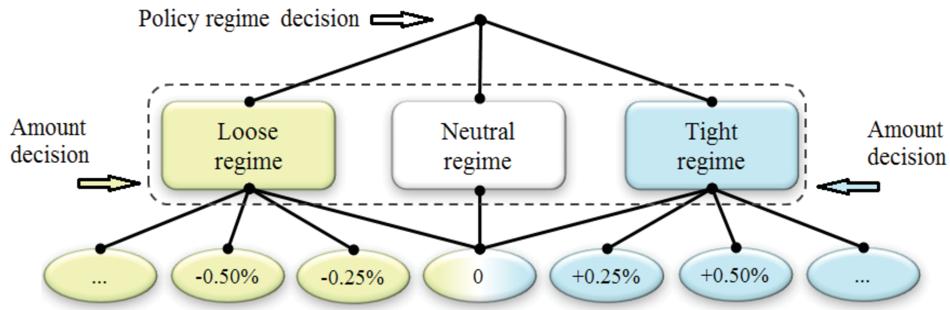


Notes: ⇓/⇔/⇑ denote the periods of policy easing/maintaining/tightening. The period of easing or tightening is the period during which the rate only moves only in one direction (down or up, respectively) between the first and last sequential unidirectional changes. The period of policy maintaining is the period between the rate reversals. The data correspond to the reference rate of the National Bank of Poland.

[1] For example, between 60 and 80 percent in the US Federal Reserve (Fed), the Bank of England (BoE) and the European Central Bank (ECB).

[2] See Figure 9 in online Appendix E for the cases of the Fed, the BoE and the ECB.

Figure 2. The decision tree of the CNOP model



The predominance and heterogeneity of zero observations poses a problem for standard discrete-choice methods such as the ordered probit (OP) model. This paper develops a more flexible cross-nested ordered probit (CNOP) model that accomodates the unobserved heterogeneity of a data-generating process (DGP) by assuming three implicit decisions, and illustrates the model in the context of interest rate adjustments. Figure 2 shows a two-stage decision tree of the proposed model, in which an ordinal dependent variable (for example, a discrete change to a policy rate) can exist in three latent regimes. The regime, that is the monetary policy stance (loose, neutral or tight), is determined by a *regime switching* (or *policy inclination*) decision, which serves the role of sample differentiation, and is endogenously driven by a direct policy reaction to current economic conditions. If the policy stance is neutral, no further actions are taken and the rate is maintained. If the stance is loose (tight), the policymakers can cut (hike) the rate by a certain amount or leave it unchanged. These unidirectional *amount* decisions determine the magnitude of the rate adjustment, intensifying or weakening the policy inclination, and are more of an institutional nature. The model simultaneously estimates the three OP equations, which represent the latent decisions, and allows for a possible correlation among them. Using this interpretation, we can classify the three types of zeros and describe how they arise: the "always" or "neutral" zeros, which are directly generated by a neutral reaction to economic conditions, and the two types of "not-always" zeros — the "loose" and "tight" zeros — which are generated by loose or tight policy inclinations and are offset by tactical and institutional reasons.

For example, despite a loose policy stance, policymakers can maintain the rate for the following reasons. First, the rate was already lowered at the last meeting (central banks are generally reluctant to move the rate on a frequent basis). Second, the dissenting policymakers at the last meeting preferred the higher rate, creating an upward pressure on the rate at the current meeting (monetary policy is commonly conducted by a committee that is often composed of heterogeneous members). Third, the recent "policy bias" statement of the central bank, indicating the most likely policy direction in the immediate future, was neutral or even tightening (policymakers are concerned about the competence and credibility of central bank communication). Fourth, the cumulative changes to the economic indicators since the last policy rate adjustment do not suggest policy easing (policymakers, who face uncertainty about the economy and incur the costs in the case of the subsequent rate reversal, prefer to wait and see and react to accumulated economic information). Finally, the policy rate has already reached the lower zero bound.

As discussed in Section 2, the proposed three-equation models can be estimated via maximum likelihood. The Monte Carlo experiments outlined in Section 3 suggest reasonable performance of the new model in the small samples (two hundred observations) and demonstrate its superiority with respect to the OP model, which typically overpredicts the most observed outcome (i.e. no-change response), produces the biased and inefficient estimates of the choice probabilities and the marginal effects of explanatory variables on these probabilities, if the underlying DGP is heterogeneous.

The conventional OP, the middle-inflated OP (Brooks *et al.* 2012, Bagozzi and Mukherjee 2012) and the new models are applied in Section 4 to explain the policy interest rate decisions of the National Bank of Poland (NBP) using a novel panel dataset, which contains the individual policy preferences (votes) of the Monetary Policy Council (MPC) members and *real-time* macroeconomic data available at the MPC meetings. According to the statistical tests and the information criteria, the real-world data overwhelmingly favor the new approach, which produces substantial improvements in statistical fit relative to the OP and middle-inflated OP models, is capable of extracting important additional information and provides an economically more reasonable inference.

In particular, the statistical rejection of the single-equation OP model provides compelling empirical evidence of the presence of heterogeneity in the DGP. The average estimated probability of a neutral policy stance is 0.41, whereas the observed frequency of status quo decisions is 0.65. Less than two-thirds of zeros seemed to be generated by a neutral policy reaction to economic conditions; the remaining zeros originate under the loose or tight policy inclination. More than fourty percent of all outcomes in the loose and tight regimes are zeros; the amount decisions tend to smooth the interest rate by weakening the up- and downward policy inclinations.

These findings suggest a considerable degree of deliberate interest-rate smoothing in the decision-making process of the NBP.

As a practical matter, the same explanatory variables can have different weights in the decisions to lower or increase the rate, which can be influenced by different direction-specific determinants. The empirical rejection of the middle-inflated OP model in favor of the CNOP model suggests that the effects of the explanatory variables on the decisions to reduce or raise the rate are asymmetric; combining these two distinct decisions into one branch of the decision tree, as implemented in the middle-inflated OP model, may seriously distort an inference.

The CNOP model also enables all variables to affect the regime switching and amount decisions in different ways. For instance, the coefficient on the previous change in the rate has a positive sign in the policy regime equation, whereas it has the negative signs in the amount equations. This enables the previous policy choice to have the same sign of the marginal effect on the probabilities of both a cut and a hike; by contrast, the single-equation structure of the OP model implies the opposite direction of these effects. A rate hike at the last meeting (relative to a status quo decision) expectedly lowers the probabilities of both a cut and a hike and raises the probability of no change according to the CNOP model; by contrast, it reduces the probabilities of a hike and no change but counterintuitively increases the probability of a cut according to the OP model. If a certain variable has an impact on both latent decisions, the OP model cannot reveal the distinct effects on the probabilities of different types of zeros (with respect to both a sign and a magnitude), incorrectly estimates that variable's total impact by focusing on the observed zeros, and produces the misleading estimates of the choice probabilities and marginal effects.

"Nobody likes change except a wet baby" — the preponderance of zero or neutral outcomes is a common phenomenon in many fields, including economics, political sciences, sociology, technometrics, medicine, psychology and biology. In studies of count and non-negative ordinal data (visits to a doctor, alcohol or tobacco consumption, disease lesions on plants, manufacturing defects, recreational demand, sexual behavior, and fertility) the abundance and heterogeneity of zero observations are well recognized. Numerous studies make a distinction between the different types of zeros — for example, no medical appointments due to chance, doctor avoidance, lack of insurance, or medical costs; no children due to infertility or choice; no illness due to strong resistance or lack of infection; and a "genuine nonuser" versus a "potential user" (for a review, see Greene and Hensher 2010 and Winkelmann 2008). Studies of survey responses using an odd-point Likert-type scale (in which respondents must indicate a negative, neutral or positive attitude) discuss the middle category endorsement and heterogeneity of indifferent responses — a true neutral option versus an ambivalent, uninformed, or inherently unordered "do-not-know" position that is commonly reported as neutral (Bagozzi and Mukherjee 2012, Hernández *et al.* 2004, Kulas and Stachowski 2009).

In decision-making experiments and micro-level studies of consumer choices, election votes and other repeated choices, the prevalence of no-change decisions is often attributed to the status quo bias — the tendency to do nothing or maintain a previous decision, athough it is not always objectively superior to other available options (Hartman *et al.* 1991, Kahneman *et al.* 1991). It is a cognitive bias that is explained by rational and irrational causes (Samuelson and Zeckhauser 1988). Due to the special features of monetary policy decision-making, such as publicity and transparency, a high level of expertise, reputation and responsibility among MPC members, and research and administrative support, we may disregard the irrational routes of observed monetary policy inertia and treat it as a rational decision. Policy inertia is often attributed to the intentional interest-rate-smoothing behavior of a central bank or the slow cyclical fluctuations of macroeconomic variables that exogenously drive policy actions (for debates on the degree of monetary policy inertia and its "illusion", see Coibion and Gorodnichenko, 2012, and Rudebusch 2002, 2006).

2 Econometric framework

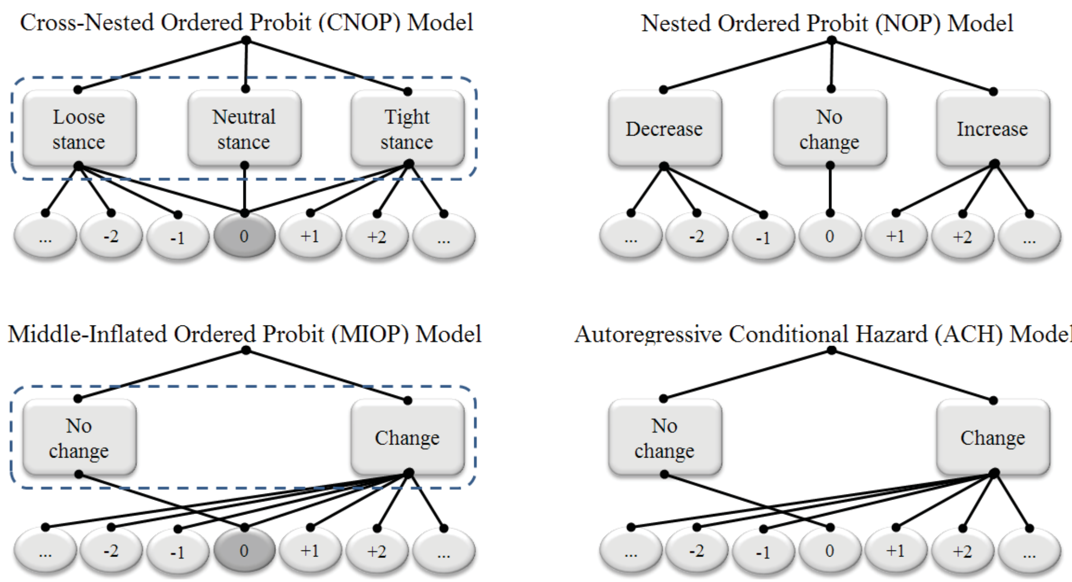
> "It is highly desirable that policy practice be formalized to the maximum possible extent." – W. Poole (2006)

The proposed CNOP model can be described as a cross-nested generalization of a two-level nested ordered probit (NOP) model with three nests (see the upper panel of Figure 3). In the case of unordered categorical data, in which the choices can be grouped into the nests of similar options, the nested logit model is prevalent. Several types of multinomial logit models with overlapping nests are proposed (Wen and Koppelman 2001, Vovsha 1997). The (cross)-nested models, specifically designed for the ordered alternatives, are not as common.[3] The NOP econometric framework is presented in online Appendix A. The difference between the decision trees of the NOP and CNOP models is that all three nests of the CNOP model overlap at a no-

[3]Small (1987) proposed "the ordered generalized extreme value model" for ordered outcomes. The model contains overlapping nests but each nest contains only two adjacent alternatives.

change response. In the NOP model, the decisions at both levels are observable and we always know to which of the three nests the observed outcomes belong, whereas the zeros are observationally equivalent in the CNOP model — we never know from which of the three regimes the zeros originate.

Figure 3. The CNOP model is a generalization of the NOP, MIOP and ACH models



The CNOP model can also be considered as a three-part zero-inflated model. The two-part zero-inflated models, developed to address both the abundant zeros and unobserved heterogeneity, include the zero-inflated Poisson model (Lambert 1992) and negative binomial model (Greene 1994) 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 responses. These models are the natural extensions of the two-part (or hurdle, or split-population) mixture models, proposed by Cragg (1971) for non-negative continuous data and subsequently developed for 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). The two-part model combines a binary choice 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). Its structure is similar to the structure of the discrete version of a sample selection model (Heckman 1979). However, in the sample selection model the first hurdle — the selection decision — determines whether a choice variable is observed, instead of whether an activity is undertaken as in the two-part model, in which all outcomes are actually observed and the first hurdle serves the role of sample differentiation (Leung and Yu 1996).

The ZIOP model is suitable for explaining decisions such as the levels of consumption, when the upper hurdle is naturally binary (to consume or not to consume) and the ordinal responses are typically non-negative. Thus, the abundant zeros are situated at one end of the ordered scale. Brooks *et al.* (2012) and Bagozzi and Mukherjee (2012) modified the ZIOP model and proposed the middle-inflated ordered probit (MIOP) model for an ordinal outcome, which ranges from negative to positive responses, and where an inflated outcome is situated in the middle instead of at one end of the choice spectrum. The difference between the two-part ACH model and the MIOP model (see the bottom panels of Figure 3) is that the two parts are separately estimated in the former and the zero observations are excluded from the second part; thus, the discrimination among the different typess of zeros is not accommodated. In the latter, the two nests overlap, assuming two types of zeros; thus, the probability of a zero is "inflated".

The three-part CNOP model is a natural generalization of the two-part ZIOP and MIOP models. A trichotomous participation decision (increase, no change, or decrease) seems to be more realistic than a binary decision (change or no change) if applied to ordinal data that assume negative, zero and positive values (such as changes to policy rates). The policymakers, who are willing to adjust the rate, have already decided the direction that they wish to move it. Combining these two distinct decisions into one category, as implemented in the two-part models, may seriously distort the inference, as documented in Section 4.

This section describes the econometric framework of the CNOP models, which are designed for an ordinal dependent variable with a minimum of three different values. 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 $\{-J, ..., 0, ..., J\}$, where $J \geq 1$ and a predominant (and potentially heterogeneous) response is coded as zero. The prevailing outcome does not have to be in the *very* middle of the ordered categories. However, if it is located at the *end* of the ordered scale, the three-part CNOP model reduces to the two-part ZIOP model.

Although the new models are suitable for large cross-sectional and longitudinal data, a sufficiently long discrete-valued time series is also applicable. The econometric framework is presented in a panel-data framework using a double subscript, where the index i denotes one of N cross-sectional units and the index t denotes one of T time periods. An application to cross-sectional or time-series data is straightforward by setting T or N to one. As an illustration, the model is interpreted in the context of interest rate decisions that are taken repeatedly at the policy-making meetings by each member of a monetary policy committee.

2.1 The cross-nested ordered probit model

The model assumes two stages (see Figure 2). At the first stage — the upper level of the decision tree — the continuous latent variable r^*_{it} represents the degree of the policymaker i's policy inclination. It is set at the meeting t in response to the observed data according to a *regime equation*

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

where $\mathbf{x}_{it}$ is the t^{th} row of the 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 that is independently and identically distributed (IID) across i and t.

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

$$r_{it} = -1 \text{ if } r^*_{it} \leq \alpha_1, \; 0 \text{ if } \alpha_1 < r^*_{it} \leq \alpha_2, \text{ and } 1 \text{ if } \alpha_2 < r^*_{it},$$

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

Under the assumption that the disturbance term ν_{it} has the cumulative distribution function (CDF) F, the probabilities of each possible outcome of r_{it} are given by

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

At the second stage — the lower level of the decision tree — three latent regimes exist. Conditional on being in the regime $r_{it} = 0$, no further policy actions are taken and the interest rate remains unchanged:

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

Thus, the conditional probability of the outcome j in the neutral policy stance is

$$\Pr(\Delta y_{it}=j|r_{it}=0)=\begin{cases}0 & \text{for} \quad j\neq 0,\\ 1 & \text{for} \quad j=0.\end{cases} \tag{3}$$

Conditional on being in the regime $r_{it}=\pm1$ ("$\pm$" denotes either "$+$" or "$-$", i.e., a tight or loose policy stance), a continuous latent variable Δy^*_{it}, representing the desired change to the rate, is determined by the direction-specific *amount equations*

$$\Delta y^*_{it}|(\mathbf{z}^{\pm}_{it},r_{it}=\pm1)=\mathbf{z}^{\pm\prime}_{it}\boldsymbol{\gamma}^{\pm}+\varepsilon^{\pm}_{it}, \tag{4}$$

where $\boldsymbol{\gamma}^{\pm}$ is a $K_{\gamma}\times 1$ vector of unknown coefficients, $\mathbf{z}^{\pm}_{it}$ is the t^{th} row of the observed $T_i\times K_{\gamma}$ data matrix $\mathbf{Z}^{\pm}_{i}$, and $\varepsilon^{\pm}_{it}$ is an IID error term with the CDF $F^{\pm}$.

The discrete change to the rate Δy_{it} is determined according to the rule

$$\Delta y_{it}|(\mathbf{z}^{\pm}_{it},r_{it}=\pm1)=j \text{ if } \mu^{\pm}_{j-1}<\Delta y^*_{it}|(\mathbf{z}^{\pm}_{it},r_{it}=\pm1)\leq\mu^{\pm}_{j} \text{ for } j=0,\pm1,...,\pm J,$$

where $-\infty\equiv\mu^{-}_{-J-1}\leq\mu^{-}_{-J}\leq...\leq\mu^{-}_{-1}\leq\mu^{-}_{0}\equiv\infty$ and $-\infty\equiv\mu^{+}_{-1}\leq\mu^{+}_{0}\leq...\leq\mu^{+}_{J-1}\leq\mu^{+}_{J}\equiv\infty$ are the $2J$ unknown thresholds to be estimated.

The conditional probability of the outcome j can be computed as

$$\Pr(\Delta y_{it}=j|\mathbf{z}^{\pm}_{it},r_{it}=\pm1)$$

$$=\begin{cases}\frac{1}{2}(1-r_{it})[F^{-}(\mu^{-}_{j}-\mathbf{z}^{-\prime}_{it}\boldsymbol{\gamma}^{-})-F^{-}(\mu^{-}_{j-1}-\mathbf{z}^{-\prime}_{it}\boldsymbol{\gamma}^{-})] & \text{if } j<0,\\ F^{\pm}(\mu^{\pm}_{j}-\mathbf{z}^{\pm\prime}_{it}\boldsymbol{\gamma}^{\pm})-F^{\pm}(\mu^{\pm}_{j-1}-\mathbf{z}^{\pm\prime}_{it}\boldsymbol{\gamma}^{\pm}) & \text{if } j=0,\\ \frac{1}{2}(1+r_{it})[F^{+}(\mu^{+}_{j}-\mathbf{z}^{+\prime}_{it}\boldsymbol{\gamma}^{+})-F^{+}(\mu^{+}_{j-1}-\mathbf{z}^{+\prime}_{it}\boldsymbol{\gamma}^{+})] & \text{if } j>0.\end{cases} \tag{5}$$

Assuming that ν_{it}, ε^{-}_{it} and ε^{+}_{it} are independent, the full probabilities are given by combining the probabilities in (2), (3) and (5):

$$\Pr(\Delta y_{it}=j|\mathbf{x}_{it},\mathbf{z}^{-}_{it},\mathbf{z}^{+}_{it})=\begin{cases}I_{j\leq0}\Pr(r_{it}=-1|\mathbf{x}_{it})\Pr(\Delta y_{it}=j|\mathbf{z}^{-}_{it},r_{it}=-1)\\ +I_{j=0}\Pr(r_{it}=0|\mathbf{x}_{it})\Pr(\Delta y_{it}=j|\mathbf{x}_{it},r_{it}=0)\\ +I_{j\geq0}\Pr(r_{it}=1|\mathbf{x}_{it})\Pr(\Delta y_{it}=j|\mathbf{z}^{+}_{it},r_{it}=1)\end{cases}$$

$$
= \begin{cases} I_{j\leq 0}F(\alpha_1 - \mathbf{x}_{it}'\boldsymbol{\beta})[F^-(\mu_j^- - \mathbf{z}_{it}^{-\prime}\boldsymbol{\gamma}^-) - F^-(\mu_{j-1}^- - \mathbf{z}_{it}^{-\prime}\boldsymbol{\gamma}^-)] \\ +I_{j=0}[F(\alpha_2 - \mathbf{x}_{it}'\boldsymbol{\beta}) - F(\alpha_1 - \mathbf{x}_{it}'\boldsymbol{\beta})] \\ +I_{j\geq 0}[1 - F(\alpha_2 - \mathbf{x}_{it}'\boldsymbol{\beta})][F^+(\mu_j^+ - \mathbf{z}_{it}^{+\prime}\boldsymbol{\gamma}^+) - F^+(\mu_{j-1}^+ - \mathbf{z}_{it}^{+\prime}\boldsymbol{\gamma}^+)], \end{cases} \tag{6}
$$

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

I assume that F, F^- and $F^+$are standard normal. The model is not identified without some extra (arbitrary) assumptions. I also assume that the intercept components of $\boldsymbol{\beta}$, $\boldsymbol{\gamma}^-$ and $\boldsymbol{\gamma}^+$ are zero. The probabilities in (6), however, are absolutely *estimable* — they are invariant to the identifying assumptions — and can be estimated using a partial (pooled) ML estimator of the vector of the parameters $\boldsymbol{\theta} = (\boldsymbol{\alpha}', \boldsymbol{\beta}', \boldsymbol{\mu}^{-\prime}, \boldsymbol{\gamma}^{-\prime}, \boldsymbol{\mu}^{+\prime}, \boldsymbol{\gamma}^{+\prime})'$ that solves

$$
\max_{\boldsymbol{\theta}\epsilon\boldsymbol{\Theta}} \sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{j=-J}^{J} q_{itj} \ln[\Pr(\Delta y_{it} = j|\mathbf{x}_{it}, \mathbf{z}_{it}^-, \mathbf{z}_{it}^+, \boldsymbol{\theta})], \tag{7}
$$

where q_{itj} is an indicator function such that $q_{itj} = 1$ if $\Delta y_{it} = j$ and $q_{itj} = 0$ otherwise. All parameters in $\boldsymbol{\theta}$ are separately identified (up to scale) via the functional form due to the nonlinearity of the OP equations. In practice, however, the standard normal CDF is often an approximately linear function over an extensive range of its argument. Thus, the simultaneous estimation of three equations may be subject to the collinearity and weak identification problems, if $\mathbf{X}$, $\mathbf{Z}^-$ and $\mathbf{Z}^+$ have all (or too many) variables in common. In this case, exclusion restrictions may be necessary to prevent weak identification. The specifications with the complete overlap of variables in the three latent equations are unlikely to be of empirical interest and supported by the data.

Using the fixed T and $N \to \infty$ asymptotics, the estimator in (7) is consistent and square-root-asymptotically normal without any additional assumptions besides the standard identification, regularity and stationarity assumptions, provided the density of $\Delta y_{it}|\mathbf{x}_{it}, \mathbf{z}_{it}^-, \mathbf{z}_{it}^+$ is correctly specified. In contrast to the full ML estimator, in which the *full* conditional density of $\Delta\mathbf{Y}_i|\mathbf{X}_i, \mathbf{Z}_i^-, \mathbf{Z}_i^+$ is assumed to be correctly specified under the null, the partial ML estimator works even if the error terms ν_{it}, ε_{it}^- and ε_{it}^+ are autocorrelated, and $\mathbf{X}_i$, $\mathbf{Z}_i^-$ and $\mathbf{Z}_i^+$ contain the lags of the covariates and the lagged Δy_{it} (see Wooldridge 2010, pp. 486-490). However, the usual asymptotic standard errors and test statistics obtained from the pooled estimation are only valid under the assumption of no serial correlation among the disturbance terms. Without dynamic completeness, the standard errors must be adjusted for serial dependence, for example, by using a robust to density misspecification sandwich estimator of the asymptotic variance of $\boldsymbol{\theta}$

$$\widehat{Avar}(\widehat{\boldsymbol{\theta}}) = \left(-\sum_{i=1}^{N}\sum_{t=1}^{T}\mathbf{H}_{it}(\widehat{\boldsymbol{\theta}})\right)^{-1}\left(\sum_{i=1}^{N}\left[\sum_{t=1}^{T}\mathbf{s}_{it}(\widehat{\boldsymbol{\theta}})\sum_{t=1}^{T}\mathbf{s}_{it}(\widehat{\boldsymbol{\theta}})'\right]\right)\left(-\sum_{i=1}^{N}\sum_{t=1}^{T}\mathbf{H}_{it}(\widehat{\boldsymbol{\theta}})\right)^{-1}, \tag{8}$$

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. 490-492). The asymptotic standard errors of $\widehat{\boldsymbol{\theta}}$ are the square roots of the diagonal elements of (8).

2.2 The correlated cross-nested ordered probit model

The CNOP model can be extended by relaxing the assumption that the mechanisms generating the regime and amount decisions are independent, i.e. that the error terms $\boldsymbol{\nu}$, $\boldsymbol{\varepsilon}^-$ and $\boldsymbol{\varepsilon}^+$ are uncorrelated. To obtain a correlated version of the model (CNOPc), I 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. This correlation may emerge, for instance, from the common but unobserved covariates.

The probabilities to observe an outcome j for the CNOPc model can be written as

$$\Pr(\Delta y_{it} = j|\mathbf{z}_{it}^-, \mathbf{z}_{it}^+, \mathbf{x}_{it})$$

$$= \begin{cases} I_{j\leq 0}[F_2(\alpha_1 - \mathbf{x}_{it}'\boldsymbol{\beta};\mu_j^- - \mathbf{z}_{it}^{-\prime}\boldsymbol{\gamma}^-;\rho^-) - F_2(\alpha_1 - \mathbf{x}_{it}'\boldsymbol{\beta};\mu_{j-1}^- - \mathbf{z}_{it}^{-\prime}\boldsymbol{\gamma}^-;\rho^-)] \\ +I_{j=0}[F(\alpha_2 - \mathbf{x}_{it}'\boldsymbol{\beta}) - F(\alpha_1 - \mathbf{x}_{it}'\boldsymbol{\beta})] \\ +I_{j\geq 0}[F_2(\mathbf{x}_{it}'\boldsymbol{\beta} - \alpha_2;\mu_j^+ - \mathbf{z}_{it}^{+\prime}\boldsymbol{\gamma}^+;-\rho^+) - F_2(\mathbf{x}_{it}'\boldsymbol{\beta} - \alpha_2;\mu_{j-1}^+ - \mathbf{z}_{it}^{+\prime}\boldsymbol{\gamma}^+;-\rho^+)], \end{cases} \tag{9}$$

where $F_2(\phi_1;\phi_2;\xi)$ is the CDF of the standardized bivariate normal distribution of the two random variables ϕ_1 and ϕ_2 with the correlation coefficient ξ. To estimate the CNOPc model by the partial ML, we have to solve (7) by replacing the probabilities in (6) with the probabilities in (9) and redefining the parameter vector $\boldsymbol{\theta}$ as $\boldsymbol{\theta} = (\boldsymbol{\alpha}', \boldsymbol{\beta}', \boldsymbol{\mu}^{-\prime}, \boldsymbol{\gamma}^{-\prime}, \boldsymbol{\mu}^{+\prime}, \boldsymbol{\gamma}^{+\prime}, \rho^{-}, \rho^{+})'$.

2.3 Partial effects

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

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

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

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

$$\begin{aligned}
&\mathbf{X}^{*} = (\mathbf{W}, \mathbf{P}, \mathbf{M}, \widetilde{\mathbf{X}}, \mathbf{V}, \widetilde{\mathbf{Z}}^{+}, \widetilde{\mathbf{Z}}^{-}),\ \boldsymbol{\beta}^{*} = (\boldsymbol{\beta}_w', \boldsymbol{\beta}_p', \boldsymbol{\beta}_m', \widetilde{\boldsymbol{\beta}}', \mathbf{0}', \mathbf{0}', \mathbf{0}')',\\
&\gamma^{+*} = ({\gamma^{+}}'_w, {\gamma^{+}}'_p, \mathbf{0}', \mathbf{0}', {\gamma^{+}}'_v, \widetilde{\boldsymbol{\gamma}^{+}}', \mathbf{0}')',\ \gamma^{-*} = ({\gamma^{-}}'_w, \mathbf{0}', {\gamma^{-}}'_m, \mathbf{0}', {\gamma^{-}}'_v, \mathbf{0}', \widetilde{\boldsymbol{\gamma}^{-}}')'.
\end{aligned}$$

The PE of the row vector $\mathbf{x}_{it}^{*}$ on the full probabilities in (9) can be computed for the CNOPc model as

$$\underset{\Pr(\Delta y_{it}=j)}{\mathbf{PE}} =$$

$$
\left\{
\begin{array}{l}
-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}_{it}^{+\prime}\boldsymbol{\gamma}^{+})}{\sqrt{1-(\rho^{+})^2}}\right)f(\mu_{j-1}^{+}\right.\right. \\
\left.-\mathbf{z}_{it}^{+\prime}\boldsymbol{\gamma}^{+})-F\left(\frac{\mathbf{x}_{it}'\boldsymbol{\beta}-\alpha_2+\rho^{+}(\mu_{j}^{+}-\mathbf{z}_{it}^{+\prime}\boldsymbol{\gamma}^{+})}{\sqrt{1-(\rho^{+})^2}}\right)f(\mu_{j}^{+}-\mathbf{z}_{it}^{+\prime}\boldsymbol{\gamma}^{+})\right]\boldsymbol{\gamma}^{+*} \\
\left.+\left[F\left(\frac{\mu_{j}^{+}-\mathbf{z}_{it}^{+\prime}\boldsymbol{\gamma}^{+}+\rho^{+}(\mathbf{x}_{it}'\boldsymbol{\beta}-\alpha_2)}{\sqrt{1-(\rho^{+})^2}}\right)-F\left(\frac{\mu_{j-1}^{+}-\mathbf{z}_{it}^{+\prime}\boldsymbol{\gamma}^{+}+\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}_{it}^{-\prime}\boldsymbol{\gamma}^{-})}{\sqrt{1-(\rho^{-})^2}}\right)f(\mu_{j-1}^{-}-\mathbf{z}_{it}^{-\prime}\boldsymbol{\gamma}^{-})\right.\right. \\
\left.-F\left(\frac{\alpha_1-\mathbf{x}_{it}'\boldsymbol{\beta}-\rho^{-}(\mu_{j}^{-}-\mathbf{z}_{it}^{-\prime}\boldsymbol{\gamma}^{-})}{\sqrt{1-(\rho^{-})^2}}\right)f(\mu_{j}^{-}-\mathbf{z}_{it}^{-\prime}\boldsymbol{\gamma}^{-})\right]\boldsymbol{\gamma}^{-*} \\
\left.-\left[F\left(\frac{\mu_{j}^{-}-\mathbf{z}_{it}^{-\prime}\boldsymbol{\gamma}^{-}-\rho^{-}(\alpha_1-\mathbf{x}_{it}'\boldsymbol{\beta})}{\sqrt{1-(\rho^{-})^2}}\right)-F\left(\frac{\mu_{j-1}^{-}-\mathbf{z}_{it}^{-\prime}\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{array}
\right.
$$

where f is the probability density function of the standard normal distribution F. The PE for the CNOP model are obtained by setting $\rho^{-}=\rho^{+}=0$. The asymptotic standard errors of the PE are computed using the Delta method as the square roots of the diagonal elements of

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

2.4 Comparison of competing models

The choice of the formal statistical test to compare the performance of the competing models is dependent on whether they are nested in each other. 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 (given that each latent equation of the latter contains all covariates in the corresponding equations of the former). The latter becomes a NOP model if all coefficients on the extra CNOP variables (if any) are fixed to zero and if $\mu_{-1}^{-}\rightarrow\infty$ and $\mu_{0}^{+}\rightarrow-\infty$; therefore, $\Pr(y_{it}^{+}=0|\mathbf{z}_{it}^{+},r_{it}=1)\rightarrow 0$ and $\Pr(y_{it}^{-}=0|\mathbf{z}_{it}^{-},r_{it}=-1)\rightarrow 0$, which can be computationally implemented by letting μ_{-1}^{-} and μ_{0}^{+} be equal to the largest and smallest numbers, respectively, that are available for the estimation software. Thus, testing the NOP versus NOPc, the NOP versus CNOP, the NOP versus CNOPc, the NOPc versus CNOPc, and the CNOP versus CNOPc model can be performed with a test for nested hypotheses, such as the likelihood ratio (LR) test.

In general, the OP and MIOP models are not nested in the NOP, NOPc, CNOP or CNOPc models and vice versa. However, these models are not strictly non-nested.

All six models overlap under certain parameter restrictions if their slope coefficients are restricted to zero and only the thresholds are estimated. Therefore, testing the OP versus any of the two-level models, the MIOP versus the NOP, NOPc, CNOP and CNOPc models and 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 (Vuong 1989). It utilizes the statistical significance between the differences in the log likelihoods. Online Appendix C provides the details about the Vuong test and computation of the informational criteria and noise-to-signal ratios.

An interesting special case when the CNOP and CNOPc models nest the MIOP model occurs under certain parameter restrictions (see online Appendix B for the details) provided (i) the dependent variable only has three outcome categories, (ii) $\mathbf{Z}^{+}$ and $\mathbf{Z}^{-}$ contain all covariates in the MIOP participation equation, and (iii) $\mathbf{X}$ includes all covariates in the MIOP amount equation. In this case, a test of the CNOP versus MIOP model can be performed using the LR test, which can be interpreted as a misspecification test for the latter.

The MIOP reduces to the OP model if (i) the amount equation of the former contains all covariates in the latter, (ii) all coefficients in the participation equation of the former are fixed to zero, and (iii) the threshold parameter in the participation equation is infinitely small to ensure that all observations always occur in the "change" regime. In this special case, a test of the MIOP versus OP model can also be performed using the LR test.

3 Finite sample performance

I conducted the extensive Monte Carlo experiments to illustrate and compare the finite sample performance of the ML estimators in the competing models, namely, to assess the bias and uncertainty of the estimates (and their asymptotic standard errors), the performance of the LR and Vuong tests and the model selection criteria, and the effects of the exclusion restrictions. The details of Monte Carlo design and the results of these simulations are discussed in online Appendix D. To save the space here, I only provide a brief summary of Monte Carlo design and main findings.

Five different DGP were simulated: OP, NOP, NOPc, CNOP, and CNOPc. For each DGP, 3000 repeated samples with 250, 500 and 1000 observations were generated. Under each DGP and for each sample size, several competing models were estimated, always including the OP and NOP models as the benchmarks. Simulations demonstrate that the PE estimates from the OP and NOP models are biased when the underlying DGP is characterized by the three types of zeros, and that the CNOP and CNOPc estimates systematically provide superior probability coverage as well as less bias (see Table 1). The CNOP and CNOPc models under the true OP DGP perform much better than the OP model under the CNOP and CNOPc DGP; as the sample size increases, the relative performance of the CNOP and CNOPc models under the OP DGP improves, whereas the OP and NOP estimates under the CNOP and CNOPc DGP remain biased. I found that it requires two-three times more observations for the three-part models to achieve the same accuracy for the estimated parameters as the OP model (if all models are correctly specified). As long as the number of observations per parameter exceeds 25, the asymptotic distribution is a reasonable approximation of the finite sample distribution; in the smaller samples, the distribution of the standard error estimates is skewed to the right.

Table 1. Selected Monte Carlo results: the CNOP and NOP estimates are consistent under the OP DGP, whereas the OP and NOP estimates remain biased under the CNOP DGP as the sample size increases

Sample size	True DGP:	OP			CNOP			
	Estimated model:	OP	NOP	CNOP	OP	NOP	CNOP	CNOPc
250		42	25	21	36	36	28	23
500	Obs/par	83	50	42	71	71	56	45
1000		167	100	83	143	143	111	91
250		0.25	0.45	1.48	34.63	32.81	0.62	0.82
500	Bias	0.22	0.31	0.99	34.75	32.93	0.25	0.40
1000		0.09	0.20	0.78	34.50	32.89	0.16	0.15
250		2.06	2.95	3.71	4.86	4.44	1.96	2.34
500	RMSE	1.43	2.04	2.48	4.69	4.34	1.34	1.62
1000		1.01	1.44	1.73	4.59	4.27	0.96	1.11
250		93.2	92.0	90.4	36.0	45.9	91.0	90.3
500	CP, %	94.2	93.4	92.2	20.5	35.3	93.0	92.4
1000		94.6	94.0	93.0	13.2	27.3	94.1	93.7

Notes: No overlap scenario (each covariate belongs only to one equation). Obs/par is the number of observations per parameter. Bias is the difference between the estimated and true values of the PE, averaged over all replications and multiplied by 100. RMSE is the root mean square error of the estimated PE relative to their true values, averaged over all replications and multiplied by 100. CP is the empirical coverage probability, computed as the percentage of times the estimated asymptotic 95% confidence intervals cover the true values of the PE.

In addition, to assess the effect of exclusion restrictions, three different scenarios of the overlap among the covariates in the specifications of the three latent equations in (1) and (4) 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, not surprisingly, the greater is the number of exclusion restrictions, the more accurate are the estimates: in the case of the substantial overlap among the covariates in the three latent equations, the asymptotic estimator can experience problems with the convergence and the invertibility of the Hessian if the sample size is small (fewer than 35 observations per parameter).

4 An application to the policy interest rate

> "... monetary economists should embark on a program of continuous improvement and enhanced precision of the Fed's monetary rule." - W. Poole (2006)

I apply the OP, MIOP, CNOP and CNOPc models to explain the systematic components of the NBP policy rate decisions, and employ a novel panel of the individual policy preferences of the MPC members and the vintages of the real-time economic data available to the public one day prior to each policy-setting meeting during the period from February 1998 to April 2014.

4.1 Data

After the adoption of direct inflation targeting in 1998, the NBP policy rate — the reference rate[4] — may be undoubtedly treated as a principal instrument of Polish monetary policy. The reference rate is administratively set by the MPC, which consists of ten members who make policy rate decisions by formal voting once per month (since 2010, eleven times per year). The Council members are appointed for a non-renewable term of six years but the Chair may serve for two consecutive terms.[5]

The MPC always moves policy rates by discrete adjustments — multiples of 25 basis points (bp), i.e. a quarter of one percent. At a policy meeting, each MPC member can express his preferred adjustment to the rate and make a proposition to be

[4] The rate on the 28-day (from 1998 to 2003), 14-day (from 2003 to 2005) and 7-day (since 2005 to present) NBP money market bills.

[5] The first and second terms lasted, respectively, from February 1998 through January 2004 and from February 2004 through January 2010. The third term began in February 2010.

voted on. If no proposition is made, no voting occurs and the rate remains unchanged. Otherwise, the Chair selects the largest proposed move and the members vote on it; the Chair has the casting vote if there is no majority. If the first voted proposition commands a majority, the other propositions are not voted on; otherwise, the members vote on the alternative proposal (if any). The available voting records report all proposed motions (but do not report who made the motions) and all individual votes, for or against (but do not explicitly indicate the desired rate change of the members who voted against). This information enables us to identify the direction of dissent, i.e. whether a dissenting policymaker prefers the higher or lower rate. It also enables us to identify the desired amount of change (if any), if only one or no proposition is made. Fortunately, only a few cases (less than 0.8% of all observations) exist in which more than one motion is proposed and the largest move is passed. In these cases, I conjecture which dissenting members prefer the alternative smaller move(s) and which members prefer a status quo by considering their previous rate-setting behaviors.

From 1998 to mid-2002, the rate of inflation in Poland decreased from 14% to 2%; since this period, it has fluctuated in range of 0.3%–4.8%. The current NBP inflation target of 2.5%±1% was established in January 2004 and was not changed after this date. Since May 2002, the policy rate changed 36 times by 25 *bp*, 12 times by 50 *bp*, and two times by 75 *bp*. Previously, it was more volatile in response to more volatile inflation; the rate changed two times by 50 *bp*, eight times by 100 *bp*, seven times by 150 *bp*, and three times by 250 *bp*. To provide a reliable statistical inference with these data limitations, the individual policy preferences (reported in Table 19 in online Appendix E) are consolidated into three categories: increase, no change and decrease. The sample contains the desired policy actions expressed during the first round of voting (if any) by 31 policymakers at 190 MPC meetings. The two MPC members (Wiesiołek and Osiatyński) are excluded from the sample because they were involved in policy decisions on two and six occasions, respectively. Among the 1719 observations employed in the estimations, the policymakers preferred to leave the rate unchanged 1125 times (65%), to increase the rate 253 times (15%) and to reduce it 341 times (20%). Table 2 provides the definitions and sources of all variables. The sample descriptive statistics are summarized in Table 20 in online Appendix E. All employed macroeconomic variables are stationary at the 0.01 significance level according to the augmented Dickey-Fuller unit root test, as shown in Table 21 in online Appendix E.

Table 2. Definitions of the variables

Mnemonics	Variable description (source of data)
Δy_i	Dependent variable - change to the NBP reference rate, preferred by MPC member i and consolidated into three categories: 1 if an increase, 0 if no change, -1 if a decrease (NBP and AC).
Δcpi	Monthly change to the consumer price index (CPI), annual rate in percent (GUS).
Δcpi^{tar}	Δcpi if CPI is above the NBP inflation target, and zero otherwise (GUS and NBP).
$\Delta ecbr$	Change to the European Central Bank policy rate (since 2/1999, in 1998 - to the Bundesbank policy rate, and zero in 1/1999), announced at the last policy meeting, annualized percent (ECB and Bundesbank).
$\Delta rate_i$	Change to the NBP reference rate, preferred by MPC member i, annualized percent (NBP and AC).
$spread$	Difference between the 1-year and 1-week Poland interbank offer rate, 5-business-day moving average, annualized percent (Thompson Reuters).
$bias$	Indicator of "policy bias" or "balance of risks" statements of the MPC: -1 if "easing", 0 if "neutral", and 1 if "restrictive" (NBP and AC).
$dissent_i$	Measure of dissent of MPC member i at a meeting, equal to -1/0/1 if member i prefers the lower/same/higher interest rate than the rate set by the MPC (NBP and AC).
I^{2002}	One since 4/2002, and zero otherwise.
I^{2010}	One since 2/2010, and zero otherwise.

Notes: GUS - Central Statistical Office of Poland, AC - author's calculations.

4.2 Model specification

Given the NBP strategy of direct inflation targeting, the policy regime decision in the CNOP model is assumed to be driven by a direct reaction to the changes in the economic conditions controlled by: (i) Δcpi_t — the recent monthly change to the current rate of inflation; (ii) Δcpi_t^{tar}, which is equivalent to Δcpi_t if the inflation is above the target, and zero otherwise (to allow for an asymmetric reaction to inflation changes depending on whether the inflation rate is above or below the target); (iii) $\Delta ecbr_t$ — the change to the ECB policy rate made at the last policy meeting (as a proxy for the recent economic trends in the European Union); (iv-v) $\Delta cpi_t^{tar} I_t^{2010}$ and $\Delta ecbr_t I_t^{2010}$, where I_t^{2010} is an indicator variable, which is which is one since February 2010, and zero otherwise (to allow for a different policy reaction in the post-crisis period during the third MPC term); (vi) $spread_t$ — the spread between the long- and short-term market interest rates (as a low-dimension market-based aggregator of publicly available information on inflationary expectations that are not reflected in the current inflation rate); (vii) $\Delta rate_{i,t-1}$ — the original (unconsolidated) change to the policy rate proposed by the MPC member i at the previous meeting (sequential decisions are not independent — the recent policy choice affects subsequent actions); and (viii) $bias_{t-1}$ — an indicator of the "policy bias" or "balance of risks" statements of the MPC at the previous meeting (to address the policymakers' concerns about the competence and credibility of central bank communication). The expected sign of the coefficients on these variables is positive — the larger is the value of a covariate, the larger is the probability of a tight policy stance and the smaller is the probability of a loose stance.

The amount decisions, which fine-tune and smooth the rate, are conditional on the tight or loose policy stance and controlled by (i) $\Delta rate_{i,t-1}$ (the larger is the hike/cut at the previous meeting, the lower is the probability of the second hike/cut in a row); (ii) $bias_{t-1}$ (the tightening/easing bias is expected to increase/decrease the probability of a higher rate); (iii) $spread_t$ (the rate hike is much more likely if the 12-month interbank rate is above the 1-week rate, rather than vice versa); (iv) the indicator variable I_t^{2002}, which is one since April 2002, and zero otherwise (to account for higher levels and stronger moves in the inflation and the policy rate prior to April 2002); and I_t^{2010} (to allow for a different policy reaction during the third MPC term). The expected sign of the coefficients is positive for $spread_t$ and $bias_{t-1}$, negative for $\Delta rate_{i,t-1}$, and should be opposite in the tight and loose regimes for I_t^{2002} and I_t^{2010}: a positive sign in the tight stance but a negative sign in the loose stance enable rate moves to be triggered by smaller changes to the explanatory variables after April 2002 and February 2010, respectively.

There are no inter-individual differences in the values of the macroeconomic explanatory variables. To better account for the individual heterogeneity of policy preferences (not fully controlled by $\Delta rate_{i,t-1}$), I augment this specification by including the individual fixed effects (FE). For parsimony reasons, I only allow for variation in the intercepts — thirty individual dummy variables are included in each latent equation.[6] Slope heterogeneity is not a concern because our interest is the estimation of the average effects of the explanatory variables, not the individual policy reactions. Under the assumption that the slope coefficients randomly differ across the individuals, the pooled ML estimator yields consistent estimates of these aggregate effects while simultaneously providing a greater statistical power and a more reliable inference. We should not expect any significant fixed T asymptotic bias of our estimator; with our temporal size (T_i is 55 on average) we are in the realm of a time-series analysis.[7]

[6] A dummy for Gronkiewicz-Waltz, who was the first MPC Chair (1998–2000) and the only MPC member in the sample who never dissented, and I_t^{2010} are omitted to avoid the dummy variable trap.

[7] Using Monte Carlo simulations, Greene (2004) investigated bias in the discrete-choice panel models, including the OP model. As T increases from 2 to 20, the 160% bias reduces to 6%.

The fixed effects are more appropriate than the random effects because we do not have a sample of individuals who were randomly obtained from a large population but instead possess a full set of the MPC members. However, the FE specification with its 110 parameters is likely subject to a weak identification problem: fewer than 16 observations per parameter are contained in the sample, and the sets of the covariates in the three latent equations overlap substantially. To prevent an overparameterization and obtain more reliable estimates, I also estimated a more parsimonious specification. I constructed the individual-specific variable $dissent_{i,t}$, which indicates a direction of member i's dissent at a meeting t: it is equal to 1/0/-1, if the member prefers the higher/same/lower rate than the MPC. The lags of $dissent_{i,t}$ reflect the dynamics of the deviation of member i's desired rate from the rate set by the MPC at the previous meetings. I included three lags in the regime equation, two lags in the amount equation under the loose regime, and three lags under the tight regime (with the expected positive sign of all coefficients; if a member preferred a higher/lower policy rate at the previous meeting, he is likely to be more hawkish/dovish at the subsequent meetings).[8]

4.3 Estimation results

The three lags of $dissent_{i,t}$ adequately capture the heterogeneity of policy preferences. The alternative specification has a slightly lower log likelihood than the FE specification (-629.2 vs -626.2), but far fewer parameters (30 vs 110), and is strongly preferred by the information criteria (the AIC is 1318 vs 1472, the BIC is 1482 vs 2072). The FE specification is heavily overparameterized: fourty individual dummies have a coefficient that is not statistically significantly different from zero at the 0.05 level, as reported in Table 23 in online Appendix E. Therefore, the specification with the lags of $dissent_{i,t}$ (henceforth the *baseline* specification) was employed in the further analysis. It saves 80 degrees of freedom, has an advantage of a greater statistical power, and can produce more efficient estimates of interest. Importantly, the estimated policy reactions to economic conditions are robust to different ways of accounting for individual heterogeneity: the coefficients of all common variables in the regime equation and of all but three variables in the amount equations are remarkably similar in both specifications.

[8] The third lag of $dissent_{i,t}$ is not included in the amount equation under the loose regime because its coefficient is not statistically significantly different from zero at the 0.22 level (see Table 22 in online Appendix E) and the LR test fails to reject its redundancy (the p-value is 0.31).

Table 3. Modeling changes to policy rate: the coefficients from the CNOP model

Variables	Policy regime equation	Amount equations	
		Loose regime	Tight regime
Δcpi_t	-0.22 (0.24)		
Δcpi_t^{tar}	7.99 (1.11)***		
$\Delta cpi_t^{tar} * I_t^{2010}$	-7.75 (1.16)***		
$\Delta ecbr_t$	11.02 (1.73)***		
$\Delta ecbr_t * I_t^{2010}$	-9.58 (1.78)***		
$\Delta rate_{i,t-1}$	2.44 (0.44)***	-1.08 (0.16)***	-3.10 (0.55)***
$spread_t$	1.93 (0.17)***	0.70 (0.14)***	0.67 (0.18)***
$dissent_{i,t-1}$	0.66 (0.17)***	1.64 (0.32)***	1.62 (0.28)***
$dissent_{i,t-2}$	0.23 (0.17)	0.74 (0.15)***	0.47 (0.23)**
$dissent_{i,t-3}$	0.51 (0.14)***		0.78 (0.28)***
$bias_{t-1}$	0.51 (0.14)***	6.99 (0.56)***	2.02 (0.21)***
I_t^{2002}		-1.37 (0.28)***	0.80 (0.21)***
I_t^{2010}		-6.93 (0.73)***	1.42 (0.54)***
$threshold_1$	-1.27 (0.12)***	-0.79 (0.25)***	2.51 (0.32)***
$threshold_2$	2.50 (0.19)***		

Notes: For the definitions of the variables refer to Table 2. ***/**/* denote the statistical significance at the 1/5/10 percent level. The robust asymptotic standard errors are shown in parentheses.

As shown in Table 3, all coefficients from the baseline specification have an expected sign and are statistically significant at the 0.01 level, with the exception of the coefficient on $dissent_{i,t-2}$ (the p-value is 0.17) and Δcpi_t (the p-value is 0.35) in the regime equation. Only the signs of the coefficients — not their values — are of practical interest. The values are only identified up to scale, whereas the signs unambiguously imply the signs of the PE on the probabilities of a rate hike or cut. Our expectation that the policy reaction to changes in inflation is dependent on whether inflation level is above or below the target is confirmed: the reaction is not statistically significant if the inflation is below the target.

Observing a large fraction of zeros does not always indicate that existing models are unsuitable. We can test which alternative model is favored by real-world data: (i) the standard OP model (including all covariates from the CNOP model; see Table 4); (ii) the two-part MIOP model (in which its dichotomous participation equation includes all covariates in the CNOP dichotomous amount equations and the trichotomous amount equation includes all covariates in the CNOP trichotomous regime equation; see Table 4); (iii) the three-part CNOP model (with the baseline specification); and (iv) the correlated version of the CNOP model (see Table 24 in online Appendix E). The NOP model is not listed because, in the case of the three outcome categories, it reduces to the OP model.

Table 4. Modeling changes to policy rate: the coefficients from the OP and MIOP models

Variables	OP	MIOP	
		Participation equation	Amount equation
Δcpi_t	0.26 (0.11)**		-0.14 (0.18)
Δcpi_t^{tar}	0.97 (0.17)***		2.43 (0.72)***
$\Delta cpi_t^{tar} * I_t^{2010}$	-0.45 (0.28)		-1.72 (0.84)**
$\Delta ecbr_t$	1.80 (0.25)***		4.25 (0.43)***
$\Delta ecbr_t * I_t^{2010}$	-0.08 (0.59)		-2.83 (0.68)***
$\Delta rate_{i,t-1}$	-0.60 (0.08)***	0.11 (0.17)	-0.27 (0.60)
$spread_t$	0.77 (0.06)***	-0.04 (0.08)	1.44 (0.12)***
$dissent_{i,t-1}$	1.16 (0.11)***	-0.28 (0.35)	1.32 (0.19)***
$dissent_{i,t-2}$	0.42 (0.10)***	-0.04 (0.19)	0.47 (0.13)***
$dissent_{i,t-3}$	0.45 (0.09)***	-0.13 (0.19)	0.58 (0.12)***
$bias_{t-1}$	1.20 (0.06)***	-0.02 (0.38)	1.30 (0.09)***
I_t^{2002}	-0.22 (0.11)**	1.55 (0.31)***	
I_t^{2010}	0.18 (0.09)**	0.89 (0.94)	
$threshold_1$	-1.37 (0.09)***	0.01 (0.17)	-1.20 (0.14)***
$threshold_2$	2.37 (0.11)***		2.99 (0.12)***

Notes: For the definitions of the variables refer to Table 2. ***/**/* denote the statistical significance at the 1/5/10 percent level. The robust asymptotic standard errors are shown in parentheses.

Table 5 shows the summary statistics and comparison of the five competing models. The two- and three-equation models demonstrate a significant increase in the likelihood compared to the single-equation OP model. The CNOP and CNOPc models are overwhelmingly superior to the OP and MIOP models according to all information criteria and are favored by the Vuong tests (at the significance level 10^{-20}). The CNOPc model exhibits an insignificant increase in the likelihood compared with the CNOP according to the LR test (the p-value is 0.999). The CNOP model is preferred by all information criteria. I also estimated it with the same set of variables in both amount equations (by including $dissent_{i,t-3}$ under the loose regime) to test whether the rate hikes and cuts are generated by different processes. In our case with only three outcome categories of the dependent variable, the CNOP nests the MIOP model under certain "symmetrical" restrictions on the parameters in the amount equations (see online Appendix B for the discussion). The LR test strongly rejects the symmetrical restrictions and prefers the CNOP model (the p-value is 10^{-37}).

Table 5. Comparison of competing models: the CNOP model is favored by real-world data

Model	OP	MIOP	CNOP	CNOPc
Log likelihood	-813.3	-758.2	-629.2	-629.2
# of parameters	15	22	30	32
AIC	1656.5	1560.4	**1318.3**	1322.3
BIC	1738.3	1680.3	**1481.8**	1496.7
Hit rate	0.77	0.80	0.83	0.83
Vuong test vs OP		-3.97***	-11.02***	-11.02***
Vuong test vs MIOP			-9.38***	-9.37***
LR test vs CNOP				0.002

Notes: ***/**/* denote the statistical significance at the 1/5/10 percent level.

Table 6. Comparison of competing models: the CNOP model has better hit rates and noise-to-signal ratios

Actual outcome	Hit rate			Adjusted noise-to-signal ratio		
	OP	MIOP	CNOP	OP	MIOP	CNOP
Decrease	0.60	0.71	**0.78**	0.12	0.07	**0.06**
No change	0.87	**0.89**	**0.89**	0.48	0.42	**0.31**
Increase	0.56	0.51	**0.64**	**0.06**	0.07	**0.06**

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 in online Appendix C.

The CNOP model also demonstrates a substantial improvement in the percentage of correct predictions (for rate cuts and hikes) and noise-to-signal ratios (for cuts and zeros), as shown in Table 6. The noise-to-signal ratios for hikes and the hit rates for zeros are similar across the three models, although slightly better in the CNOP model. Interestingly, the OP and MOP models predict more zeros (1224 and 1228) than the CNOP model (1171), but they *correctly* predict only 977 and 1004 zeros, respectively, whereas the CNOP model *correctly* predicts 1005 zeros.

To give the MIOP model additional chances, I also estimated it (a) including all CNOP covariates in both parts (the log likelihood is -725.1; see Table 25 in online Appendix E) and (b) taking additionally all covariates in the participation equation by their absolute values to account for the binary (change or no change) nature of the first-stage decision (the log likelihood is -713.6; see Table 26 in online Appendix E). In both cases, the CNOP remains overwhelmingly superior to the MIOP model according to the information criteria[9] and Vuong tests (at the significance level 10^{-9}).

The model comparison relies heavily on statistical criteria. Are we simply fine-tuning the OP and MIOP models or are the resulting improvements economically meaningful? The three models produce a conflicting inference and have incompatible and opposite estimates of the effect of some explanatory variables on choice probabilities. The most striking differences across the models are in the effects of the previous change to the rate $\Delta rate_{i,t-1}$.

We expect a positive coefficient on $\Delta rate_{i,t-1}$ in the OP model. In the case of a rate hike, the probability of a hike/cut at the next meeting should be larger/smaller than for the case of a cut. The coefficient has a negative sign in the OP model, which nonsensically implies that the larger is the proposed hike at the last meeting, the more likely is a cut at the next meeting. The CNOP model assumes that the rate change is the combined result of the two distinct decisions, on which a given variable may have different and even opposite effects. We expect a high level of persistency in the latent policy regime due to the slow cyclical fluctuations of macroeconomic indicators, which exogenously drive the policy stance. Central banks are typically conservative and dislike frequent reversals in the direction of movement in interest rates. Therefore, we expect a positive coefficient on $\Delta rate_{i,t-1}$ in the policy regime equation. Policymakers are cautious and tend to wait and see after they have moved the rate; an adjustment is typically followed by a status quo decision. The CNOP amount decisions are unidirectional, either non-positive or non-negative, if the policy stance is loose or tight, respectively. Thus, we expect a negative coefficient on $\Delta rate_{i,t-1}$ in the amount equations. The coefficient has a positive sign in the regime equation but the negative

[9] The AIC and BIC are 1318 and 1482 for the CNOP model but only 1508 and 1666 for the MIOP model in the case (a) and 1485 and 1643 in the case (b), respectively.

signs in the amount equations, which implies that the larger is the hike at the last meeting, the larger is the probability of a tight regime at the next meeting and, conditional on the tight/loose stance, the smaller is the probability of a hike/cut and the larger is the probability of no change.

Table 7. Comparison of competing models: the CNOP model provides the economically more meaningful estimates of the partial effects of $\Delta rate_{i,t-1}$ on choice probabilities

	OP	MIOP	CNOP
Pr($\Delta y_{i,t}$ = "increase")	-0.003 (0.001)***	-0.001 (0.001)	0.000 (0.000)
Pr($\Delta y_{i,t}$ = "no change")	-0.021 (0.003)***	-0.009 (0.010)	0.084 (0.024)***
Pr($\Delta y_{i,t}$ = "decrease")	0.025 (0.004)***	0.009 (0.010)	-0.084 (0.024)***

Notes: ***/**/* denote the statistical significance at the 1/5/10 percent level. The robust asymptotic standard errors are shown in parentheses. The partial effects are computed as a change in the probabilities when $\Delta rate_{i,t-1}$ changes from -25 *bp* to 0 *bp*, the inflation rate is above the target, and the other variables are fixed at their sample median values.

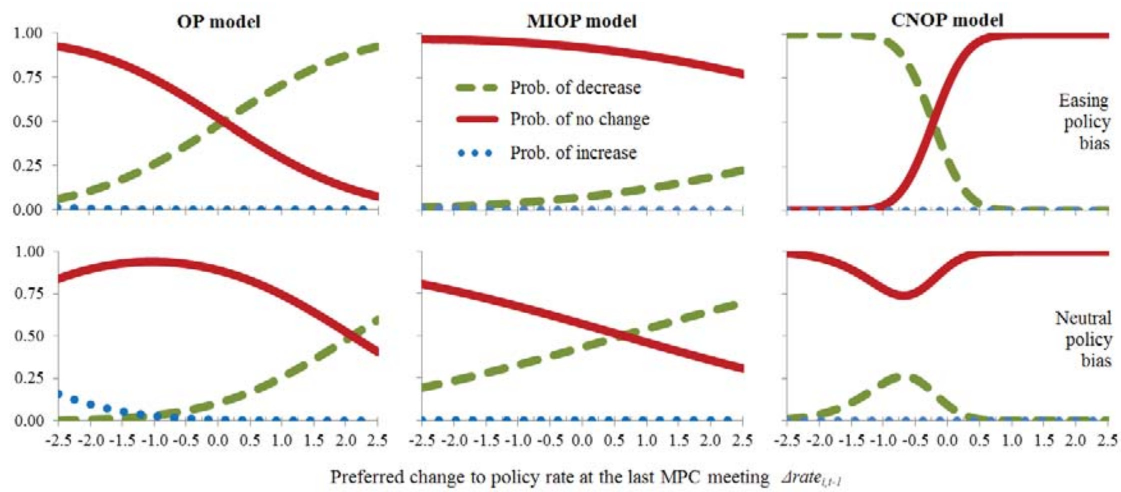
The differences in the PE of $\Delta rate_{i,t-1}$ on the choice probabilities obtained across the three models are intriguing: the CNOP model has the opposite signs of the PE compared with the OP and MIOP models, as shown in Table 7.[10] According to the CNOP model, if $\Delta rate_{i,t-1}$ changes from -25 to 0 *bp*, holding all other variables fixed, the probability of a rate cut decreases by 0.084, the probability of a hike increases insignificantly, and the probability of no change increases by 0.084. By contrast, the OP and MIOP models produce a conflicting and misleading inference: the probability of a cut increases by 0.025 and 0.009, the probability of a hike decreases by 0.003 and 0.001, and the probability of no change decreases by 0.021 and 0.009, respectively.

The impact of $\Delta rate_{i,t-1}$ on choice probabilities from the three models is also graphically compared in Figure 4. The predicted probabilities exhibit sharp contradictions. For example, if the policy bias is easing and $\Delta rate_{i,t-1}$ increases, the probability of no change increases in the CNOP model but decreases in the OP and MIOP models. Similarly, the three models make a conflicting inference regarding the probability of a rate reduction. The OP and MIOP models fail to provide an accurate assessment of the relationship between the explanatory variables and outcome

[10] The partial effects of all explanatory variables are reported in Table 27 in online Appendix E.

probabilities and produce an absurd inference. The capability of the CNOP model to disentangle the opposite directions of the effect of $\Delta rate_{i,t-1}$ on the regime and amount decisions produces an economically more reasonable inference.

Figure 4. Comparison of competing models: the CNOP model provides the more reasonable estimates of choice probabilities



Notes: The probabilities are computed for the range of the preferred change to the rate $\Delta rate_{i,t-1}$ and two values of $bias_{t-1}$ (easing and neutral) at the last MPC meeting, if the inflation rate is above the target and the other variables are fixed at their sample median values.

Figure 5 shows the estimated probabilities of latent policy regimes, which are averaged for each meeting across all MPC members. The probability profiles differ considerably in the periods of policy easing, maintaining and tightening, as demostrated in Figure 6. Averaged over all meetings, the probabilities of the loose, neutral and tight policy stances are 0.33, 0.41 and 0.26, respectively, whereas the observed frequencies of the cuts, no-change decisions and hikes are 0.20, 0.65 and 0.15, respectively. All the zeros are not generated by a neutral policy stance.

Figure 5. Actual policy decisions and estimated probabilities of latent policy regimes

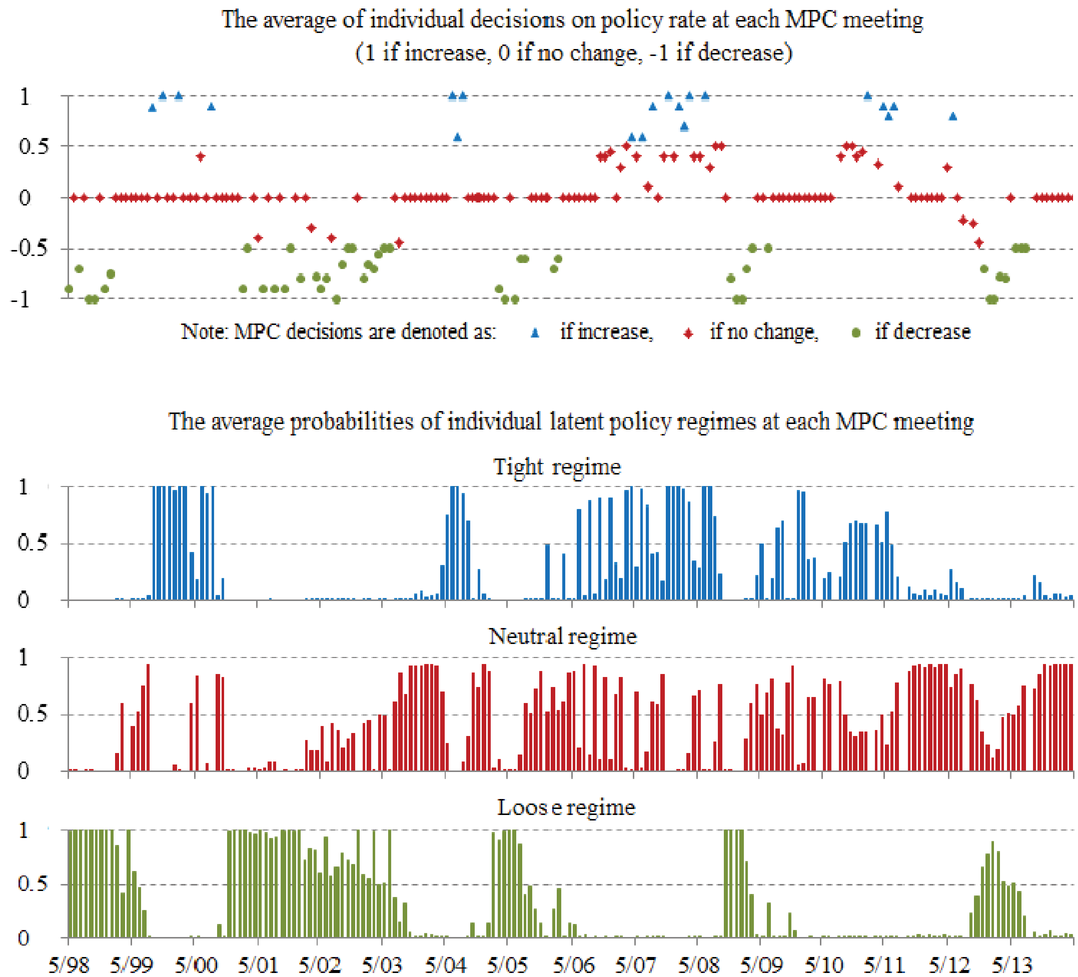
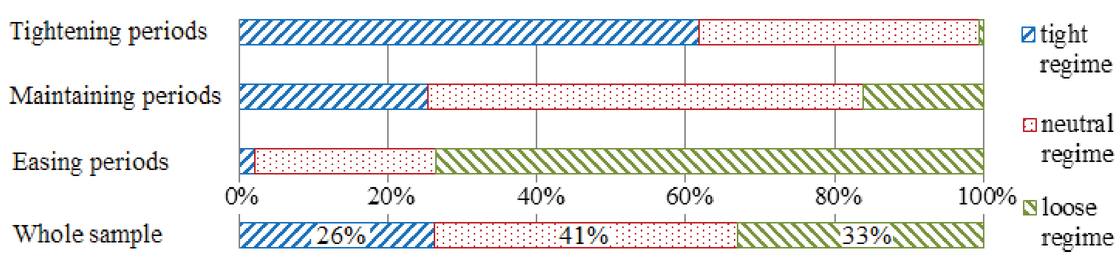


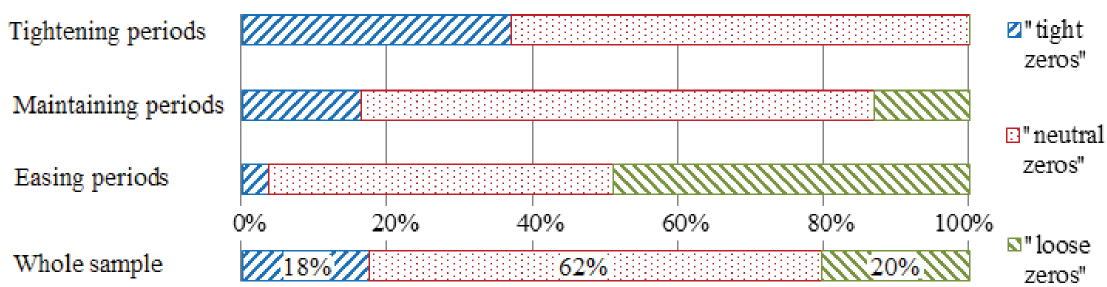
Figure 6. Probabilities of latent regimes in different policy periods remarkably differ



Notes: The estimates are obtained from the baseline CNOP model. For the definitions of the policy periods, refer to Figure 1.

These findings are refined in Figure 7, which reports the decomposition of unconditional probability of no change into three conditional parts, $\Pr(\Delta y_{i,t} = 0|r_{i,t} = -1)$, $\Pr(\Delta y_{i,t} = 0|r_{i,t} = 0)$ and $\Pr(\Delta y_{i,t} = 0|r_{i,t} = 1)$, which correspond to the loose, neutral and tight zeros. This decomposition substantially varies and, as we hypothesized, the identified three types of zeros are unproportionally distributed across different policy periods. During policy easing and tightening, the fractions of neutral zeros are 0.47 and 0.63, respectively. The fraction of neutral zeros is only 0.70 even among the zeros that are clustered between the rate reversals during policy maintaining. For the entire sample, the portions of the loose, neutral and tight zeros are 0.20, 0.62 and 0.18, respectively.[11] According to the CNOP model, less than two-thirds of the status quo decisions appeared to be generated by a neutral policy reaction to the economic conditions. The policy-making process in the NBP appears to be inertial by choice: 40% and 44% of all outcomes in the loose and tight regimes, respectively, are the zeros.

Figure 7. The decomposition of $\Pr(\Delta y_{i,t}=0)$ into the probabilities conditional on the loose, neutral or tight regimes remarkably differs in different policy periods



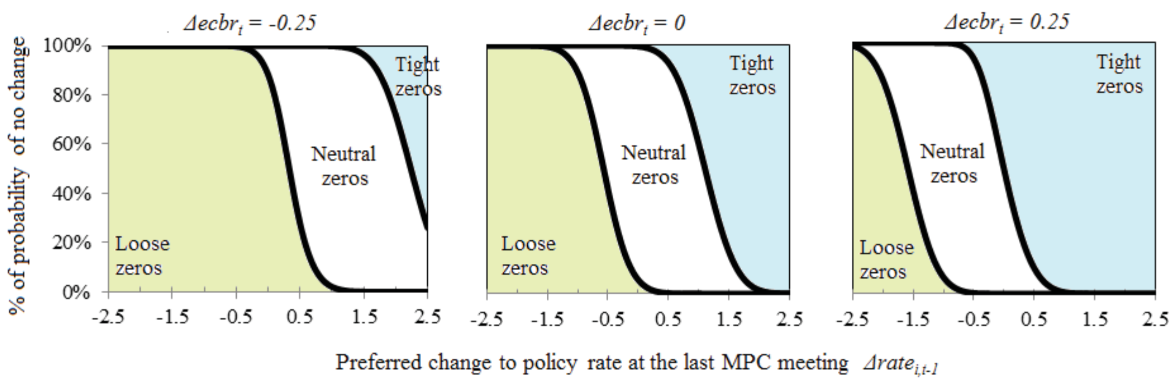
Notes: The estimates are obtained from the baseline CNOP model. For the definitions of the policy periods, refer to Figure 1.

The PE on the unconditional probability of no change can also be decomposed into three components. For example, the 0.084 (with the 0.024 robust standard error) PE of $\Delta rate_{i,t-1}$ on $\Pr(\Delta y_{i,t} = 0)$ is the combined result of the -0.098 (0.025), 0.178 (0.040) and 0.003 (0.002) effects conditional on the loose, neutral and tight policy regimes, respectively (see Table 29 in online Appendix E). To graphically illustrate how the decomposition of $\Pr(\Delta y_{i,t} = 0)$ is dependent on data, it can be plotted as a function of two explanatory variables, holding all others fixed. Fot example, Figure

[11] The average predicted probability of a no-change decision during the observed no-change outcomes is decomposed similar as 0.19/0.64/0.17 (see Table 28 in online Appendix E).

8 shows that if the individual policy choice at the last MPC meeting was to leave the rate unchanged, the inflation is above the target, the last ECB policy decision was a 25-bp cut, and the other variables are fixed at their sample median values, then $\Pr(\Delta y_{i,t} = 0)$ is composed, on average, of 88% of loose and 12% of neutral zeros. However, if the ECB left the policy rate unchanged, then it is composed of 6% of loose and 94% of neutral zeros. If the ECB decision was a 25-bp hike, then $\Pr(\Delta y_{it} = 0)$ is composed of 49% of neutral and 51% of tight zeros.

Figure 8. The decomposition of $\Pr(\Delta \mathrm{y}_{i,t}{=}0)$ into three components conditional on the loose, neutral and tight policy regimes as a function of policy rate choice at the last MPC meeting $\Delta rate_{i,t-1}$ and recent ECB policy decision $\Delta ecbr_t$



Notes: The probabilities are computed for the range of $\Delta rate_{i,t-1}$ and three values of $ecbr_t$, if the inflation rate is above the target, holding all other variables at their sample median values. The estimates are obtained from the baseline CNOP model. For the definitions of the variables, refer to Table 2.

4.4 Sensitivity analysis

The sensitivity of the obtained empirical results is examined along several dimensions. All key empirical findings — the parameter estimates from the CNOP model (see Table 30 in online Appendix F), the comparison of the PE estimates from the OP, MIOP and CNOP models (see Table 31 in online Appendix F) and the model performance comparison (see Table 32 in online Appendix F) — are highly robust with respect to the following modifications of the baseline specification: (a) alternative definitions of the individual policy rate preferences (the individual preferences expressed in the last voting round versus the first round (if any) as in the baseline specification; this modification affected the definitions of the dependent variable $\Delta y_{i,t}$ and several explanatory variables: $\Delta rate_{i,t-1}$, $dissent_{i,t-1}$, $dissent_{i,t-2}$ and $dissent_{i,t-3}$);

(b) alternative definition of the previous policy choice (the rate change set by the MPC versus the individual preferred rate change as in the baseline specification); (c) alternative indicator of the policy bias (the two separate 0/1 indicators for the easing and restrictive biases versus one -1/0/1 indicator as in the baseline specification); (d) different measures of inflation (the measures of the expected and core inflation versus the current headline inflation as in the baseline specification); (e) different maturities of short-term rates in the spread between the one-year and short-term market interest rates (the one-month and two-week Poland interbank offer rates and the NBP reference rate versus the one-week rate as in the baseline specification); (f) inclusion of the different measures of the economic activity as a potentially influential omitted variable (the various monthly indicators from the *Business Tendency Survey* of the NBP); (g) use of the different subsamples (elimination of the third term of the MPC, i.e. the last 51 meetings after January 2010; elimination of high-inflation period prior to April 2002, i.e. the first 49 meetings of the first term of the MPC).

5 Concluding remarks

> "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 (1999)

Ordinal dependent variables with negative, zero and positive values are often characterized by abundant observations in the middle neutral category. Observing a large fraction of zeros does not necessarily imply that conventional discrete-choice models are not suitable. If zeros are generated by different groups of population or by separate decision-making processes and positive and negative outcomes are driven by the distinct forces, treating all observations as originating from the same data-generating process and applying a standard single-equation model would be a misspecification. The standard models are hindered by overfitting of the most popular choice; in addition, a failure of the data homogeneity assumption and the way, in which zero values are treated, usually result in the biased and inefficient estimates of the choice probabilities and the marginal effects of the explanatory variables on these probabilities.

To address these issues, this paper develops a new mixture model with overlapping latent regimes by combining three ordered probit equations. In the empirical application to policy interest rate, the new model not only demonstrates that the presence of heterogeneity in the data generating process is convincing and dominates the conventional models but also provides a qualitatively different and economically more reasonable inference.

The proposed cross-nested ordered probit model can be applied to a variety of datasets (changes to consumption, prices, or rankings) and survey responses (when respondents are asked to indicate a negative, neutral or positive attitude). The GAUSS codes and replication files are available upon request.

Acknowledgements

I gratefully acknowledge research support by grant #R10-0221 from the Global Development Network, a grant from the National Bank of Poland, and the Zvi Griliches Excellence Award from the EERC. I am also grateful to Jérôme Adda, Michael Beenstock, James D. Hamilton, Mark N. Harris, Peter R. Hansen, Helmut Lütkepohl, Massimiliano Marcellino, Chiara Monfardini, Simon van Norden, Dobromil Serwa, Grzegorz Szafrański, and Francis Vella, as well as the participants of the MOOD conference in Rome, the ESEM conference in Málaga, the IAAE conference in London, the ESAM conference in Hobart, and the seminars at the National Bank of Poland, the University of Bologna, the New Economic School and the Higher School of Economics in Moscow, and the University of Melbourne for the useful discussions and comments on the previous versions of this paper.

References

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

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

Coibion, O., Gorodnichenko, Y., 2012. Why are target interest rate changes so persistent? American Economic Journal – Macroeconomics 4, 126-162.

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

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

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

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

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

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

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

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

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

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

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

Kelley, M. E., Anderson, S. J., 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., 1998. The accidental theorist and other dispatches from the dismal science. W. W. Norton & Company, Inc.

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

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

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

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

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

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

Rudebusch, G. D., 2002. Term structure evidence on interest rate smoothing and monetary policy inertia. Journal of Monetary Economics 49, 1161-1187.

Rudebusch, G. D., 2006. Monetary policy inertia: fact or fiction? International Journal of Central Banking 2(4), 85-135.

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

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

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

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

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

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

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

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