Modelling category inflation with multiple inflation processes*

Sarah Brown,† Mark N. Harris‡ and Christopher Spencer¶

†*Department of Economics, University of Sheffield, 9 Mappin Street, Sheffield, S1 4DT, UK. Tel.: +44 (0)114 222 3404. (sarah.brown@sheffield.ac.uk)*
‡*School of Economics and Finance, Curtin Business School, Curtin University, Perth WA 6845, Australia. Tel.: +61 (0)8 9266 9692. (mark.harris@curtin.edu.au)*
¶*School of Business and Economics, Ashby Road, Loughborough University, Loughborough, Leicestershire, LE11 3TU, UK. Tel.: +44 (0)1509 222715 (c.g.spencer@lboro.ac.uk)*

Abstract

Zero-inflated ordered probit ($ZIOP$) and middle-inflated ordered probit ($MIOP$) models are finding increasing favour in the discrete choice literature. Both models consist of a mixture of binary and single ordered probit equations, the combination of which accounts for an "excessive" build-up of observations in a given choice category. We propose generalisations to these models – which collapse to their $ZIOP/MIOP$ counterparts under a set of simple parameter restrictions – with respect to the inflation process. The appropriateness and implications of our generalisations are demonstrated by using two key empirical applications from the economics and political science literatures. Likelihood ratio (LR) and Lagrange multiplier (LM) specification tests lead us to support the newly proposed generalised models over the $ZIOP/MIOP$ ones, and suggest a role for these new models in modelling zero- and middle-inflation processes.

Keywords: Discrete ordered data, Lagrange multiplier test, middle-inflation, zero-inflation..

JEL Classification numbers: C12, C35.

* The authors thank Benjamin Bagozzi and Bumba Mukherjee for making their data available on the *Harvard Dataverse*, which can be found at https://dataverse.harvard.edu/dataverse/pan (accessed 25 October 2016).

I Introduction and motivation

Recent advances in discrete choice modelling have witnessed the development of so-called inflated ordered probit models. These draw inspiration from the suite of hurdle and double-hurdle models for continuous and count outcome variables - developed to address an excess of zero observations (Cragg 1971, Mullahey 1986, Lambert 1992, Heilbron 1994, Mullahey 1997) - and are typically motivated by the fact that in many ordered choice situations, a large proportion of empirical observations fall into a single particular choice category which appears "inflated" relative to the others. Significantly, the importance of not accounting for such category inflation is underlined by the fact that it can lead to mis-specification, biased estimates, incorrect inference and erroneous policy advice.

Such models have been applied in fields such as economics, political science, and medical statistics, and can be divided into two main variants. First, the *zero-inflated ordered probit* ($ZIOP$) model, in which an excess of observations is observed at one end of the choice spectrum. The popularity of the $ZIOP$ modeling framework is reflected in its recent incorporation into mainstream statistical software (e.g., $STATA$ 15, *Limdep/NLogit*), and has been used to explain a variety of phenomena including: the willingness to pay for renewable energy (Akcura 2015); conflict events (Bagozzi et al. 2015); sports participation (Downward et al. 2011); car sharing (Habib et al. 2012); smoking participation (Harris and Zhao 2007, Gurmu and Dagne 2012); the demand for physical and mental health treatment in the US (Meyerhoefer and Zuvekas 2010); depression and labour market outcomes including absenteeism (Peng et al. 2013); vehicle injury severity (Jiang et al. 2013); and visits to museums and historical sites (Falk and Katz-Gerro 2016).

The second variant is the more recently developed *middle-inflated ordered probit* ($MIOP$) model, which is characterized by a middle outcome being inflated. This type of model has been used to investigate: attitudes towards EU membership (Bagozzi and Mukherjee 2012); monetary policy decisions (Brooks et al. 2012); voters' left-right perception of political parties in Japan (Miwa 2015); community level environmental policy (Zirogiannis et al. 2015); and attitudes towards immigration (Bagozzi et al. 2014).

This paper proposes generalizations to these models that preserve the ordering of outcomes whilst still explicitly accounting for the maintained inflation process. In a setting with J categorical outcomes, instead of having a single 'splitting equation' (see Harris and Zhao 2007), our generalizations require J–1 of these to be estimated. We demonstrate that these generalised models collapse to their associated $ZIOP$ and $MIOP$ counterparts under certain linear parameter restrictions, such that all of the parameter vectors of the J–1 splitting equations are equal. The models are then applied to the data and specifications used in the original contributions of Harris and Zhao (2007) and Bagozzi and Mukherjee (2012). We first revisit the work of Harris and Zhao (2007) - the original paper on the $ZIOP$ model - which explores tobacco consumption behavior at the individual level. Attention then turns to the seminal work of Bagozzi and Mukherjee (2012), who use a $MIOP$ framework to model the presence of "face-saving" middle-category responses in a commonly studied Eurobarometer survey question (European Commission 2002a,b), which measures attitudes towards European Union (EU) membership in EU candidate countries. LR and LM tests favor the generalised models in both applications. This finding, we propose, is important, particularly when recalling that Harris and Zhao (2007) and Bagozzi and Mukherjee (2012) claim to have demonstrated the superiority of the $ZIOP$ and $MIOP$ approaches over the OP one. This paper thus establishes that further improvements can be realized by increasing the flexibility of the $ZIOP$ and $MIOP$ models. Moreover, although our applications use survey data, the statistical framework developed above is applicable to other types of ordered response data where category inflation is hypothesized.

By way of contextualising our contribution, we note that our focus is on inflation in a single categorical outcome deriving from multiple sources. However, category inflation need not be characterised by only one of the outcome categories being inflated. Here, Greene, Harris, and Hollingsworth (2015) estimate a discrete ordered model of self-assessed health in which two outcomes are subject to category inflation. Related work by Cai, Xia, and Zhou (2018) explores the consequences of 'generalized' category inflation for multinomial, ordinal, Poisson, and zero-truncated Poisson outcomes and allow for inflation in multiple categories from a single source;[1] however, unlike our contribution, no testing framework is proposed.

In sum we contribute to the literature in several important ways. Building on the growing trend of discrete choice models with category inflation, we suggest a generalization to the inflation process. This both lends itself to a specification test of such models and adds to a new strand of inflated ordered probit models, that are likely to have widespread applicability across the social and related sciences.[2] For example, the $MIOP$ application focuses on a type of survey question where the response options range from feeling negative to positive about an issue, such that a middle category captures feelings of neutrality or indifference. Such questions are commonplace in questionnaires, which suggests there is potentially considerable scope for the analysis of such data using our proposed models. We now say more about the motivation underlying our methodological approach.

Accounting for the presence of category inflation in an ordered setting raises salient issues regarding how it should be modelled. To motivate our analysis, a useful first step is to consider that even if a categorical ordered outcome is characterised by a considerable amount of observations relative to all others, a $ZIOP$ or $MIOP$ modelling approach may not be warranted. Instead, a standard ordered probit model may be sufficient, in that any category can be 'inflated' through adjustment of the relevant threshold parameters.[3] Adopting such a modelling strategy would amount to explicitly assuming that all model categories are generated by a single data generation process (DGP).

This highlights a defining feature of the $ZIOP$ and $MIOP$ modelling approach: a prior assumption that inflation in a given category is generated by two distinct DGPs. It also leads to a second equally important characteristic of $ZIOP$ and $MIOP$ modelling that is commonly overlooked in the literature: namely, a given category need *not* exhibit a build-up of observations to warrant using an $ZIOP$ or $MIOP$ approach. All that is required is a

[1] The $ZIOP$ model (Harris and Zhao 2007) was initial proposed as a "zero-inflation" extension of the zero-inflated Poisson model (Lambert 1992).

[2] We have made the Gauss code used to estimate all generalised models and specification tests in this paper publicly available. For the MIOP model go to:
https://drive.google.com/drive/folders/1V8JSWUlAeINuoAUQhZ_jji00jE_qHfXw?usp=sharing
Estimation code for the ZIOP model can be found here:
https://drive.google.com/drive/folders/1Wb3CcUU254PBo-OOs_-hsnJG9idh-lbB?usp=sharing

[3] We are grateful to a referee for pointing this out.

belief that one of the observed categories is generated by two distinct DGPs. This need not manifest itself in a noticeable spike in the number observations for a given category, or to cite Harris and Zhao (2007) in the context of the $ZIOP$, "...an excess of zero observations" (p.1074). In this regard whilst an empirical build-up of observations in a given category may lead researchers to suspect that a $ZIOP$ or $MIOP$ modelling approach is appropriate, their application should be strictly hypothesis driven; in turn this will have significant implications for the choice of the model's exclusion restrictions.

Our starting point is to assume that a $ZIOP$ or $MIOP$ modelling approach *is* warranted where the data are assumed to be generated by two DGPs. However, this assumption motivates two important questions. First, can category inflation be the product of more than two DGPs, and if so, how can this be modelled? Second, if the inflated category is generated by more than two DGPs, is it possible to test if using a $ZIOP$ or $MIOP$ approach is too restrictive? Our contribution explicitly addresses these questions by developing a framework that maintains the ordering of categorical outcomes, accounts for the presence of category inflation with $N > 2$ DGPs, and nests the $ZIOP$ and $MIOP$ as a special case under certain parameter restrictions. This latter feature is particularly significant. The DGPs which comprise the $ZIOP$ and $MIOP$ are captured by latent equations. As these processes are unobserved by the researcher, a valid question relates to whether the process driving the category inflation is correctly specified. The extant literature provides no sufficient guidance here. Our generalizations can be used as specification tests of the $ZIOP$ and $MIOP$ models, by permitting us to determine if using a $ZIOP$ or a $MIOP$ model is overly restrictive. Just as significantly, our generalised frameworks represent attractive natural extensions to the $ZIOP$ and $MIOP$ models in their own right.[4] If the $MIOP$ were to additionally incorporate categorical outcomes at the ends of the choice spectrum, the $ZIOP$ could be viewed as a 'special case' of the $MIOP$; the same would apply applies to its respective generalisations. Here, our decision to present zero- and middle- inflated models separately

[4]Gillman et al. (2013) develop a framework based on a very specific case of the generalised $MIOP$ model proposed here: the three outcome case with particular regard to monetary policy. No attempt is made by Gillman et al. (2013) to generalise the model to J outcomes. Further, the possibility that the model can be applied in a $ZIOP$ setting is completely overlooked.

follows a convention that is already established in the empirical literature.

II Generalized Inflated Ordered Probit Models

An inflated ordered probit modelling strategy is appropriate where the response variable of interest is categorical and ordered, and in the extant literature is characterized by the combination of a single binary equation - often termed a "splitting equation" - with a single ordered probit (OP) "outcome equation". The combination of these allows the empirical regularity of a build-up observations in a given category to arise from two distinct data generating processes (DGPs). For a discrete ordered variable with J outcomes, a $ZIOP$ approach is appropriate where a build-up of observations occurs at either end of the choice spectrum, such that for $j = 0, 1, 2, .., J - 1$ ordered categories, the build-up is witnessed in either category 'zero' ($j = 0$) or category $j = J$-1. The $MIOP$ approach is a natural extension to the $ZIOP$ framework, allowing for category inflation associated with a build-up of observations in one of the middle categories - that is, one of the $j = 1, 2, ..., J$-2, outcomes. In what follows we extend these models, maintaining a single ordered probit (OP) outcome equation, but introducing J-1 binary splitting equations, as opposed to a single one. As demonstrated below, this innovation implies that for the generalized versions, the build-up of observations in the inflated category arises due to J distinct DGPs, instead of merely two. This distinction in the inflation process turns out to be very important for the empirical applications.

Consider a discrete random variable y that assumes the discrete ordered values of $y \in 0, 1, ..., J-1$, where we note that for ease of comparison, our notation throughout is consistent with that used in Harris and Zhao (2007). A standard OP approach would map a single latent variable to the observed outcome y via so-called boundary parameters, with the latent variable being related to a set of covariates. Let r denote a binary variable indicating the split between regimes 0 and 1. r is related to a latent variable r^* via the mapping: $r = 1$ for $r^* > 0$ and $r = 0$ for $r^* \leq 0$. The latent variable r^* represents the propensity to be in regime 1 and is defined as

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

where $\mathbf{x}$ is a k_x vector of covariates that determine the choice between the two regimes, $\boldsymbol{\beta}$ a vector of unknown coefficients, and ε a standard-normally distributed error term. Accordingly, the probability of being in regime 1 is given by

$$\Pr(r = 1 \,|\mathbf{x}) = \Pr(r^* > 0|\mathbf{x}) = \Phi(\mathbf{x}'\boldsymbol{\beta}), \tag{2}$$

where $\Phi(.)$ is the cumulative distribution function (CDF) of the univariate standard normal distribution. Outcomes in regime 1 are represented by a discrete variable $\widetilde{y}$ ($\widetilde{y} = 0, 1, ..., J-1$) that is generated by an OP model via a second underlying latent variable $\widetilde{y}^*$

$$\widetilde{y}^* = \mathbf{z}'\boldsymbol{\gamma} + u, \tag{3}$$

with $\mathbf{z}$ being a k_z vector of explanatory variables with unknown weights $\boldsymbol{\gamma}$, and u a standard normal error term. Under the assumption that ε and u identically and independently follow standard Gaussian distributions, the full probabilities for y are

$$\Pr(y) = \begin{cases} \Pr\left(y = 0 \,|\mathbf{z}, \mathbf{x}\right) = \Pr\left(r = 0|\mathbf{x}\right) + \Pr(r = 1, \widetilde{y} = 0|\mathbf{z}, \mathbf{x}) \\ \Pr\left(y = j \,|\mathbf{z}, \mathbf{x}\right) = \Pr\left(r = 1|\mathbf{x}\right) \Pr\left(\widetilde{y} = j \,|\mathbf{z}, \mathbf{x}\right), \quad (j = 1, ..., J-1) \end{cases} \tag{4}$$

which, by independence of ε and u is given by

$$\Pr(y) = \begin{cases} \Pr\left(y = 0 \,|\mathbf{z}, \mathbf{x}\right) = \left[1 - \Phi\left(\mathbf{x}'\boldsymbol{\beta}\right)\right] + \Phi\left(\mathbf{x}'\boldsymbol{\beta}\right) \Phi\left(\mu_0 - \mathbf{z}'\boldsymbol{\gamma}\right) \\ \Pr\left(y = j \,|\mathbf{z}, \mathbf{x}\right) = \Phi\left(\mathbf{x}'\boldsymbol{\beta}\right) \begin{bmatrix} \Phi\left(\mu_j - \mathbf{z}'\boldsymbol{\gamma}\right) \\ -\Phi\left(\mu_{j-1} - \mathbf{z}'\boldsymbol{\gamma}\right) \end{bmatrix}, \; (j = 1, ..., J-2) \\ \Pr\left(y = J-1 \,|\mathbf{z}, \mathbf{x}\right) = \Phi\left(\mathbf{x}'\boldsymbol{\beta}\right) \left[1 - \Phi\left(\mu_{J-2} - \mathbf{z}'\boldsymbol{\gamma}\right)\right]. \end{cases} \tag{5}$$

The framework depicted in expression (5) is the $ZIOP$ model. Here, the probability that a zero observation has been inflated is captured by a combination of the probability of zero from the OP process plus the probability of zero from the splitting equation. This central feature of the model also holds when the model is extended to allow for correlated errors, *viz.*,

$$\Pr(y) = \begin{cases} \Pr(y = 0 \,|\mathbf{z}, \mathbf{x}) = [1 - \Phi(\mathbf{x}'\boldsymbol{\beta})] + \Phi_2(\mathbf{x}'\boldsymbol{\beta}, \mu_0 - \mathbf{z}'\boldsymbol{\gamma}; -\rho) \\ \Pr(y = j \,|\mathbf{z}, \mathbf{x}) = \begin{bmatrix} \Phi_2\left(\mathbf{x}'\boldsymbol{\beta}, \mu_j - \mathbf{z}'\boldsymbol{\gamma}; -\rho\right) \\ -\Phi_2\left(\mathbf{x}'\boldsymbol{\beta}, \mu_{j-1} - \mathbf{z}'\boldsymbol{\gamma}; -\rho\right) \end{bmatrix}, \ (j = 1, ..., J-2) \\ \Pr(y = J-1 \,|\mathbf{z}, \mathbf{x}) = \Phi_2\left(\mathbf{x}'\boldsymbol{\beta}, \mathbf{z}'\boldsymbol{\gamma} - \mu_{J-2}; \rho\right). \end{cases} \tag{6}$$

where ρ is the correlation coefficient ($-1 \leq \rho \leq 1$), and Φ_2 denotes the CDF of the bivariate normal distribution. We refer to the correlated model in (6) as the $ZIOPC$.

Given this assumed form for the probabilities and an independent and identically distributed sample of size $i = 1, \ldots, N$ from the population on $(y_i, \mathbf{z}, \mathbf{x})$, this, and all other models derived below satisfy all of the usual regularity conditions for maximum likelihood estimation. In estimation, to ensure the required ordering of the boundary parameters we specify them as

$$\mu_j = \mu_{j-1} + \exp\left(\xi_j\right), \quad j = 1, 2, ..., J-1 \tag{7}$$

where μ_0 is freely estimated (Greene and Hensher 2010). The full parameter set $\boldsymbol{\theta} = \left(\boldsymbol{\gamma}', \boldsymbol{\beta}', \boldsymbol{\mu}', \rho\right)'$ of the model can be consistently and efficiently estimated using the log-likelihood function

$$\ell(\boldsymbol{\theta}) = \sum_{i=1}^{N} \sum_{j=0}^{J-1} h_{ij} \ln\left[\Pr\left(y_i = j \,|\mathbf{z}, \mathbf{x}, \boldsymbol{\theta}\right)\right], \tag{8}$$

where (8) the indicator function h_{ij} is

$$h_{ij} = \begin{cases} 1 \text{ if individual } i \text{ chooses outcome } j \\ 0 \text{ otherwise.} \end{cases} \quad (i = 1, ..., N; \ j = 0, 1, ..., J-1). \tag{9}$$

In our empirical applications the common sandwich estimator (White 1982) is used to compute standard errors of parameters.[5] Standard errors of secondary estimated quantities,

[5] **CHECK MODEL TYPE THIS QUOTE APPLIES TO**. As stated in Greene and Hensher (2010), page 31, '. . . in almost any case, the sandwich estimator provides an appropriate asymptotic covariance matrix

such as partial effects and summary probabilities are estimated using the delta method. All subsequent models differ only with respect to the probabilities entering the likelihood and the contents of $\boldsymbol{\theta}$. Both latent equations are estimated simultaneously and not sequentially, such that only the joint outcome of the two DGPs captured by (5) is observed. Such a latent class model is an example of a partial observability one (also see Poirier (1980) where this concept is applied in the context of a bivariate probit model) involving two latent equations.[6]

Diagrammatically, the $ZIOP$ model is illustrated in the left hand panel of Figure 1, and comprises the binary probit 'splitting equation', which comprises regimes $r = 0$ and $r = 1$; and an ordered probit (OP) model comprising J categorical outcomes labelled $y = 0, 1, 2, ...J–1$. In many empirical applications, the splitting equation is treated as distinguishing between individuals who are willing to participate ($r = 1$) or not ($r = 0$) in the consumption of a good, typically a social bad. Non-participation decisions may be governed by factors such as health concerns, religious beliefs, ethical considerations, or societal norms, but *not* the price of the good or income constraints. Many real-world examples reflect such behavior: consider decisions not to consume drugs and recreational substances such as alcohol, tobacco, and cannabis. However, non-consumption may still arise if individuals who are are prepared to consume the good in regime $r = 1$ are unable to do so because of income or price constraints. Zero consumption of the good is thus driven by a mixture of non-participants, and participants who are unable to consume.

Now consider the latent class model depicted on the right side of Figure 1, which comprises a single OP model comprising J categorical outcomes labelled $y = 0, 1, 2, ...J–1$, and $J–1$ splitting equations: here, for each $j > 0$ category in the OP model, the individual is 'tempered' towards choosing the zero outcome by a category-specific splitting equation. We refer to this econometric model as the "generalized $ZIOP$" (hereafter $GZIOP$ model). As with $ZIOP$ estimation, all equations in this model are unobserved by the researcher and estimated simultaneously. The observed data is generated due to the joint outcome of J

for an estimator that is biased in an unknown direction'.

[6]The ZIOP model satisfies all of the usual regularity conditions for maximum likelihood estimation and, accordingly, all the usual well-behaved properties of the maximum likelihood estimator follow (Harris and Zhou, 2007). The GZIOP also meets these criteria. This also applies where a middle category is inflated.

DGPs, namely the sum of J–1 binary probit equations and a single OP one; this contrasts with the $ZIOP$ model, which is characterised by two DGPs. In what follows, we demonstrate that the $GZIOP$ model still embodies the important attribute of zero-inflation and collapses to the $ZIOP$ under a certain set of parameter restrictions.

The J–1 splitting equations of the $GZIOP$ have the form

$$r_j^* = \mathbf{x}'\boldsymbol{\beta}_j + \varepsilon_j, \tag{10}$$

which allow for the aforementioned differentiated tempering effects across the $j = 1, 2, ..., J-1$ outcome equation propensities. The associated observability criteria is now given by

$$y_j = \widetilde{y} r_j \tag{11}$$

Under independence, generalizing the $ZIOP$ in this manner yields the $GZIOP$ model which has probabilities of the form

$$\Pr(y) = \begin{cases} \Pr\left(y = 0 \,|\mathbf{z}, \mathbf{x}\right) = \left(\begin{array}{c} \Pr\left(\widetilde{y} = 0|\mathbf{z}\right) \\ + \Pr\left(\widetilde{y} = j|\mathbf{z}\right) \Pr\left(r_j = 0|\widetilde{y} = j, \mathbf{x}\right) \end{array} \right), \quad j = 1, ..., J-1 \\ \Pr\left(y = j \,|\mathbf{z}, \mathbf{x}\right) = \Pr\left(\widetilde{y} = j|\mathbf{z}\right) \Pr\left(r_j = 1|\widetilde{y} = j, \mathbf{x}\right), \quad j > 0 \end{cases} \tag{12}$$

such that

$$\Pr(y) = \begin{cases} \Pr\left(y = 0 \,|\mathbf{z}, \mathbf{x}\right) = \begin{cases} \Phi\left(\mu_0 - \mathbf{z}'\boldsymbol{\gamma}\right) + \sum_{j=1}^{J-2} \left(\begin{array}{c} \Phi\left(\mu_j - \mathbf{z}'\boldsymbol{\gamma}\right) \\ -\Phi\left(\mu_{j-1} - \mathbf{z}'\boldsymbol{\gamma}\right) \end{array} \right) \Phi\left(-\mathbf{x}'\boldsymbol{\beta}_j\right) \\ + \left[1 - \Phi\left(\mu_{J-2} - \mathbf{z}'\boldsymbol{\gamma}\right)\right] \Phi\left(-\mathbf{x}'\boldsymbol{\beta}_{J-1}\right) \end{cases} \\ \Pr\left(y = j \,|\mathbf{z}, \mathbf{x}\right) = \left[\Phi\left(\mu_j - \mathbf{z}'\boldsymbol{\gamma}\right) - \Phi\left(\mu_{j-1} - \mathbf{z}'\boldsymbol{\gamma}\right)\right] \Phi\left(\mathbf{x}'\boldsymbol{\beta}_j\right), \; j = 1, ..., J-2 \\ \Pr\left(y = J-1 \,|\mathbf{z}, \mathbf{x}\right) = \left[1 - \Phi\left(\mu_{J-2} - \mathbf{z}'\boldsymbol{\gamma}\right)\right] \Phi\left(\mathbf{x}'\boldsymbol{\beta}_{J-1}\right) \end{cases} \tag{13}$$

which embodies the required zero-inflation due to the terms $\Pr\left(\widetilde{y} = j|\mathbf{z}\right) \Pr\left(r_j = 0|\widetilde{y} = j, \mathbf{x}\right)$ for $j = 1, ..., J-1$. Zero-inflation is also maintained under the likely scenario of correlated errors, where joint probabilities now become

$$
\Pr(y) = \begin{cases} \Pr(y=0\,|\mathbf{z},\mathbf{x}) = \begin{cases} \Phi(\mu_0 - \mathbf{z}'\boldsymbol{\gamma}) + \sum_{j=1}^{J-2} \begin{bmatrix} \Phi_2\left(\mu_j - \mathbf{z}'\boldsymbol{\gamma}, -\mathbf{x}'\boldsymbol{\beta}_j; \rho_j\right) \\ -\Phi_2\left(\mu_{j-1} - \mathbf{z}'\boldsymbol{\gamma}; -\mathbf{x}'\boldsymbol{\beta}_j; \rho_j\right) \end{bmatrix} \\ +\Phi_2\left(\mathbf{z}'\boldsymbol{\gamma} - \mu_{J-2}, -\mathbf{x}'\boldsymbol{\beta}_{J-1}; \rho_{J-1}\right) \end{cases} \\ \Pr(y=j\,|\mathbf{z},\mathbf{x}) = \Phi_2\left(\mu_j - \mathbf{z}'\boldsymbol{\gamma}, \mathbf{x}'\boldsymbol{\beta}_j; \rho_j\right) - \Phi_2\left(\mu_{j-1} - \mathbf{z}'\boldsymbol{\gamma}, \mathbf{x}'\boldsymbol{\beta}_j; \rho_j\right)\ ,\ j = 1, ..., J-2 \\ \Pr(y=J-1\,|\mathbf{z},\mathbf{x}) = \Phi_2\left(\mathbf{z}'\boldsymbol{\gamma} - \mu_{J-2}, \mathbf{x}'\boldsymbol{\beta}_{J-1}; \rho_{J-1}\right) \end{cases} \tag{14}
$$

The correlated $ZIOP$ model defined by the set of equations in (14) is referred to as the $GZIOPC$. Unlike the $ZIOPC$ the model is characterized by J–1 correlation coefficients denoted ρ_j $\forall$ $j = 1, 2, 3...J$–1. One *could* allow for a more complex correlation structure amongst all of the stochastic elements of the generalised variants. The generalisation in (14) allows for correlations between the stochastic elements relating to the inflation and outcome equations; this follows the approach taken in the original literature. However, it would also be possible to allow for correlations *across* the splitting equations in the generalised variants. Whilst theoretically this poses no additional issues (apart from more complicated expressions for the probabilities), this is arguably not appropriate here. This is because the correlations across inflation equations would necessarily correspond to different individuals. Thus there is less *a priori* expectation that these should be related, as compared to those equations relating to the *same* individual.

Using the model of the equations in (14) we now show that the generalized $ZIOP$ variants outlined above collapse to their original counterparts under a set of simple linear parameter restrictions. This implies that the model on the right side of Figure 1 nests the model depicted on the left. In the generalised model(s) identification requires the data to identify J–1 splitting equations as opposed to a single one. One implication of this model characteristic is that compared to the non-generalised model variants, the choice of exclusion restrictions assumes a more prominent role, as several splitting equations require identification instead of one. More generally, behavioral identification in our generalised models requires that there are no empty sets of individuals in expression (3) that are pushed towards an inflated outcome for each of the model's J–1 splitting equations. This issue is revisited when the finite sample properties of our models are explored in Section IV.

In both of our empirical applications, no evidence of identification issues are found to be present. A possible generalisation of the model could entail different sets of variables in the various splitting equations, although the original $ZIOP$ model would no longer be nested. In general, weak identification is likely to be evidenced by instances of model non-convergence and/or estimated model probabilities close to zero. As Greene, Rose, and Hensher (2015) note in the context of a latent class ordered choice model: "Signature features of a model that has been over-fit will be exceedingly small estimates of the class probabilities, wild values of the structural parameters and huge estimated standard errors." (p.719).

Consider imposing the linear set of restrictions that $\boldsymbol{\beta}_1 = \boldsymbol{\beta}_2 = \cdots = \boldsymbol{\beta}_{J-1}$ and $\rho_1 = \rho_2 = \cdots = \rho_{J-1}$ on (14). This yields

$$\begin{cases} \Pr\left(y=0\left|\mathbf{z},\mathbf{x}\right.\right) = \begin{cases} \Phi\left(\mu_0 - \mathbf{z}'\boldsymbol{\gamma}\right) + \sum_{j=1}^{J-2} \begin{bmatrix} \Phi_2\left(\mu_j - \mathbf{z}'\boldsymbol{\gamma}, -\mathbf{x}'\boldsymbol{\beta};\rho\right) \\ -\Phi_2\left(\mu_{j-1} - \mathbf{z}'\boldsymbol{\gamma}; -\mathbf{x}'\boldsymbol{\beta};\rho\right) \end{bmatrix} \\ +\Phi_2\left(\mathbf{z}'\boldsymbol{\gamma} - \mu_{J-2}, -\mathbf{x}'\boldsymbol{\beta};\rho\right) \end{cases} \\ \Pr\left(y=j\left|\mathbf{z},\mathbf{x}\right.\right) = \Phi_2\left(\mu_j - \mathbf{z}'\boldsymbol{\gamma},\mathbf{x}'\boldsymbol{\beta};\rho\right) - \Phi_2\left(\mu_{j-1} - \mathbf{z}'\boldsymbol{\gamma},\mathbf{x}'\boldsymbol{\beta};\rho\right) \ , \ j=1,...,J-2 \\ \Pr\left(y=J-1\left|\mathbf{z},\mathbf{x}\right.\right) = \Phi_2\left(\mathbf{z}'\boldsymbol{\gamma} - \mu_{J-2},\mathbf{x}'\boldsymbol{\beta};\rho\right) \end{cases} \tag{15}$$

where we note that while the expressions for $\Pr\left(y=j\left|\mathbf{z},\mathbf{x}\right.\right)$ and $\Pr\left(y=J-1\left|\mathbf{z},\mathbf{x}\right.\right)$ immediately collapse to those in expression (6), the $\Pr\left(y=0\right)$ expression in (15) can be constructed using 1 minus the sum of the $\Pr\left(y=J-1\left|\mathbf{z},\mathbf{x}\right.\right)$ and all $\Pr(y=j\left|\mathbf{z},\mathbf{x}\right.)$, $\forall\ j=1,2,...J-2$ terms to give

$$\Pr\left(y=0\left|\mathbf{z},\mathbf{x}\right.\right) = \left[1-\Phi\left(\mathbf{x}'\boldsymbol{\beta}\right)\right] + \Phi_2\left(\mathbf{x}'\boldsymbol{\beta},\mu_0 - \mathbf{z}'\boldsymbol{\gamma};-\rho\right). \tag{16}$$

This also yields the result in (6), and is straightforward to verify. Using (15) and (16) yields

$$\Pr\left(y=0\right) = 1 - \overbrace{\sum_{j=1}^{J-2}\left[\Phi_2\left(\mu_j - \mathbf{z}'\boldsymbol{\gamma},\mathbf{x}'\boldsymbol{\beta};-\rho\right) - \Phi_2\left(\mu_{j-1} - \mathbf{z}'\boldsymbol{\gamma},\mathbf{x}'\boldsymbol{\beta};-\rho\right)\right]}^{\Pr(y=j\ \forall\ j=1,2,...J-2)} - \overbrace{\Phi_2\left(\mathbf{z}'\boldsymbol{\gamma} - \mu_{J-2},\mathbf{x}'\boldsymbol{\beta};\rho\right)}^{\Pr(y=J-1)} \tag{17}$$

which can be expanded as follows

$$\Pr(y=0) = 1 - \left\{ \begin{array}{c} \left[\Phi_2\left(\mu_1 - \mathbf{z}'\boldsymbol{\gamma}, \mathbf{x}'\boldsymbol{\beta}; -\rho\right) - \Phi_2\left(\mu_0 - \mathbf{z}'\boldsymbol{\gamma}, \mathbf{x}'\boldsymbol{\beta}; -\rho\right)\right] \\ + \left[\Phi_2\left(\mu_2 - \mathbf{z}'\boldsymbol{\gamma}, \mathbf{x}'\boldsymbol{\beta}; -\rho\right) - \Phi_2\left(\mu_1 - \mathbf{z}'\boldsymbol{\gamma}, \mathbf{x}'\boldsymbol{\beta}; -\rho\right)\right] \\ + \left[\Phi_2\left(\mu_3 - \mathbf{z}'\boldsymbol{\gamma}, \mathbf{x}'\boldsymbol{\beta}; -\rho\right) - \Phi_2\left(\mu_2 - \mathbf{z}'\boldsymbol{\gamma}, \mathbf{x}'\boldsymbol{\beta}; -\rho\right)\right] \\ \vdots \\ + \left[\Phi_2\left(\mu_{J-2} - \mathbf{z}'\boldsymbol{\gamma}, \mathbf{x}'\boldsymbol{\beta}; -\rho\right) - \Phi_2\left(\mu_{J-3} - \mathbf{z}'\boldsymbol{\gamma}, \mathbf{x}'\boldsymbol{\beta}; -\rho\right)\right] \\ + \left[\Phi(\mathbf{x}'\boldsymbol{\beta}) - \Phi\left(\mu_{J-2} - \mathbf{z}'\boldsymbol{\gamma}, \mathbf{x}'\boldsymbol{\beta}; \rho\right)\right] \end{array} \right\} \quad (18)$$

After cancelling terms and algebraic manipulation, it can be verified that

$$\Pr(y=0) = \left[1 - \Phi\left(\mathbf{x}'\boldsymbol{\beta}\right)\right] + \Phi_2\left(\mu_0 - \mathbf{z}'\boldsymbol{\gamma}, \mathbf{x}'\boldsymbol{\beta}; -\rho\right). \quad (19)$$

Substituting (19) into (15) results in *GZIOPC* probabilities that are identical to the *ZIOPC* probabilities in expression (5). That is, the *GZIOPC* collapses to – and therefore nests – the *ZIOPC*. Further, setting $\rho = 0$ in (19) yields probabilities that are identical to the *ZIOP* probabilities in expression (5), *viz.*

$$\Pr(y=0) = \left[1 - \Phi\left(\mathbf{x}'\boldsymbol{\beta}\right)\right] + \Phi(\mathbf{x}'\boldsymbol{\beta})\Phi\left(\mu_0 - \mathbf{z}'\boldsymbol{\gamma}\right). \quad (20)$$

The *GZIOPC* also collapses to the *ZIOP*, albeit under the alternative set of parameter restrictions $\boldsymbol{\beta}_1 = \boldsymbol{\beta}_2 = \boldsymbol{\beta}_3 ... = \boldsymbol{\beta}_{J-1}$ and $\rho_j = 0 \; \forall \; j = 1, 2, ...J-1$. Lastly, imposing the latter set of restrictions implicitly reduces the *GZIOPC* model to its uncorrelated counterpart in (13), the *GZIOP*. The sets of parameter restrictions described above provide tests of: (*i*) the more flexible functional form of the *GZIOPC* model versus the simpler nested forms of the usual *ZIOPC* and *ZIOP* models; and (*ii*) the *GZIOP* versus the *ZIOP* model. A noteworthy property of the generalised variant proposed here is that it is not constrained by the "parallel regression" assumption inherent in the ordered probit, *ZIOP* and *ZIOPC* models; this also applies to models with middle-inflation, which we now discuss.

Building on the *ZIOP* model, two contributions – Bagozzi and Mukherjee (2012) and

Brooks et al. (2012) – independently suggested the *middle-inflated ordered probit* ($MIOP$) model to allow for inflation in an arbitrary middle category.[7] Whilst each of these contributions restricts the analysis to three categorical outcomes, our modelling framework applies to instances where $J > 3$; in keeping with our discussion of the $ZIOP$ model and its generalisation, the presence of $j = 0, 1, 2, ...J–1$ categories is also considered . Diagrammatically, the $MIOP$ this is depicted on the left hand side of Figure 2. It comprises a single splitting equation and an OP model, both of which are unobserved by the researcher. Here, m denotes an inflated middle category, which can assume any of the values in the set $j \in \{1, 2...J - 2\}$; the splitting equation now distinguishes between observational units in the inflated middle category ($r = 0$) and those in all other categories ($r = 1$).

Following logic analogous to that used for the $GZIOP$, we can generalize the $MIOP$. Here, we stress that due to its similarity with the zero-inflation case, a formal exposition of the $MIOPC$ and its relationship to the $GMIOPC$ is given in Appendix B; the same principles apply. The generalised model (hereafter $GMIOP$) is illustrated in the right-hand panel of Figure 2: it shows that for any given propensity towards a given category $j \neq m$ in the OP equation, there is a movement towards the inflated middle category, m. Naturally, the $MIOP$ and its generalisation are related in an analogous way to that of the $ZIOP$ and the $GZIOP$, and we can also consider model variants with correlated errors which we label $MIOPC$ and $GMIOPC$. This means that the model depicted on the right of Figure 2 can nest the model depicted on the left under appropriate parameter restrictions. Testing the restrictions associated with these model variants entails testing (i) the more flexible functional form of the $GMIOPC$ model versus the simpler nested forms of the $MIOPC$ and $MIOP$ models and (ii) the $GMIOP$ versus the $MIOP$ model. As with the $GZIOP$ model, the $GMIOP$ is still an inflated ordered probit model. The ordering of outcomes is still preserved, middle-inflation arises due to $J–1$ distinct DGPs as opposed to just one, and all (latent) equations in the model are estimated simultaneously. Appendix C also establishes

[7] Bagozzi and Mukherjee (2012) were the first to use the term 'middle-inflated'. Brooks et al. (2012) refer to their model merely as an 'inflated ordered probit'. In this contribution we use the former nomenclature, and suggest that the term *inflated ordered probit* (IOP) model may be better viewed as encompassing both the $ZIOP$ and the $MIOP$ model classes.

that our proposed generalisations are coherent and demonstrates that our models neither nest, nor are nested by the *generalised ordered probit* ('GOP') model (Terza 1985).

To test the hypotheses associated with the various sets of parameter restrictions described above for our inflated models, two approaches are used. First, we use the standard LR test. Second, an LM test is proposed.[8] This is an appealing specification test for the $ZIOPC$ and $MIOPC$ models *versus* their generalized alternatives, as it only requires estimation of the simpler nested models. It involves evaluation of the score vector of the more general model evaluated at parameter values under the null hypothesis (*i.e.*, at the $ZIOPC$ or $MIOPC$ ones). For instance, testing between $GZIOPC$ versus $ZIOPC$ models yields an LM statistic is given by

$$LM^{ZIOPC} = (\boldsymbol{\nabla\beta}, \boldsymbol{\nabla\gamma}, \nabla\mu_0, \boldsymbol{\nabla\xi}, \boldsymbol{\nabla\rho})' \left[\mathbf{I}\left(\hat{\theta}_R\right)\right]^{-1} (\boldsymbol{\nabla\beta}, \boldsymbol{\nabla\gamma}, \nabla\mu_0, \boldsymbol{\nabla\xi}, \boldsymbol{\nabla\rho}) \sim \chi^2_q \qquad (21)$$

which is evaluated at the relevant parameter restrictions as defined by the appropriate null hypothesis. Under H_0, LM^{ZIOPC} will be a chi-squared variate where q denotes the appropriate number of parameter restrictions. If the alternative model is the uncorrelated generalised version, one would omit the relevant partition of the score vector $(\boldsymbol{\nabla\rho})$**.** As is common practice, the outer product of gradients ($OPGs$) is used to estimate the inverse of the variance of the score vector, $\left[\mathbf{I}\left(\hat{\theta}_R\right)\right]^{-1}$ (Greene 2012). For the zero-inflated and middle-inflated approaches, derivations of the score vectors for the LM test can be found in Appendices A and B, respectively. Reassuringly, the results of the LR and LM tests are very similar in all of our empirical applications, suggesting that the log-likelihood function is well-behaved, and further, that standard asymptotic theory performs well.

Finally, it is also possible to consider subsets of a given generalised model as the model under the alternative and adapt the LM test appropriately. This would likely increase power in that particular direction. For example, only subsets of parameters may vary. In the absence of any prior information, such an approach is not recommended, as such tests

[8] Our testing framework focuses on instances where one inflated model nests another. In relation to the problem of zero-inflation in the Poisson counts literature, Wilson (2015) argues that the widespread practice of using the Vuong test as a test of zero inflation in a non-nested setting is erroneous.

would invariably be based on mis-specified alternative models that would likely adversely affect the test performance.

III Data

To explore the performance of our generalizations and testing framework, we consider two key empirical examples from the literature. Each uses an inflated ordered probit approach to model responses in a large-scale survey data set. Our $GZIOP$ application re-visits the original contribution of Harris and Zhao (2007). Their health economics based application analyzes the determinants of participating in tobacco consumption. A zero-inflated application is deemed appropriate in that zero tobacco consumption may be construed as being determined by two DGPs: non-participation due to, for example, health and legal concerns; and further, non-participation due to being at a corner solution associated with a standard consumer demand problem, whereby individuals will not smoke if the price rises above a certain threshold, or income falls below a certain threshold. Their data is drawn from the 1995, 1998 and 2001 surveys of the Australian *National Drug Strategy Household Survey* (NDSHS, 2001), and comprises a total of over $40,000$ respondents. Removal of missing values leads to an estimation sample of $28,813$ individuals. Information on individuals' consumption of tobacco is available via a discrete variable measuring the intensity of consumption. Specifically, respondents are asked: "How often do you now smoke cigarettes, pipes or other tobacco products?", where the responses take the form of one of the following choices: not at all ($y = 0$); smoking less frequently than daily ($y = 1$); smoking daily with less than 20 cigarettes per day ($y = 2$); and smoking daily with 20 or more cigarettes per day ($y = 3$). In terms of consumption frequencies, 76% of observations are non smokers, 4% smoke weekly or less, 13.8% smoke daily but less than 20 per day, and 6.2% smoke daily and consume more than 20 cigarettes a day.

Covariates in the splitting (or "participation") equation include factors relating to individuals' attitudes towards smoking and health concerns, and include variables that reflect education levels and other standard socio-demographic variables such as income, marital status, age, gender and ethnic background. In the OP (or "outcome") equation, covariates include standard demand-schedule variables such as income and own- and cross-drug prices (in the results presented below, $Ln(P_{A/M/T})$ refers to the natural log of the price of alcohol / marijuana / tobacco, respectively), in addition to standard socio-demographic factors such as those related to a respondent's age, to capture any heterogeneity in consumption behavior among smokers. The specification shares 13 common variables in the splitting and OP equations, and is characterized by: $N = 28,813$; $J = 4$; $k_x = 16$; and $k_z = 18$.

For the $GMIOP$ attention turns to the work of Bagozzi and Mukherjee (2012), who use a $MIOP$ framework to analyze individual responses in a data set that explores respondents' attitudes towards European Union (EU) membership in EU accession countries; significantly, the data set in question has also been the subject of scrutiny in other contributions to the political science literature (Gabel 1998; Carey 2002; Elgün and Tillman 2007). When asked about their attitudes towards joining the EU, respondents choose from one of three alternatives: *a bad thing*; *neither good nor bad*; or *a good thing.* The associated response frequencies for these are 10.83%, 33.07% and 56.10%, respectively. The authors hypothesize that the middle category contains responses from two distinct sources: "informed" respondents with good knowledge of the impact of EU membership; and "uninformed" respondents, who select *neither good nor bad* as a "face-saving measure". This results in middle category inflation, warranting a $MIOP$ approach. Here, we emphasize the hypothesis driven nature of category inflation in this application: in keeping with the discussion in Section I the inflated category is not characterised by an excess of middle category observations relative to other categories. A fourth '*do not know*' category is treated as being a "neither good nor bad" response by Bagozzi and Mukherjee (2012) which is common in the literature. The authors report that their findings remain unchanged when "do not know" responses are dropped from estimations. The model thus comprises a splitting equation which captures the impact of covariates on the likelihood that respondents are either informed or uninformed; and an outcome equation (OP) which estimates the impact of a second variable set on the probabilities of observing each ordered survey response category, which is estimated conditional on the respondent being informed. The specification shares 8 common variables in the two equations, and is characterized by: $N = 9,113$; $J = 3$; $k_x = 12$; and $k_z = 16$.

The splitting equation covariates capture if a respondent is knowledgeable about the EU and its impact. Variables specific to this equation measure: How often a respondent watches the news (*media*); the extent of an individual's knowledge of the EU based on a subset of true-false questions asked as part of the survey ('*True EU knowledge*'); and whether or not respondents were aware of their country's bid for EU membership ('*EU-bid knowledge*'). The common variables that appear in the splitting equation are: An ordinal measure coded as 1 if the respondent reports discussing politics with friends as "never", 2 if "occasionally," and as 3 if "frequently" ('*discuss politics*'); a geographical location dummy ('*rural*'); a gender dummy coded as 1 for female on the basis that women are less likely to support EU membership as they are more vulnerable to the costs of integration that occur when states join the EU ('*female*'); age ('*age*'); whether the individual is studying at a college or university ('*student*'); and indicator variables for educational attainment ('*educ high*', '*educ high-mid*', '*educ low-mid*').

Variables exclusive to the outcome (OP) equation comprise: an income measure to test the hypothesis that individuals with higher incomes are more likely to view EU membership in a positive way since they benefit from European integration ('*income*'); variables that account for a respondent's occupational status ('*professional*'; '*executive*'; '*manual*'; '*farmer*'); whether or not they are unemployed ('*unemployed*'); and variables capturing the extent to which domestic political institutions are trusted ('*political trust*'), and if respondents are xenophobic (*xenophobia*). However, prior to conducting estimations on both datasets, the performance of our proposed models is explored using Monte Carlo (MC) experiments.

IV Finite sample performance

To ascertain finite sample performance of our tests, we consider a range of Monte Carlo (MC) experiments. These experiments are based on the same data and specifications of the $ZIOP$ and $MIOP$ models considered in Harris and Zhao (2007) and Bagozzi and Mukherjee (2012); such that sample sizes for each are $N = 28,813$ and $9,113$, respectively. The number of repetitions was set to $2,000$, where all simulation 'noise' had effectively settled after $1,000$ repetitions. The results are in Table 1. Panel A presents the empirical size of the tests, and the first column identifies the true DGP and the respective degrees of freedom for each test (*df*). For each DGP, three tests – each between a generalised model and a null, non-generalised one – are performed.

Panel A presents results for the zero-inflated application, and tests between: $GZIOP$ vs $ZIOP$; $GZIOPC$ vs $ZIOPC$; and $GZIOPC$ vs $ZIOP$, with $J = 3$ outcomes. Row 1 has a $ZIOP$ DGP with $df = 13, 14, 15$, respectively. At nominal 5% size, we see that all empirical sizes are very close to this, even when the null model is the $ZIOPC$. Row 2 repeats the exercise, but for a true DGP of $ZIOPC$. Here the empirical size is again very close to the nominal one (at 5.8%). The tests also have good 'power' in correctly rejecting the uncorrelated versions of the model (38% and 49%, respectively). Row 3 considers the implications of extending the choice set to a larger number of outcomes, one of which is now relatively sparsely populated; as can be seen, the empirical sizes remain very close to the nominal ones.

As rejection of the null models may reflect other forms of model mis-specification, we also generate under ordered probit and parallel regression (Brant 1990) models. These *quasi*-power experiments reflect likely forms of serious model mis-specification encountered with our type of data. The OP model is based on an equation of the form of expression (3). For the parallel regression model, the data is generated by multiple $\boldsymbol{\gamma}_j$ vectors generated by independent binary models for all observed values of j. The results are presented in rows 4 and 5, respectively. All tests have good general 'power' (24%–36%) against the OP DGP. Against the parallel regression model, all tests similarly exhibited reasonable 'power' (at around 14%).

To complement the zero-inflated experiments in Panel A of Table 1, we also considered a variant of the $GZIOP$ model with no tempering for one of the outcomes.[9] Such a model does not collapse to the null $ZIOP$ model under any set of simple linear parameter restrictions. In experiments, this model variant failed to converge in nearly 50% of instances. When

[9] The splitting equation corresponding to the $j = 3$ outcome was removed.

convergence was achieved, the LM test rejected the null model in 100% of instances, and the estimated tempering probabilities of the true zero amount were very close to zero. Clear evidence of model mis-specification in the tempering equation of the non-tempered outcome presented itself in the form of very large coefficients and extremely high standard errors. These findings add to the evidence that the LM test performs well as a general specification test: they suggest that model failure in estimation would also indicate a mis-specified model, as would obtaining tempering probabilities in the splitting equations that are very close to zero.

Panel B of Table 1 presents similar results for the middle-inflated experiments and tests: $GMIOP$ *vs* $MIOP$; $GMIOPC$ *vs* $MIOPC$; and $GMIOPC$ *vs* $MIOP$. Row 6 corresponds to a $MIOP$ DGP and is based on the full sample ($N = 9,113$) and has $J = 3$. Here, all empirical sizes are very close to nominal ones. Row 7 considers a $MIOPC$ DGP. At 6%, empirical size is again very close to the nominal one. These tests have reasonable 'power' at picking-up the mis-specified uncorrelated model, with rejection probabilities of around 18% and 24%**.** The effect of reducing the *df* is explored here in row 8, where the $MIOP$ is re-estimated and statistically insignificant variables are removed. This respectively yields *df* $= 7, 8, 9$; again, all tests are correctly sized.

As our tests are all asymptotic, the implications for their properties of estimating using a smaller sample are also explored. This is achieved by taking the (already) relatively small sample in the $MIOP$ example and randomly removing 50% of the observations, yielding $N = 4,556$. The re-sized sample marginally worsens the performance of the tests, with all of them being slightly over-sized at around 7%–8%. Finally, *quasi*-power experiments were once again performed by generating under an OP model and parallel regressions (rows 9 and 10). Again, all tests behave exceptionally well as general specification ones, as indicated by high rejection probabilities of up to nearly 80% in some instances. In summary, for both the zero-inflated and middle inflated experiments, all LM tests appear correctly sized, and typically have good 'power' in identifying mis-specified models. Finally, we note that rejection of uncorrelated versions, may simply be a sign of a mis-specified correlated model.

Power experiments

Using the observed data, we also conducted power experiments based on the null models of $ZIOP$ and $MIOP$ versus their generalised forms. In all experiments our approach is characterised by taking the estimated value of $\boldsymbol{\beta}$ in each null model, setting $\boldsymbol{\beta}_j = \boldsymbol{\beta}\ \forall\ j$ in the corresponding generalised set-up, and perturbing a single parameter β_0 in a single splitting equation by successively larger increments. For brevity, we only report power runs for the non-correlated DGPs and their associated LM tests. The power analysis results are shown in Panels A ($ZIOP$) and B ($MIOP$) of Figure 3, and cover experiments performed using alternative df and sample size.

In the $ZIOP$ experiments two curves are charted, both of which utilize the full data sample: one corresponds to J=3 categorical outcomes (df=13); and another to J=4 (df=26). The curve corresponding to the higher df has uniformly higher power, where we note that increasing the number of categorical outcomes from three to four is responsible for the increase in df. Whilst relatively larger parameter perturbations are required to induce rejections under J=3, both tests have the 'usual' shaped power curves and our analysis suggests both tests have good power.

For the $MIOP$ experiments, which are all characteried by $J = 3$ categories, we initially focus on two experiments that use the full sample but which are differentiated by dropping insignificant variables from the splitting equations. This has the impact of reducing the df from $df = 12$ to $df = 7$. Panel B shows that the difference in df has no discernible effect on power and is arguably to be expected given the nature of our peturbations: specifically, in each $MIOP$ experiment the single perturbed parameter differs from only a single estimated parameter. This is unlike the $ZIOP$ experiments in Panel A, where each experiment is distinguished by a different number of splitting equations: in the $d = 26$ experiment, there are three such equations, and the single perturbed parameter thus differs from two single estimated parameters; however, in the $d = 13$ experiment, the presence of only two splitting equations means that the single perturbed parameter differs from only a single estimated parameter. One might have anticipated greater differences in power gains here, given the large difference in df in the $ZIOP$ experiments; when the df is smaller, model failure associated with a single parameter could be interpreted as being more severe.[10]

Finally, we also conducted experiments with a small sample size (small N, $df = 12$) under the null of $MIOP$. With the reduced sample, a reduction in power is observed relative to other $MIOP$ experiments, in that relatively larger parameter perturbations are required to lead to model rejection. Despite the relative reduction in power, we note that all $MIOP$ tests have the 'usual' shaped power curves, and like the $ZIOP$ experiments, exhibit good power. Significantly, our results demonstrate that the ability of the tests to identify $ZIOP$ ($MIOP$) model mis-specification in the direction of the $GZIOP$ ($GMIOP$) one(s) is an increasing function of both the number and size of perturbations from the null. The ability to identify model mis-specification also responds to changes in the *df* of the test and the sample size. Differences in the way that the *df* are obtained may have effects on the power of the tests. However, as with all MC experiments, the results may be dependent upon the particular experiments considered. We now turn to model estimation.

V Estimation

As noted above, such zero-inflated models are examples of latent class models which exhibit partial observability: observationally equivalent outcomes can arise from distinct DGPs. For example, in Harris and Zhao (2007) an individual makes a participation decision, and for participants, a consumption decision is made. The fact that consumption can still be zero for some participants gives rise to zero-inflation.

To rationalise the GZIOP model, the ordered consumption levels would be driven by an OP process and the propensity for zero-consumption corresponds to non-participation. Significantly the theory of rational addiction (Becker and Murphy 1988) assumes that some individuals are rational in going "cold-turkey" – that is, switching from positive consumption levels – as captured by the latent ordered probit equation – to zero, as captured by the j=1,2,J-1 binary equations. To accommodate this requires that corresponding to each posi-

[10] Although not reported here, significant power gains also occurred in cases where (i) a full, single vector was perturbed and (ii) all vectors were perturbed. Both of these alternative scenarios showed comparatively higher power compared to the single-parameter experiment. This is because the single parameter experiment represents the scenario where the test is most likely not to perform well, as it is closest to the null.

tive consumption level is a separate binary equation which splits individuals into two types: those remaining at their inherent consumption level, and 'quitters' who are "pushed" towards zero. The $GZIOPC$ model developed above allows for this possibility. The is nothing that imply that all members have the propensity to quit. In essence, one is testing whether a single equation – in Harris and Zhao (2007) representing participation – is sufficiently general to represent *all* of the types of zero that could arise.

It is informative to consider the behavioral assumptions required for model identification. The $ZIOP$ model is only identified if the inflated category observed in the empirical data is composed of two types of observations. In the smoking application, these respective observations correspond to the non-smokers associated with the inflation equation in expression (1), and infrequent smokers associated with the consumption equation in expression (3). The identification of the generalised model is somewhat stricter. The inflated category observed in the data is instead composed of individuals with an inherent consumption level of zero in the consumption equation in (3), and J-1 distinct groups of smokers with positive inherent consumption levels in (3), who are "pushed" towards zero consumption by the J-1 splitting equations given by (10). Behavioral identification in the $GZIOP$ therefore requires that there are no empty sets of individuals in expression (3) that are pushed towards zero-consumption via (10), for all $j \geq 1$. In our empirical application this is attributable to factors such as health status, medical considerations, income, and wealth. Here, it is reasonable to expect that if the *total* population from which the sample is drawn is characterised by no empty sets of individuals, the use of large scale datasets - as used in our empirical applications - will mitigate the problem of failing to identify all of these sets of individuals, especially when $J-1$ is large. In practice, the presence of empty sets may manifest itself in the form of one or more of the r_j^* splitting equations being characterised by negligible tempering probabilities. That is, the model will appear to be 'weakly identified'. Significantly, our empirical applications exhibit little evidence of this form of weak identification, in that all of the estimated tempering probabilities associated with the J-1 splitting equations diverge from zero. We also note that if evidence of such empty sets is found, the generalised model may be re-specified by omitting the affected r_j^* splitting equations, and re-estimating without them. Whilst the resulting specification will still be an inflated model, it will no longer be 'generalised', in that the standard $ZIOP$ (and $MIOP$) model will no longer be nested. This would consequently mean that that our proposed LR and LM tests are inappropriate. Whilst not the focus of this paper, the possibility of refining the $GZIOP$ (or $GMIOP$) in the way described above suggests that the generalised class of inflated model developed in this contribution forms part of a much broader model class for analysing category inflation.

Table 2 reports the results of the LM tests. All of the $ZIOP$ variants are overwhelmingly rejected in favour of the $GZIOP$ models. Moreover, the $GZIOP$ is rejected in favour of its correlated variant, the $GZIOPC$. In addition to the LM tests, Table 2 reports the corresponding LR tests, which closely mirror the LM ones. We stress here that rejection does not necessarily imply that the generalised variant is "correct": it is possible to reject a false model against many alternatives, even if none of the alternative models are correct (Davidson and MacKinnon 1987). Our findings are also re-visited in the Monte Carlo section in Section IV, in which a number of experiments are performed. Our findings indicate that LM tests are correctly sized, and have good power in identifying mis-specified models. The closeness of the LR and LM test statistics suggests that in the case of the present application, the log-likelihood function is well-behaved and standard asymptotic theory performs well.

Given the evidence to support the presence of correlated errors, Table 3 presents the $GZIOPC$ and $ZIOPC$ output equation parameters for comparison purposes. Doing so enables us to directly compare how model inference changes as a result of using a generalised model instead of its nested equivalent. With respect to the ρ's, although they are all negative and strongly significant across specifications, some noteworthy differences in size do arise. More importantly however, are differences across the structural parameters. While income is positive across both specifications, it is more significant in the $GZIOPC$ model, as well as being over twice the size. Whilst this implies a standard demand function result with tobacco consumption increasing with income levels, it also indicates a more powerful effect for income in the generalised model. In contrast, cross-drug prices corresponding to alcohol, marijuana, and tobacco all have noticeably smaller parameters in the $GZIOPC$ than for the $ZIOPC$. This suggests that individuals' demand for tobacco is less responsive to changes in drug prices than previously estimated. Other variables are similar in size and significance.

Of particular interest is a comparison of the parameter estimates in the single splitting equation of the $ZIOPC$, as compared to estimates associated with its generalized variant $GZIOPC$. These estimates are presented in Table 4. For the $GZIOPC$ we witness some very large changes across $j = 1, 2$ and 3 as compared to $ZIOPC$; here, we recall that implicitly the restriction of the latter is that these are all equal across j.

It is interesting to put an economic interpretation on these differences. Consider the $ZIOPC$ and $GZIOPC$ results: Ln(income) has a small (-0.067) but significant effect in the $ZIOPC$ model. The negative effect found here implies that higher income individuals are associated with a higher propensity for zero (*i.e..*, non-consumption) arising from the splitting equation. Harris and Zhao (2007) argue that income, being a proxy for social status/class, will be negatively correlated with smoking participation rates. As with the $ZIOPC$, negative (positive) coefficients in the $GZIOPC$ splitting equations are also associated with higher (lower) probabilities of tempering towards zero consumption. For the $GZIOPC$, Ln(income) is insignificant and positive for $j = 1$ (0.067), highly significant, negative and slightly smaller for $j = 2$ (-0.075), and highly significant and smaller still for $j = 3$ (-0.181); similar results, not reported here, also arise when the related models with independent errors are compared. Qualitatively similar results are in fact found for all splitting equation and outcome equation variables. For those individuals with an underlying propensity for low amounts of smoking ($j = 1$), the insignificant coefficient means that higher income individuals are more likely to remain at this underlying propensity. This could imply that for higher income earners, there is less social stigma associated with "social (infrequent) smoking". However, for higher underlying intensity levels ($j = 2, 3$) the fact that the income effect becomes negative and increasingly pronounced as j increases implies that for higher underlying intensity levels, increasing income is now associated with an increasing probability of these individuals tempering this intensity down to zero consumption. In general, the large and significant negative tempering effects in the $j = 3$ equation could also imply that these factors are associated with individuals going "cold turkey", that is, moving frequently between high and zero consumption levels.

Some variables that are statistically insignificant in the single $ZIOPC$ splitting equations are highly significant in the $GZIOPC$ ones. For example, the dummy variable that corresponds to whether an individual's highest level of education is Year 12 has no effect in the $ZIOPC$ model, but for the $GZIOPC$ exerts a strong positive effect for $j = 3$. Estimation using the $ZIOPC$ can therefore be viewed as leading to splitting equation estimates that mask large Year 12 effect variations across the $j = 1, 2, 3$ categories in the $GZIOPC$. More generally, just because the effect of a splitting equation variable may be zero in a non-generalised model, it does not mean that the effect might not be significantly felt across one or more of the $j = 1, 2...J$ categories in a generalised version. Conversely, it follows that where we observe high levels of significance for a variable - consider the effect of having a degree in the $ZIOPC$, it does not mean that such effects will be felt across all of the $j = 1, 2, 3$ categories.

In general, there appears to be considerable variability in the coefficients corresponding to a given covariate in the $j = 1, 2, 3$ splitting equations in the $GZIOPC$ model. This differential effect is typically more pronounced in the $j = 3$ equation. These findings contrast with those for the single-splitting equation $ZIOPC$ model. In many cases such differences can have non-negligible ramifications with respect to the channels through which different variables impact on smoking behavior, and the associated policy implications.

Table 5 presents a selection of overall partial effects for the correlated model variants evaluated at sample means. Consider the effect of Ln(Income): The $ZIOPC$ model indicates that income has a positive effect on the overall probability of observed zero consumption, operating primarily through the "non-participation" effect. In contrast, the $GZIOPC$ indicates that income has *no* effect overall on the probability of observed zero consumption - whereby social class effects and standard demand analysis effects seemingly work in opposite directions to each other, thereby cancelling each other out. For the $ZIOPC$, income has an effect on all $j = 0, 1, 2$ outcomes, but only for high consumers in the generalized variant.

Own price effects in the $ZIOPC$ model, $Ln(P_T)$, appear large on zero consumption, with a one-unit increase leading to a 14 percentage point (pp) increased chance of this. For the $GZIOPC$ the corresponding figure is over 16.4pp. For high ($j = 3$) consumption levels the comparable figures are $-8.5pp$ and $10.1pp$, respectively. On the other hand, the effect of being married is fairly consistent across the two approaches (indeed, almost identical across $j = 1$ and 2).

To further investigate the consequences of estimating the mis-specified $ZIOP$ and $ZIOPC$ models, Table 6 presents a series of estimated probabilities averaged over all individuals, in which the extent to which non-participatory effects contribute to decision outcomes is quantified (reassuringly, the overall probabilities for all model variants match the observed sample means in the dataset). Such effects are obtained by estimating the probabilities solely associated with the underlying OP component of the respective models. These probabilities effectively "purge", or "net out", any inflation effects. For the correlated versions, the estimated OP parameters were used to estimate these in isolation from the inflation equation(s) - essentially setting the correlation coefficients to zero. Accordingly, we estimate the amount of zero-inflation in the model - $Amount$ (Zero-inflation) - as the difference between the overall predicted probability of zero consumption and the corresponding purged amount. This quantity is then used to calculate the proportion of overall zero consumption that is attributable to the effects of model inflation. Expressed as a percentage, we denote this quantity $Amount(\%)$.

As Table 6 shows, the purged probabilities differ substantially for the $GZIOP$ and $ZIOP$ models, especially for higher consumption levels. Moreover, whilst the $GZIOP$ suggests some nearly 50% of the zero observations can be attributed to zero-inflation, this figure is just over 45% for the $ZIOP$. By comparison, the correlated models both suggest greater levels of zero-inflation, with the generalized variant indicating a relatively higher contribution to overall zero consumption (72% versus 63%). These findings point to the non-generalized models underestimating the degree of overall model inflation.

V.1 $MIOP$ application: *Eurobarometer* survey data

One could envisage this as a sequential process: an individual makes a decision to be informed or not about the EU. Then, conditional on being informed, individuals express their attitude towards EU membership. For the case of the $GMIOP$, one could also envisage individuals as having an underlying propensity for a particular attitude towards EU membership, which could then be tempered by the extent to which they choose to be informed. As in the case of the $MIOP$, these inherent choices would be tempered towards the face-saving inflated option of *neither good nor bad*. Moreover individuals with an inherent propensity for believing EU membership to be a bad thing might need a "bigger push" than those with an inherent propensity for believing EU membership is *a good thing* (or *vice versa*), to move them away from their inherent propensities towards *neither good nor bad*.

Table 7 presents the LM and LR test results. For both tests, the $MIOP$ model is rejected in favour of the $GMIOP$ and $GMIOPC$, and we observe that the $GMIOP$ is rejected in favor of the $GMIOPC$. However, unlike the zero-inflated application in Section **??**, the non-generalized models are not unanimously rejected by both tests in favour of their corresponding generalized variants at conventional (5%) levels of significance. Specifically, the LM test of the $MIOPC$ versus the $GMIOPC$ fails to reject the former at the 5% level, although it is still possible to reject at the 10% level. It is possible that the tests against the $GMIOP$ model are picking-up model mis-specification due to erroneously ignoring the correlation; see Section IV. While this result supports Bagozzi and Mukherjee (2012), the $GMIOPC$ results do suggest the possible presence of an asymmetry with respect to the source of the middle-inflation. As with the $ZIOP$ application, the similarities between the LR and LM test statistics are indicative of a well-behaved log-likelihood function and standard asymptotic theory performing well.

The output equation parameters for the correlated models are presented in Table 9. The $GMIOPC$ model has parameter estimates that are typically similar in sign, significance and magnitude to the $MIOPC$. One noteworthy difference relates to the educational attainment variables, for which the *Educ low-mid* becomes statistically significant in the generalized model.

Table 9 presents the coefficient estimates for the $MIOPC$ and $GZIOPC$ models. Based on the statistical significance of the coefficients in the tempering equations, face-saving effects for the $GMIOPC$ appear to derive overwhelmingly from only one of its tempering equations:

The $j = 2$ equation associated with a propensity to view the EU as *a good thing*. Such a finding is significant: It reveals an asymmetry, where respondents with an underlying propensity to select *a bad thing* in the outcome equation are markedly less inclined to resort to face-saving measures. We also observe that virtually all coefficients in the $j = 2$ equation for the $GMIOPC$ have similar sized coefficients and significance levels to the splitting equation coefficients reported in Bagozzi and Mukherjee (2012), which here are presented as the $MIOPC$. Similar interpretations to the original contribution therefore apply.

The overall partial effects for the $MIOPC$ and $GMIOPC$ models are given in Table 10. The reported effects across all specifications are similar, being comparable in magnitude, direction of effect and significance levels. There are a few exceptions to this. For example, higher education-level effects appear more pronounced in the $GMIOPC$ model for outcomes $j = 1, 2$ whereas the effects of EU-bid knowledge ($j = 1, 2$) are comparatively stronger in the $MIOPC$ model. Overall these results align with the findings in Table 9, where face-saving effects in the $GMIOPC$ model derive from the $j = 2$ tempering equation: There are essentially no significant drivers of face-saving behavior in the $j = 0$ tempering equation, which appears to be redundant. Here, the $GMIOPC$ can be viewed as being characterised by having only a single 'viable' tempering equation. This may account for why the LM test for the $MIOPC$ model - which by construction has a single tempering equation - was not rejected. In this regard, despite there being very little to choose between with respect to the $GZIOPC$ and the $MIOPC$ models, there is a benefit to estimating the former model in that it helps to uncover asymmetries which the single-equation splitting equation of the $MIOPC$ may, by construction, mask.

Model summary probabilities are given in Table 11. Irrespective of model variant, the overall probabilities are virtually identical to the sample proportions. It is useful to pin-down the extent to which face-saving behavior impacts on respondents' choices. The overall probabilities associated with the underlying OP component of each model are again calculated alongside the corresponding probabilities "purged" of inflation effects. As was the case under zero inflation, for the correlated versions the implied independent OP is used in these calculations. Once more, the difference between the overall $j = 1$ probabilities and these purged ones, are denoted *Amount* (Middle-inflation), which can be interpreted as the amount of middle category inflation due to face-saving behavior.

Turning to the *Amount*(%) statistic, of the total responses to the *neither good nor bad* outcome, some 33% of these can be attributed to face-saving responses for the $MIOP$ model, a figure that rises to around 53% for the $GMIOP$ model. These percentages rise for the correlated versions, to 43% and 54%, respectively. As was found with the tobacco consumption application in Section **??**, the extent of overall model inflation in the non-generalized models is underestimated relative to the generalized models. In the case of the present application these differences are sizable, and, based on the results in Tables 9 to 10, are associated with movement away from the $j = 2$ tempering equation.

VI Conclusions and discussion

As these new models collapse to their nested $ZIOP/MIOP$ counterparts under a set of simple parameter restrictions, it is possible to use standard testing paradigms to test for these. We derive the appropriate Lagrange multiplier (LM) tests, which can be used without having to estimate the more general model (*c.f.*, the likelihood ratio (LR) test, for example). Using empirical applications from two key contributions from this literature we find that the tests generally fail dramatically in the case of the $ZIOP$ model, but provide mixed results for the $MIOP$ one. Hence we provide potentially superior alternatives to the established *zero-* and *middle-inflated ordered probit* models; we name these new models, respectively, the *generalized zero-inflated ordered probit* ($GZIOP$) and the *generalized middle-inflated ordered probit* ($GMIOP$). These models have non-negligible implications for model results. This, we argue, may have far-reaching policy implications depending on the application in hand.

This paper proposes generalisations to the increasingly popular $ZIOP$ and $MIOP$ models which allow for tempering from each underlying OP outcome towards the inflated one. We demonstrate that each generalized variant collapses to its associated $ZIOP$ and $MIOP$ form under certain linear parameter restrictions, such that all of the parameter vectors of the now $J-1$ splitting equations are equal. For both the $ZIOP$ and $MIOP$ models, only a single splitting equation requires estimation, whereas the generalized versions each estimate $J-1$ of these. The equality of $\boldsymbol{\beta}_j$ ensures that the model collapses to the $ZIOP/MIOP$. The models are then applied to the data and specifications used in the original contributions of Harris and Zhao (2007) and Bagozzi and Mukherjee (2012). LR and LM tests favor the generalised models in both applications. This finding, we propose, is important, particularly when recalling that Harris and Zhao (2007) and Bagozzi and Mukherjee (2012) claim to have demonstrated the superiority of the $ZIOP$ and $MIOP$ approaches over the OP one. This paper has established that further improvements can be realized by increasing the flexibility of the $ZIOP$ and $MIOP$ models.

In addition to future work applying our proposed generalized models to other empirical settings, our suggested modelling approach raises salient issues which merit further exploration. Consider the cigarette consumption example: it may be the case that tempering is characterised *not* by a simple binary decision - as captured by each of the $J-1$ splitting equations – but a movement down from high levels of tobacco consumption to lower levels, which may, or may not, include zero. Although it is possible to amend the basic set-up of our generalised models to accommodate this kind of behaviour, doing so would represent a move towards a latent class-type set-up that would require even stricter conditions for identification. Most significantly however, amending our proposed generalisations in such a way would yield models that no longer constitute generalisations of the original models proposed by Harris and Zhao (2007) and Bagozzi and Mukherjee (2012), which are the focus of the current contribution. However, as zero- and middle-inflated models have been used effectively to model behavior in a wide array of social, economic, and political settings, the possibility of using these suggested innovations in similar settings represents an interesting avenue for future research.

References

Akcura, E. (2015). Mandatory versus voluntary payment for green electricity. *Ecological Economics 116*(C), 84–94.

Bagozzi, B. E., T. Brawner, B. Mukherjee, and V. Yadav (2014). Regional international organizations and individual immigration attitudes: Results from finite mixture models.

International Interactions 40(3), 350–375.

Bagozzi, B. E., D. W. Hill, W. H. Moore, and B. Mukherjee (2015). Modeling two types of peace: The zero-inflated ordered probit (ZiOP) model in conflict research. *Journal of Conflict Resolution 59*(4), 728–752.

Bagozzi, B. E. and B. Mukherjee (2012). A mixture model for middle category inflation in ordered survey responses. *Political Analysis 20*(3), 369–386.

Becker, G. and K. Murphy (1988). A theory of rational addiction. *Journal of Political Economy 96*(4), 675–700.

Brant, R. (1990). Assessing proportionality in the proportional odds model for ordinal logistic regression. *Biometrics 46*(4), 1171–1178.

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

Cai, T., Y. Xia, and Y. Zhou (2018). Generalized inflated discrete models: A strategy to work with multimodal discrete distributions. *Sociological Methods & Research*.

Carey, S. (2002). Undivided loyalties: Is national identity an obstacle to European integration? *European Union Politics 3*(4), 387–413.

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

Davidson, R. and J. MacKinnon (1987). Implicit alternatives and the local power of test statistics. *Econometrica 55*(6), 1305–1329.

Downward, P., F. Lera-Lopez, and S. Rasciute (2011). The zero-inflated ordered probit approach to modelling sports participation. *Economic Modelling 28*(6), 2469–2477.

Elgün, Ö. and E. R. Tillman (2007). Exposure to European Union Policies and support for membership in the candidate countries. *Political Research Quarterly 60*(3), 391–400.

European Commission (2002a). Candidate countries Eurobarometer 2002.22, September-October 2002.

European Commission (2002b). Eurobarometer 57: EU 15 Report. 50.

Falk, M. and T. Katz-Gerro (2016). Cultural participation in Europe: Can we identify common determinants? *Journal of Cultural Economics 40*(2), 127–162.

Gabel, M. (1998). Public support for European integration: An empirical test of five theories. *Journal of Politics 60*(2), 333–354.

Gillman, M., W. H. Greene, M. N. Harris, and C. Spencer (2013). The Tempered Ordered Probit (TOP) Model with an application to monetary policy. Center for Economic Institutions Working Paper No.2013-04, Institute of Economic Research, Hitotsubashi University.

Greene, W. (2012). *Econometric Analysis 7e* (seventh ed.). New Jersey, USA: Prentice Hall.

Greene, W., M. Harris, and B. Hollingsworth (2015). Inflated responses in measures of self-assessed health. *American Journal of Health Economics 1*(4), 461–493.

Greene, W., M. N. Harris, B. Hollingsworth, and T. A. Weterings (2014). Heterogeneity in ordered choice models: A review with applications to self-assessed health. *Journal of Economic Surveys 28*(1), 109–133.

Greene, W. and D. Hensher (2010). *Modeling Ordered Choices*. Cambridge University Press.

Greene, W. H., J. M. Rose, and D. A. Hensher (2015). *Applied Choice Analysis* (2nd ed.). Cambridge, UK: Cambridge University Press.

Gurmu, S. and G. A. Dagne (2012). Bayesian approach to zero-inflated bivariate ordered probit regression model, with an application to tobacco use. *Journal of Probability and Statistics Article ID 617678*.

Habib, K. M. N., C. Morency, M. T. Islam, and V. Grasset (2012). Modelling users' behaviour of a carsharing program: Application of a joint hazard and zero inflated dynamic ordered probability model. *Transportation Research Part A: Policy and Practice 46*(2), 241–254.

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

Heilbron, D. (1994). Zero-altered and other regression models for count data with added zeros. *Biometrical Journal 36*, 531–547.

Henriksson, R. D. and R. C. Merton (1981). On market timing and investment performance II: Statistical procedures for evaluating forecasting skills. *Journal of Business 54*(4), 513–533.

Jiang, X., B. Huang, X. Yan, R. L. Zaretzki, and S. Richards (2013). Two-vehicle injury severity models based on integration of pavement management and traffic engineering factors. *Traffic Injury Prevention 14*(5), 544–553.

Kim, T.-H., P. Mizen, and T. Chevapatrakul (2008). Forecasting changes in UK interest rates. *Journal of Forecasting 27*(1), 53–74.

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

Maddala, G. S. (1983). *Limited Dependent and Qualitative Variables in Econometrics*. Cambridge, U.K.: Cambridge University Press.

Merton, R. C. (1981). On market timing and investment performance I: An equilibrium theory of value for market forecasts. *Journal of Business 54*(3), 363–406.

Meyerhoefer, C. D. and S. H. Zuvekas (2010). New estimates of the demand for physical and mental health treatment. *Health Economics 19*(3), 297–315.

Miwa, H. (2015). Voters' left-right perception of parties in contemporary Japan: Removing the noise of misunderstanding. *Japanese Journal of Political Science 16(S1)*, 114–137.

Mullahey, J. (1986). Specification and testing of some modified count data models. *Journal of Econometrics 33*(3), 341–365.

Mullahey, J. (1997). Heterogeneity, excess zeros and the structure of count data models. *Journal of Applied Econometrics 12*, 337–350.

Peng, L., C. D. Meyerhoefer, and S. H. Zuvekas (2013). The effect of depression on labor market outcomes. NBER Working Papers 19451, National Bureau of Economic Research, Inc.

Poirier, D. J. (1980). Partial observability in bivariate probit models. *Journal of Econometrics 12*(2), 209–217.

Pudney, S. and M. Shields (2000). Gender, race, pay and promotion in the british nursing profession: Estimationof a generalised ordered probit model. *Journal of Applied Econometrics 15*, 367–399.

Rosa, C. (2009). Forecasting the direction of policy rate changes: The importance of ECB words. *Economic Notes 38*(1-2), 39–66.

Terza, J. V. (1985). Ordered probit: A generalization. *Communications in Statistics - Theory and Methods 14*(1), 1–11.

White, H. (1982). Maximum likelihood estimation of misspecified models. *Econometrica 50*(1), 1–25.

Wilson, P. (2015). The misuse of the Vuong test for non-nested models to test for zero-inflation. *Economics Letters 127*, 51–53.

Zirogiannis, N., J. Alcorn, J. Piepenburg, and J. Rupp (2015). I want in on that: Community-level policies for unconventional gas development in New York. *Agricultural and Resource Economics Review 44*(2), 394–410.

Figures and Tables

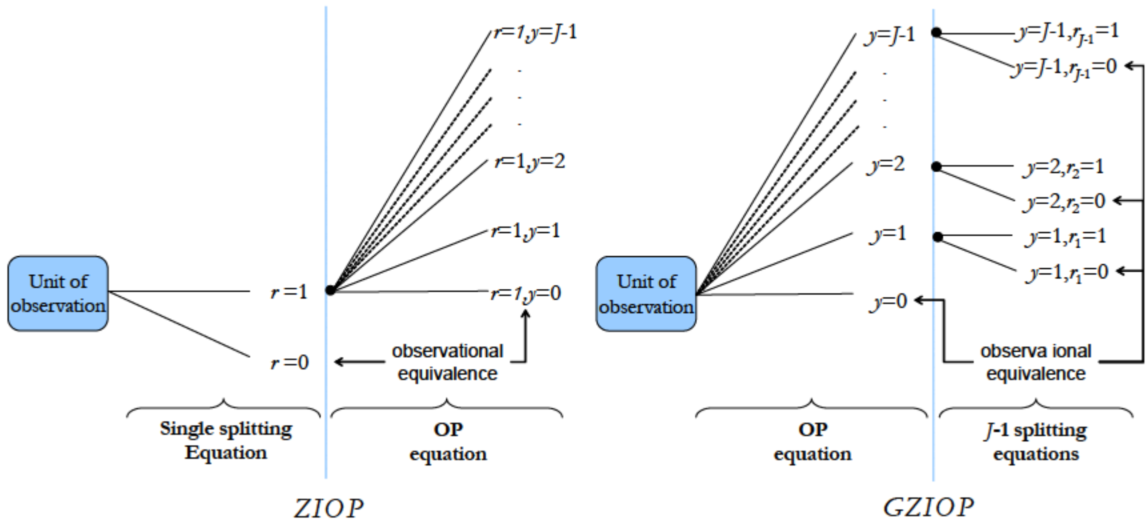

Figure 1: The Zero-Inflated Ordered Probit model ($ZIOP$) and its generalisation ($GZIOP$)

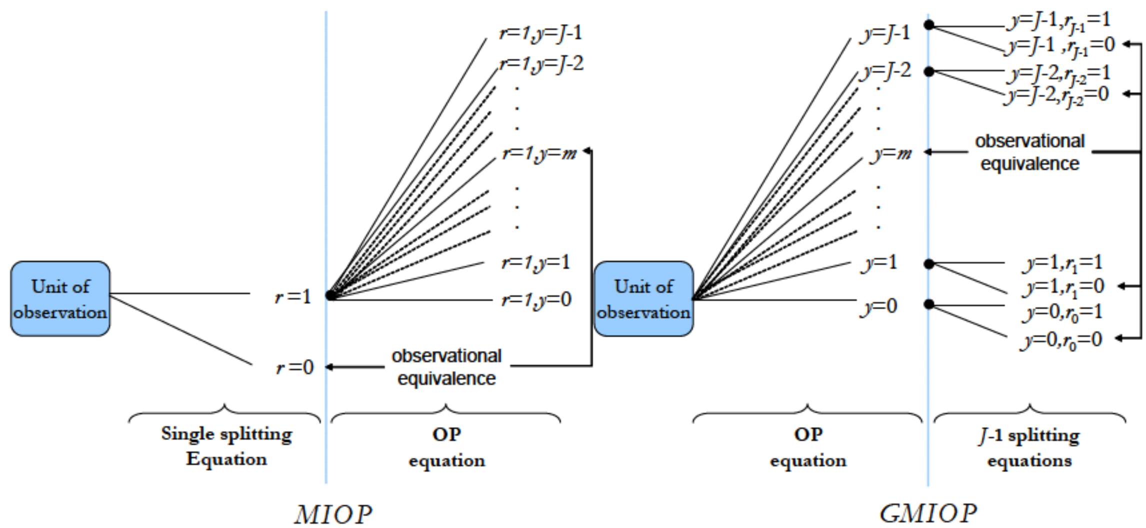

Figure 2: The Middle-Inflated Ordered Probit model ($MIOP$) and its generalisation, ($GMIOP$)

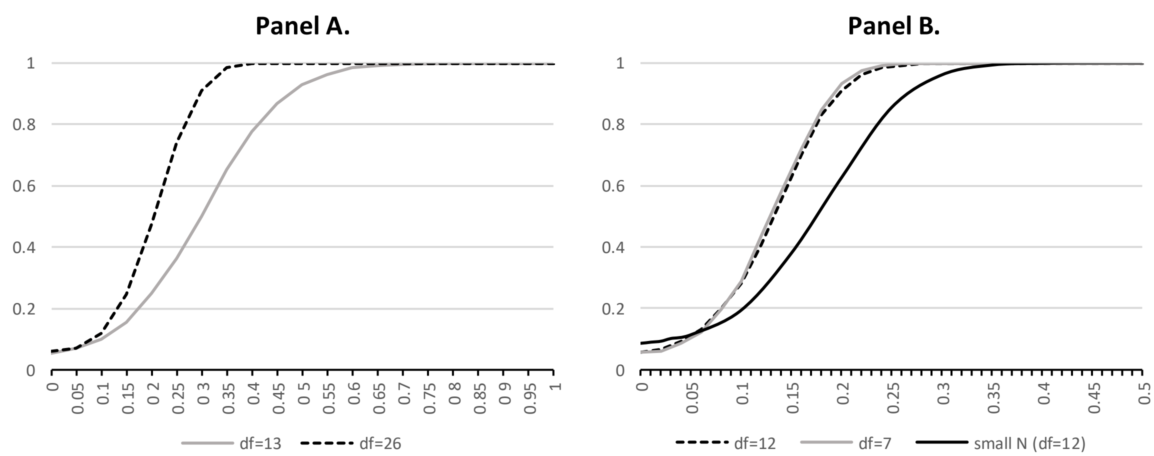

Figure 3: Empirical power curves for the *ZIOP* (Panel A) and *MIOP* (Panel B) models

Table 1: Monte Carlo rejection probabilities

Panel A	Rejection probability			
True model	$GZIOP$ *vs.* $ZIOP$	$GZIOPC$ *vs.* $ZIOPC$	$GZIOPC$ *vs.* $ZIOP$	*Notes*
1.$ZIOP\,(df = 13, 14, 15)$	0.053	0.058	0.056	$J = 3$; $N = 28,813$; $k_x = 16$; $k_z = 18$
2.$ZIOPC\,(df = 13, 14, 15)$	0.381	0.058	0.489	$J = 3$; $N = 28,813$; $k_x = 16$; $k_z = 18$
3.$ZIOP(df = 26, 28, 29)$	0.059	0.061	0.063	$J = 4$; $N = 28,813$; $k_x = 16$; $k_z = 18$
4.$OP\,(df = 13, 14, 15)$	0.252	0.358	0.239	$J = 3$; $N = 28,813$; $k_x = 16$; $k_z = 18$
5.$Parallel\,(df = 13, 14, 15)$	0.141	0.140	0.144	$J = 3$; $N = 28,813$; $k_x = 16$; $k_z = 18$

Panel B	Rejection probability			
True model	$GMIOP$ *vs.* $MIOP$	$GMIOPC$ *vs.* $MIOPC$	$GMIOPC$ *vs.* $MIOP$	*Notes*
6.$MIOP\,(df = 12, 13, 14)$	0.057	0.061	0.062	$J = 3$; $N = 9,113$; $k_x = 12$; $k_z = 16$
7.$MIOPC\,(df = 12, 13, 14)$	0.181	0.061	0.241	$J = 3$; $N = 9,113$; $k_x = 12$; $k_z = 16$
8.$MIOP\,(df = 7, 8, 9)$	0.056	0.055	0.053	$J = 3$; $N = 9,113$; $k_x =?$; $k_z =?$
9.$MIOP(df = 12, 13, 14)$	0.076	0.077	0.0795	$J = 3$; $N =?$; $k_x = 12$; $k_z = 16$
10.$OP\,(df = 12, 13, 14)$	0.484	0.788	0.657	$J = 3$; $N = 9,113$; $k_x = 12$; $k_z = 16$
11.$Parallel\,(df = 12, 13, 14)$	0.253	0.677	0.503	$J = 3$; $N = 9,113$; $k_x = 12$; $k_z = 16$

Table 2: Specification test results: competing $ZIOP$ models

Model	LM Test statistic	df	p-value	LR Test statistic	p-value
$ZIOP$ vs $GZIOP$	194	32	$4.27E-25$	178	$3.56E-22$
$ZIOPC$ vs $GZIOPC$	207	34	$1.68E-26$	202	$9.09E-26$
$ZIOP$ vs $GZIOPC$	221	35	$7.29E-29$	212	$3.33E-27$
$GZIOPvs$ $GZIOPC$	27	3	$5.89E-06$	34	$1.98E-07$

Table 3: Estimates of the output equation parameters for *ZIOPC* and *GZIOPC*

	ZIOPC		*GZIOPC*	
Ln (Income)	0.041	(0.022)*	0.101	(0.023)***
Male	0.027	(0.04)	−0.013	(0.042)
Married	−0.012	(0.057)	0.014	(0.049)
Pre-school	0.028	(0.054)	0.091	(0.063)
Capital	−0.088	(0.035)**	−0.047	(0.037)
Work	−0.227	(0.054)***	−0.26	(0.065)***
Unemployed	0.071	(0.078)	0.118	(0.085)
Study	−0.602	(0.073)***	−0.619	(0.085)***
English-speaking	0.121	(0.073)*	0.114	(0.078)
Degree	−0.759	(0.078)***	−0.728	(0.075)***
Diploma	−0.217	(0.047)***	−0.279	(0.052)***
Year 12	−0.332	(0.049)***	−0.376	(0.052)***
School	−0.437	(0.082)***	−0.435	(0.099)***
$Ln(P_A)$	−1.49	(0.363)***	−1.033	(0.272)***
$Ln(P_M)$	0.028	(0.052)	0.013	(0.037)
$Ln(P_T)$	−0.739	(0.096)***	−0.518	(0.081)***
Age	1.185	(0.055)***	0.957	(0.064)***
Age^2	−1.084	(0.057)***	−0.743	(0.077)***
μ_0	−8.844	(1.753)***	−5.595	(1.377)***
μ_1	−8.577	(1.752)***	−5.335	(1.376)***
μ_2	−7.509	(1.743)***	−3.908	(1.373)***
ρ	−0.424	(0.136)***	–	–
ρ_1	–	–	−0.857	(0.274)***
ρ_2	–	–	−0.647	(0.138)***
ρ_3	–	–	−0.831	(0.178)***
$\ell(\boldsymbol{\theta})$	−21,623		−21,522	

Robust standard errors in parentheses.***, ** and * denote significance at 1%, 5% and 10% level respectively.

Table 4: Estimates of the splitting equation parameters for $ZIOPC$ and $GZIOPC$; tobacco consumption[a]

	$ZIOPC$		$GZIOPC$					
			$j = 1$		$j = 2$		$j = 3$	
Ln(Income)	−0.067	(0.02)***	0.067	(0.054)	−0.075	(0.023)***	−0.181	(0.038)***
Male	0.238	(0.03)***	0.319	(0.098)***	0.06	(0.032)*	0.323	(0.063)***
Married	−0.4	(0.031)***	−0.331	(0.098)***	−0.277	(0.037)***	−0.379	(0.068)***
Pre-school	−0.143	(0.046)***	−0.095	(0.084)	−0.083	(0.049)*	−0.292	(0.092)***
Capital	0.015	(0.029)	0.112	(0.067)*	0.029	(0.031)	−0.027	(0.052)
Work	0.022	(0.04)	0.017	(0.097)	0.066	(0.045)	0.164	(0.081)***
Unemployed	0.15	(0.077)*	0.103	(0.379)	0.179	(0.092)*	−0.094	(0.124)
Study	0.456	(0.125)***	0.759	(0.248)***	0.364	(0.113)***	0.723	(0.232)***
English	0.148	(0.067)**	0.152	(0.121)	0.044	(0.074)	0.067	(0.109)
Degree	−0.203	(0.053)***	0.204	(0.173)	−0.195	(0.098)*	0.3	(0.115)***
Diploma	−0.071	(0.035)**	−0.013	(0.129)	−0.038	(0.048)	0.157	(0.07)**
Year 12	−0.044	(0.042)	0.07	(0.141)	−0.05	(0.059)	0.268	(0.078)***
School	−0.014	(0.206)	0.154	(0.417)	−0.267	(0.216)	0.476	(0.31)
Young female	0.076	(0.038)**	0.008	(0.051)	0.056	(0.027)**	−0.014	(0.049)
Ln(Age)	−1.627	(0.073)***	−1.589	(0.172)***	−1.425	(0.084)***	−2.132	(0.161)***
Constant	6.49	(0.348)***	4.356	(0.702)***	5.599	(0.419)***	9.972	(0.77)***

[a]See notes to Table 3.

Table 5: Selected overall partial effects *ZIOPC* and *GZIOPC*; smoking data[a]

	ZIOPC				*GZIOPC*			
	$j=0$	$j=1$	$j=2$	$j=3$	$j=0$	$j=1$	$j=2$	$j=3$
Ln(Income)	0.013	−0.003	−0.008	−0.001	0.007	0.003	−0.005	−0.005
	(0.005)***	(0.001)***	(0.003)***	(0.002)	(0.005)	(0.002)	(0.004)	(0.003)**
Male	−0.077	0.009	0.044	0.024	−0.069	0.015	0.013	0.042
	(0.007)***	(0.002)***	(0.004)***	(0.003)***	(0.008)***	(0.003)***	(0.006)**	(0.004)***
Married	0.124	−0.016	−0.071	−0.037	0.128	−0.016	−0.063	−0.049
	(0.007)***	(0.002)***	(0.004)***	(0.003)***	(0.008)***	(0.004)***	(0.006)***	(0.004)***
Pre school	0.038	−0.006	−0.022	−0.01	0.035	−0.005	−0.008	−0.022
	(0.009)***	(0.002)***	(0.006)***	(0.004)**	(0.01)***	(0.004)	(0.008)	(0.006)**
Capital	0.012	0.002	−0.005	−0.009	0.007	0.005	0.001	−0.013
	(0.007)*	(0.001)	(0.004)	(0.003)***	(0.007)	(0.003)*	(0.006)	(0.004)***
Work	0.036	0.004	−0.016	−0.024	0.044	0.001	−0.017	−0.029
	(0.009)***	(0.002)	(0.005)***	(0.004)***	(0.011)***	(0.005)	(0.009)*	(0.006)***
Unemployed	−0.059	0.005	0.032	0.021	−0.072	0.005	0.057	0.01
	(0.019)***	(0.004)	(0.011)***	(0.007)***	(0.025)***	(0.018)	(0.018)***	(0.011)
English-speaking	−0.068	0.004	0.036	0.027	−0.063	0.007	0.024	0.032
	(0.015)***	(0.003)	(0.009)***	(0.006)***	(0.015)***	(0.005)	(0.012)**	(0.009)***
Degree	0.205	0.001	−0.102	−0.104	0.226	0.012	−0.135	−0.102
	(0.01)***	(0.003)	(0.006)***	(0.005)***	(0.012)***	(0.005)**	(0.009)***	(0.007)***
Diploma	0.062	0	−0.031	−0.031	0.076	0	−0.043	−0.033
	(0.008)***	(0.002)	(0.005)***	(0.004)***	(0.01)***	(0.005)**	(0.008)	(0.005)***
Year 12	0.076	0.002	−0.037	−0.041	0.091	0.004	−0.058	−0.037
	(0.01)***	(0.002)	(0.006)***	(0.004)***	(0.011)***	(0.006)	(0.009)***	(0.006)***
Young female	−0.023	0.003	0.013	0.007	−0.012	0	0.013	−0.002
	(0.011)**	(0.002)*	(0.006)**	(0.003)**	(0.01)	(0.003)	(0.006)**	(0.007)
$Ln(P_A)$	0.28	0.017	−0.13	−0.168	0.327	0.003	−0.127	−0.202
	(0.068)***	(0.006)***	(0.032)***	(0.041)***	(0.083)***	(0.013)	(0.036)***	(0.051)***
$Ln(P_M)$	−0.005	0	0.002	0.003	−0.004	0	0.002	0.003
	(0.01)	(0.001)	(0.005)	(0.006)	(0.012)	(0)	(0.005)	(0.007)
$Ln(P_T)$	0.139	0.009	−0.064	−0.083	0.164	0.001	−0.064	−0.101
	(0.018)***	(0.003)***	(0.009)***	(0.011)***	(0.023)***	(0.007)	(0.012)***	(0.014)***

[a]See notes to Table 3.

Table 6: Summary probabilities from the *ZIOP* and *GZIOP* models; and *ZIOPC* and *GZIOPC* models[a]

		Independent errors				Correlated errors			
		Overall		Purged		Overall		Purged	
Outcome	Sample	*ZIOP*	*GZIOP*	*ZIOP*	*GZIOP*	*ZIOPC*	*GZIOPC*	*ZIOPC*	*GZIOPC*
$j = 0$	0.7475	0.7474	0.7479	0.4029	0.3831	0.7475	0.7478	0.2787	0.2058
		(0.002)***	(0.002)***	(0.016)***	(0.027)***	(0.002)***	(0.002)***	(0.032)***	(0.024)***
$j = 1$	0.0432	0.0434	0.0432	0.0944	0.1091	0.04340	0.04325	0.0779	0.06252
		(0.001)***	(0.001)***	(0.003)***	(0.015)***	(0.001)***	(0.001)***	(0.006)***	(0.007)***
$j = 2$	0.1448	0.1454	0.1448	0.3398	0.3721	0.1453	0.1448	0.3467	0.4368
		(0.002)***	(0.002)***	(0.010)***	(0.022)***	(0.002)***	(0.002)***	(0.013)***	(0.033)***
$j = 3$	0.0645	0.0639	0.0642	0.1629	0.1357	0.06379	0.06414	0.2967	0.2949
		(0.001)***	(0.001)***	(0.006)***	(0.019)***	(0.001)***	(0.001)***	(0.046)***	(0.046)***
		ZIOP	*GZIOP*			*ZIOPC*	*GZIOPC*		
Amount (Zero-inflation)		0.3444	0.3648			0.4324	0.4597		
		(0.016)***	(0.027)***			(0.030)***	(0.025)***		
Amount(%)		46.09%	48.77%			62.72%	72.48%		

[a]See notes to Table 3.

Table 7: Specification test results: competing $MIOP$ models

Model	LM Test statistic	df	p-value	LR Test statistic	p-value
$MIOP$ *vs* $GMIOP$	32.1	12	0.001	39.3	0.000
$MIOPC$ *vs* $GMIOPC$	20.4	13	0.086	26.2	0.016
$MIOP$ *vs* $GMIOPC$	37.0	14	0.001	46.0	0.000
$GMIOP$ *vs* $GMIOPC$	9.5	2	0.009	6.7	0.035

Table 8: Estimates of the output equation parameters for $MIOPC$ and $GMIOPC$

	$MIOPC$		$GMIOPC$	
Rural	0.028	(0.022)	0.043	(0.029)
Female	0.091	(0.037)	0.126	(0.056)
Age	−0.001	(0.001)	0.001	(0.002)
Student	0.165	(0.085)*	0.229	(0.129)**
Educ high	0.102	(0.066)	−0.106	(0.111)
Educ high-mid	0.059	(0.074)	−0.010	(0.136)
Educ low-mid	0.027	(0.050)	−0.208	(0.094)**
Political trust	0.847	(0.051)***	0.861	(0.059)***
Xenophobia	−0.528	(0.049)***	−0.547	(0.054)***
Discuss politics	−0.029	(0.026)	−0.021	(0.037)
Professional	−0.089	(0.072)	−0.084	(0.072)
Executive	0.115	(0.102)	0.118	(0.102)
Manual	−0.124	(0.045)***	−0.126	(0.046)***
Farmer	−0.043	(0.081)	−0.060	(0.084)
Unemployed	0.108	(0.054)**	0.111	(0.055)**
Income	0.067	(0.007)***	0.070	(0.007)***
μ_0	−0.616	(0.115)***	−0.405	(0.113)***
μ_1	0.138	(0.123)	0.131	(0.110)
ρ	−0.744	(0.162)***	–	–
ρ_1	–	–	0.231	(0.277)
ρ_2	–	–	−0.685	(0.188)***
$\ell(\boldsymbol{\theta})$	$-7,921.7745$		$-7,908.6544$	

Table 9: Estimates of the splitting equation parameters for $MIOPC$ and $GMIOPC$

	$MIOPC$		$GMIOPC$			
			$j = 0$		$j = 2$	
Rural	−0.082	(0.036)**	0.018	(0.087)	−0.111	(0.047)**
Female	−0.332	(0.073)***	0.079	(0.164)	−0.403	(0.096)***
Age	−0.006	(0.002)***	0.006	(0.005)	−0.008	(0.003)***
Student	−0.309	(0.149)**	1.093	(7.817)	−0.421	(0.176)**
Educ high	−0.199	(0.123)*	−1.123	(1.069)	0.094	(0.192)***
Educ high-mid	−0.449	(0.131)***	−0.639	(1.030)	−0.384	(0.179)**
Educ low-mid	−0.434	(0.095)***	−1.200	(1.078)	−0.134	(0.131)
Constant	0.586	(0.207)***	2.033	(1.469)	0.565	(0.252)**
Discuss politics	0.187	(0.048)***	0.104	(0.114)	0.178	(0.059)***
EU-bid knowledge	0.398	(0.091)***	−0.153	(0.291)	0.408	(0.098)***
True EU knowledge	0.126	(0.019)***	−0.021	(0.032)	0.129	(0.022)***
Media	0.044	(0.024)*	−0.139	(0.087)	0.057	(0.025)**

Table 10: Overall partial effects $MIOPC$ and $GMIOPC$

	$MIOPC$			$GMIOPC$		
Common variables	$j=0$	$j=1$	$j=2$	$j=0$	$j=1$	$j=2$
Rural	−0.005	0.012	−0.007	−0.006	0.014	−0.008
	(0.004)	(0.006)**	(0.006)	(0.004)	(0.006)**	(0.007)
Female	−0.015	0.052	−0.036	−0.017	0.054	−0.037
	(0.006)**	(0.01)***	(0.011)***	(0.006)***	(0.01)***	(0.011)***
Age	$8.8e-05$	0.001	−0.001	$1.6e-04$	0.001	−0.001
	$(1.9e-04)$	$(2.8e-04)$***	$(3.6e-04)$***	$(2.1e-04)$	$(3.8e-04)$***	$(4.3e-04)$***
Student	−0.028	0.035	−0.007	−0.025	0.032	−0.007
	(0.014)*	(0.023)	(0.027)	(0.022)	(0.034)	(0.03)
Educ high	−0.017	0.023	−0.006	−0.025	0.042	−0.018
	(0.011)	(0.02)	(0.022)	(0.013)*	(0.024)*	(0.024)
Educ high-mid	−0.01	0.081	−0.071	−0.016	0.099	−0.083
	(0.013)	(0.019)***	(0.023)***	(0.015)	(0.026)***	(0.026)***
Educ low-mid	−0.005	0.083	−0.079	−0.011	0.105	−0.094
	(0.009)	(0.014)***	(0.016)***	(0.011)	(0.018)***	(0.018)***
Discuss	0.005	−0.033	0.028	0.008	−0.037	0.03
	(0.004)	(0.007)***	(0.008)***	(0.005)	(0.008)***	(0.008)***
Outcome equation only variables						
Political trust	−0.142	−0.144	0.285	−0.137	−0.162	0.299
	(0.008)***	(0.011)***	(0.016)***	(0.009)***	(0.017)***	(0.018)***
Xenophobia	0.088	0.09	−0.178	0.089	0.105	−0.194
	(0.009)***	(0.01)***	(0.018)***	(0.01)***	(0.012)***	(0.019)***
Professional	0.015	0.015	−0.03	0.012	0.015	−0.027
	(0.013)	(0.012)	(0.025)	(0.012)	(0.014)	(0.026)
Executive	−0.019	−0.02	0.039	−0.017	−0.02	0.037
	(0.016)	(0.016)	(0.032)	(0.015)	(0.018)	(0.033)
Manual	0.021	0.021	−0.042	0.020	0.024	−0.044
	(0.007)***	(0.008)***	(0.015)***	(0.007)***	(0.009)***	(0.016)***
Farmer	0.007	0.007	−0.015	0.009	0.011	−0.02
	(0.015)	(0.016)	(0.031)	(0.016)	(0.018)	(0.033)
Unemployed	−0.018	−0.018	0.036	−0.017	−0.02	0.037
	(0.009)**	(0.009)**	(0.017)**	(0.009)**	(0.01)**	(0.019)**
Income	−0.011	−0.011	0.023	−0.011	−0.013	0.024
	(0.001)***	(0.001)***	(0.002)***	(0.001)***	(0.002)***	(0.002)***
Splitting equation only variables						
EU-bid knowledge	$4.6e-05$	−0.081	0.081	0.006	−0.071	0.065
	$(1.3e-04)$	(0.017)***	(0.017)***	(0.013)	(0.018)***	(0.016)***
True EU knowledge	$1.5e-05$	−0.025	0.025	−0.001	−0.022	0.024
	$(3.9e-05)$	(0.003)***	(0.003)***	(0.002)	(0.003)***	(0.003)***
Media	$5.2e-06$	−0.009	0.009	−0.005	−0.005	0.011
	$(1.5e-05)$	(0.005)*	(0.005)*	(0.003)	(0.005)	(0.005)**

Table 11: Summary probabilities from the $MIOP$ and $GMIOP$ models; EU data

		Independent errors				Correlated errors			
		Overall		Purged		Overall		Purged	
Outcome	Sample	$MIOP$	$GMIOP$	$MIOP$	$GMIOP$	$MIOPC$	$GMIOPC$	$MIOPC$	$GMIOPC$
$j = 0$	0.108	0.108	0.108	0.128	0.189	0.108	0.108	0.109	0.145
		(0.003)***	(0.003)***	(0.004)***	(0.028)***	(0.003)***	(0.003)***	(0.003)***	(0.015)***
$j = 1$	0.331	0.331	0.331	0.222	0.155	0.331	0.331	0.190	0.153
		(0.005)***	(0.005)***	(0.015)***	(0.040)***	(0.005)***	(0.005)***	(0.018)***	(0.021)***
$j = 2$	0.561	0.561	0.561	0.650	0.656	0.561	0.561	0.701	0.702
		(0.005)***	(0.005)***	(0.013)***	(0.022)***	(0.005)***	(0.005)***	(0.018)***	(0.021)***
		$MIOP$	$GMIOP$			$MIOPC$	$GMIOPC$		
Amount (Middle-inflation)		0.109	0.176			0.141	0.176		
		(0.014)***	(0.040)***			(0.018)***	(0.021)***		
Amount(%)		32.83%	53.14%			42.59%	53.67%		

Online Appendix

A Lagrange multiplier (LM) test of the $ZIOPC$ model(s)

A highly appealing specification test for the $ZIOPC$ models versus their generalized alternatives is the LM test, as this only requires estimation of the simpler nested models. This involves evaluation of the score vector of the more general model evaluated at parameter values under the null (*i.e.*, at $ZIOPC$ ones). Here we present the score for the case of correlated errors. As noted above, the $GZIOPC$ model of equation (14) can form the basis of an LM test of the $GZIOPC$ versus the $ZIOP$ and $ZIOPC$ models. The former is tested using $H_0 : \boldsymbol{\beta}_j = \boldsymbol{\beta}$ and $\rho_j = 0, \forall j$ and the latter by $H_0 : \boldsymbol{\beta}_j = \boldsymbol{\beta}$ and $\rho_j = \rho, \forall j$.

Using the matrix version of the general result for bivariate normal distributions that

$$\frac{\partial \Phi_2 (a, b; \rho)}{\partial a} = \phi (a) \Phi \left(\frac{b - \rho a}{\sqrt{1 - \rho^2}} \right), \tag{A.1}$$

where $\Phi_2 (a, b; \rho)$ denotes the standardized bivariate normal cumulative density function (CDF), we can define the following quantities of interest. First, define $\Phi^+_{b,j}$ as

$$\Phi^+_{b,j} = \Phi \left(\frac{(\mu_j - \mathbf{z}'\boldsymbol{\gamma}) - \rho_j (-\mathbf{x}'\boldsymbol{\beta}_j)}{\sqrt{1 - \rho_j^2}} \right) - \Phi \left(\frac{(\mu_{j-1} - \mathbf{z}'\boldsymbol{\gamma}) - \rho_j (-\mathbf{x}'\boldsymbol{\beta}_j)}{\sqrt{1 - \rho_j^2}} \right) \tag{A.2}$$

for $j = 1, \ldots, J - 2$ and

$$\Phi^+_{b,J-1} = \Phi \left(\frac{(\mathbf{z}'\boldsymbol{\gamma} - \mu_{J-2}) - \rho_{J-1} (\mathbf{x}'\boldsymbol{\beta}_{J-1})}{\sqrt{1 - \rho_{J-1}^2}} \right) \tag{A.3}$$

for $j = J - 1$; and then $\Phi^-_{b,j}$ as

$$\Phi^-_{b,j} = \Phi \left(\frac{(\mu_j - \mathbf{z}'\boldsymbol{\gamma}) + \rho_j (\mathbf{x}'\boldsymbol{\beta}_j)}{\sqrt{1 - \rho_j^2}} \right) - \Phi \left(\frac{(\mu_{j-1} - \mathbf{z}'\boldsymbol{\gamma}) + \rho_j (\mathbf{x}'\boldsymbol{\beta}_j)}{\sqrt{1 - \rho_j^2}} \right) \tag{A.4}$$

for $j = 1, \dots, J-2$ and

$$\Phi^{-}_{b,J-1} = \Phi\left(\frac{\left(\mathbf{z}'\boldsymbol{\gamma} - \mu_{J-2}\right) + \rho_{J-1}\left(-\mathbf{x}'\boldsymbol{\beta}_{J-1}\right)}{\sqrt{1-\rho^2_{J-1}}}\right) \tag{A.5}$$

for $j = J-1$. Labelling the probabilities of the $GZIOPC$ model P^{GZIOPC}, and using expressions (A.2) to (A.5), the score with respect to the elements of $\boldsymbol{\beta}$ can be written as

$$\frac{\partial \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\beta}_j} = \left[\begin{array}{c} \sum\limits_{y_i=0} -\mathbf{x}\phi\left(-\mathbf{x}'\boldsymbol{\beta}_j\right)\Phi^{+}_{b,j} + \sum\limits_{y_i=0} -\mathbf{x}\phi\left(-\mathbf{x}'\boldsymbol{\beta}_{J-1}\right)\Phi^{-}_{b,J-1} + \\ \sum\limits_{y_i>0}^{y_i=J-2} \mathbf{x}\phi\left(\mathbf{x}'\boldsymbol{\beta}_j\right)\Phi^{-}_{b,j} + \sum\limits_{y_i=J-1} \mathbf{x}\phi\left(\mathbf{x}'\boldsymbol{\beta}_{J-1}\right)\Phi^{+}_{b,J-1} \end{array}\right] \div P^{GZIOPC}_{j=y_i} \tag{A.6}$$

for $\boldsymbol{\beta}_j$, $j = 1, \dots, J-1$. Similarly, defining $\phi^{+}_{a,j}$ as

$$\phi^{+}_{a,j} = \phi\left(\mu_j - \mathbf{z}'\boldsymbol{\gamma}\right)\Phi\left(\frac{\left(-\mathbf{x}'\boldsymbol{\beta}_j\right) - \rho_j\left(\mu_j - \mathbf{z}'\boldsymbol{\gamma}\right)}{\sqrt{1-\rho^2_j}}\right) - \phi\left(\mu_{j-1} - \mathbf{z}'\boldsymbol{\gamma}\right)\Phi\left(\frac{\left(-\mathbf{x}'\boldsymbol{\beta}_j\right) - \rho_j\left(\mu_{j-1} - \mathbf{z}'\boldsymbol{\gamma}\right)}{\sqrt{1-\rho^2_j}}\right) \tag{A.7}$$

for $j = 1, \dots, J-2$ and

$$\phi^{+}_{a,J-1} = \phi\left(\mathbf{z}'\boldsymbol{\gamma} - \mu_{J-2}\right)\Phi\left(\frac{\mathbf{x}'\boldsymbol{\beta}_{J-1} - \rho_{J-1}\left(\mathbf{z}'\boldsymbol{\gamma} - \mu_{J-2}\right)}{\sqrt{1-\rho^2_{J-1}}}\right) \tag{A.8}$$

for $j = J-1$; and then $\phi^{-}_{a,j}$ as

$$\phi^{-}_{a,j} = \phi\left(\mu_j - \mathbf{z}'\boldsymbol{\gamma}\right)\Phi\left(\frac{\mathbf{x}'\boldsymbol{\beta}_j + \rho_j\left(\mu_j - \mathbf{z}'\boldsymbol{\gamma}\right)}{\sqrt{1-\rho^2_j}}\right) - \phi\left(\mu_{j-1} - \mathbf{z}'\boldsymbol{\gamma}\right)\Phi\left(\frac{\mathbf{x}'\boldsymbol{\beta}_j + \rho_j\left(\mu_{j-1} - \mathbf{z}'\boldsymbol{\gamma}\right)}{\sqrt{1-\rho^2_j}}\right) \tag{A.9}$$

for $j = 1, \dots, J-2$ and

$$\phi^{-}_{a,J-1} = \phi\left(\mathbf{z}'\boldsymbol{\gamma} - \mu_{J-2}\right)\Phi\left(\frac{\left(-\mathbf{x}'\boldsymbol{\beta}_{J-1}\right) + \rho_{J-1}\left(\mathbf{z}'\boldsymbol{\gamma} - \mu_{J-2}\right)}{\sqrt{1-\rho^2_{J-1}}}\right) \tag{A.10}$$

for $j = J-1$ permits us to write the score with respect to $\boldsymbol{\gamma}$ as

$$
\frac{\partial \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\gamma}} = \left[\begin{array}{c} \sum_{y_i=0} \left[-\mathbf{z}\phi\left(\mu_0 - \mathbf{z}'\boldsymbol{\gamma}\right) + \sum_{j=1}^{J-2} -\mathbf{z}\phi_{a,,j}^{+} + \mathbf{z}\phi_{a,J-1}^{-} \right] + \\ \sum_{y_i>0}^{y_i=J-2} \left[-\mathbf{z}\phi_{a,,j}^{-} \right] \times 1\left[y_i = j \right] + \\ \sum_{y_i=J-1} \mathbf{z}\phi_{a,J-1}^{+} \end{array} \right] \div P_{j=y_i}^{GZIOPC}. \tag{A.11}
$$

As stated in Section II the required ordering of the boundary parameters is ensured by specifying them as

$$
\mu_j = \mu_{j-1} + \exp\left(\xi_j\right), \quad j = 1, 2, ...J-2 \tag{A.12}
$$

where μ_0 is freely estimated (Greene and Hensher 2010). Accordingly, the associated scores with respect to $\mu_0, \xi_1, \xi_2, \ldots \xi_{J-2}$ are given by,

$$
\begin{aligned}
\frac{\partial \ell(\boldsymbol{\theta})}{\partial \mu_0} &= \left[\sum_{y_i=0} \phi\left(\mu_0 - \mathbf{z}'\boldsymbol{\gamma}\right) + \phi_{a,,j}^{+} - \phi_{a,J-1}^{-} \right] \div P_{j=0}^{GZIOPC} \\
&+ \left[\sum_{y_i>0}^{y_i=J-2} \left[\phi_{a,,j}^{-} \right] \times 1\left[y_i = j \right] \right] \div P_{j=y_i}^{GZIOPC} \\
&- \left[\sum_{y_i=J-1} \phi_{a,J-1}^{+} \right] \div P_{j=J-1}^{GZIOPC}
\end{aligned} \tag{A.13}
$$

$$
\begin{aligned}
\frac{\partial \ell(\boldsymbol{\theta})}{\partial \xi_1} &= \left[\sum_{y_i=0} \left\{ \begin{array}{c} \sum_{j=1} \exp\left(\xi_1\right)\phi\left(\mu_1 - \mathbf{z}'\boldsymbol{\gamma}\right)\Phi\left(\frac{\mathbf{x}'\boldsymbol{\beta}_1 + \rho_j\left(\mu_1 - \mathbf{z}'\boldsymbol{\gamma}\right)}{\sqrt{1-\rho_1^2}}\right) + \\ \sum_{j=2}^{J-2} \exp\left(\xi_1\right)\phi_{a,,j}^{+} - \exp\left(\xi_1\right)\phi_{a,J-1}^{-} \end{array} \right\} \right] \div P_{j=0}^{GZIOPC} \\
&+ \left[\sum_{y_i=1} \exp\left(\xi_1\right)\phi\left(\mu_1 - \mathbf{z}'\boldsymbol{\gamma}\right)\Phi\left(\frac{\mathbf{x}'\boldsymbol{\beta}_j + \rho_j\left(\mu_1 - \mathbf{z}'\boldsymbol{\gamma}\right)}{\sqrt{1-\rho_1^2}}\right) \right] \div P_{j=1}^{GZIOPC} \\
&+ \left[\sum_{y_i>1}^{y_i=J-2} \exp\left(\xi_1\right)\phi_{a,,j}^{-} \right] \div P_{j=y}^{GZIOPC} + \left[\sum_{y_i=J-1} -\exp\left(\xi_1\right)\phi_{a,J-1}^{+} \right] \div P_{j=J}^{GZIOPC}
\end{aligned} \tag{A.14}
$$

$$\frac{\partial \ell(\boldsymbol{\theta})}{\partial \xi_2} = \left[\sum_{y_i=0} \left\{ \begin{array}{c} \sum_{j=2} \exp(\xi_2)\,\phi(\mu_2 - \mathbf{z}'\boldsymbol{\gamma})\,\Phi\left(\frac{\mathbf{x}'\boldsymbol{\beta}_2+\rho_2(\mu_2-\mathbf{z}'\boldsymbol{\gamma})}{\sqrt{1-\rho_2^2}}\right) \\ +\sum_{j=2}^{J-2} \exp(\xi_2)\,\phi^{+}_{a,,j} - \exp(\xi_2)\,\phi^{-}_{a,J-1} \end{array} \right\}\right] \div P^{GZIOPC}_{j=y_i} \tag{A.15}$$

$$+ \left[\sum_{y_i=2} \exp(\xi_2)\,\phi(\mu_2 - \mathbf{z}'\boldsymbol{\gamma})\,\Phi\left(\frac{\mathbf{x}'\boldsymbol{\beta}_2 + \rho_2(\mu_2 - \mathbf{z}'\boldsymbol{\gamma})}{\sqrt{1-\rho_2^2}}\right)\right] \div P^{GZIOPC}_{j=2}$$

$$+ \left[\sum_{y_i>2}^{y_i=J-2} \exp(\xi_2)\,\phi^{-}_{a,,j}\right] \div P^{GZIOPC}_{j=y} + \left[\sum_{y_i=J-1} -\exp(\xi_2)\,\phi^{+}_{a,J-1}\right] \div P^{GZIOPC}_{j=J-1}$$

$$\vdots$$

$$\frac{\partial \ell(\boldsymbol{\theta})}{\partial \xi_{J-1}} = \left[\sum_{y_i=J-1} -\exp(\xi_{J-1})\,\phi^{+}_{a,J-1}\right] \div P^{GZIOPC}_{j=J-1} \tag{A.16}$$

Finally, the derivatives of the elements of $\boldsymbol{\rho}$ $\forall j = 1, 2, ...J-2$ are given by

$$\frac{\partial \ell(\boldsymbol{\theta})}{\partial \rho_j} = \left[\sum_{y_i=0} \left[\phi_2\left(\mu_j - \mathbf{z}'\boldsymbol{\gamma}, -\mathbf{x}'\boldsymbol{\beta}_j; \rho_1\right) - \phi_2\left(\mu_{j-1} - \mathbf{z}'\boldsymbol{\gamma}, -\mathbf{x}'\boldsymbol{\beta}_j; \rho_j\right)\right]\right] \div P^{GZIOPC}_{j=0} \tag{A.17}$$

$$+ \left[\sum_{y_i=j} -\left[\phi_2\left(\mu_j - \mathbf{z}'\boldsymbol{\gamma}, \mathbf{x}'\boldsymbol{\beta}_j; -\rho_1\right) - \phi_2\left(\mu_{j-1} - \mathbf{z}'\boldsymbol{\gamma}, \mathbf{x}'\boldsymbol{\beta}_j; -\rho_j\right)\right]\right] \div P^{GZIOPC}_{j=y_i}$$

whereas for ρ_{J-1} we have

$$\frac{\partial \ell(\boldsymbol{\theta})}{\partial \rho_{J-1}} = \left[\sum_{y_i=0} -\phi_2\left(\mathbf{z}'\boldsymbol{\gamma} - \mu_{J-2}, -\mathbf{x}'\boldsymbol{\beta}_{J-1}; -\rho_{J-1}\right)\right] \div P^{GZIOPC}_{j=0} \tag{A.18}$$

$$+ \left[\sum_{J-1} \phi_2\left(\mathbf{z}'\boldsymbol{\gamma} - \mu_{J-2}, \mathbf{x}'\boldsymbol{\beta}_{J-1}; \rho_{J-1}\right)\right] \div P^{GZIOPC}_{j=J-1}$$

In estimation we ensure a well-defined ρ_j, $j = 1, \ldots, J-1$, such that for $-1 < \rho_j < 1$ we use the hyperbolic tangent function transformation, $\rho_j = \tanh \rho_j^*$, where ρ_j^* is freely estimated. If such a transformation is followed, then the above derivatives for $\boldsymbol{\rho}$ need to be multiplied by $\partial \tanh \rho_j^* / \rho_j^* = 1 - \tanh^2 \rho_j^*$. Using all of the above quantities, the LM statistic is given by

$$LM^{ZIOP}_{correlated} = (\boldsymbol{\nabla\beta}, \boldsymbol{\nabla\gamma}, \nabla\mu_0, \boldsymbol{\nabla\xi}, \boldsymbol{\nabla\rho})' \left[\mathbf{I}\left(\hat{\theta}_R\right)\right]^{-1} (\boldsymbol{\nabla\beta}, \boldsymbol{\nabla\gamma}, \nabla\mu_0, \boldsymbol{\nabla\xi}, \boldsymbol{\nabla\rho}) \tag{A.19}$$

which is evaluated at the relevant parameter restrictions as defined by the appropriate null hypothesis. Under H_0, $LM^{ZIOP}_{correlated} \sim \chi^2_q$, where q is the number of parameter restrictions under the appropriate H_0. If the alternative model is the uncorrelated generalised version, one would omit the relevant partition of the score vector $(\boldsymbol{\nabla\rho})$. As is common practice, we use the outer product of gradients (OPGs) to estimate the inverse of the variance of the score vector, $\left[\mathbf{I}\left(\hat{\theta}_R\right)\right]^{-1}$ (Greene 2012).

B The Generalized Middle-Inflated Ordered Probit Model ($GMIOP$)

As noted, two contributions - Bagozzi and Mukherjee (2012) and Brooks et al. (2012) - independently proposed the *middle-inflated ordered probit* ($MIOP$) model to allow for inflation in any arbitrary category. Bagozzi and Mukherjee (2012) were the first to use the term 'middle-inflated'. Brooks et al. (2012) refer to their model merely as an 'inflated ordered probit'. We use the former nomenclature, and suggest that the term *inflated ordered probit* (IOP) model may be better viewed as encompassing both the $ZIOP$ and the $MIOP$ model classes. In keeping with Section II, we develop the $GMIOP$ framework in the context of J outcomes. Whilst in both original $MIOP$ contributions the empirical analysis is restricted to three outcomes, the model developed in this section naturally also applies to instances where $J > 3$.

Consider again an OP model as a starting point, where each individual i has an unobserved underlying propensity

$$y^* = \mathbf{z}'\boldsymbol{\gamma} + \eta \tag{B.1}$$

such that y^* translates into observed outcomes y via the usual OP form. We now assume that a middle category $y \in \{1, 2...J-2\}$ is associated with an "excess of observations" and/or they can be hypothesised to have arisen from two distinct data generating processes. Label this category m. Again, define r^* as an underlying latent variable that represents an overall propensity to choose the inflated category m over any other, which translates into an "observed" binary outcome. Let this be a linear (in the parameters, $\boldsymbol{\beta}$) function of observed characteristics $\mathbf{x}_i$ and a standard normal random error term ε

$$r^* = \mathbf{x}'\boldsymbol{\beta} + \varepsilon. \tag{B.2}$$

Again, a two-regime scenario arises where for observations in regime $r = 0$, the inflated outcome is observed; but for those in $r = 1$, any of the possible outcomes in the choice set $j = \{0, 1, 2...J-2, J-1\}$ - *including the inflated category* m – can be observed. Accordingly, overall $MIOP$ probabilities under the assumption of independent errors are given by

$$\Pr(y_i) = \begin{cases} \Pr\left(y = 0 \left| \mathbf{x}_i, \mathbf{z}_i \right.\right) = \Phi\left(\mathbf{x}_i'\boldsymbol{\beta}\right) \times \Phi\left(\mu_0 - \mathbf{z}_i'\boldsymbol{\gamma}\right) \\ \Pr\left(y = j \left| \mathbf{x}_i, \mathbf{z}_i \right.\right) = \Phi\left(\mathbf{x}_i'\boldsymbol{\beta}\right) \times \left[\Phi\left(\mu_1 - \mathbf{z}_i'\boldsymbol{\gamma}\right) - \Phi\left(\mu_0 - \mathbf{z}_i'\boldsymbol{\gamma}\right)\right] + M \\ \Pr\left(y = J-1 \left| \mathbf{x}_i, \mathbf{z}_i \right.\right) = \Phi\left(\mathbf{x}_i'\boldsymbol{\beta}\right) \times \left[1 - \Phi\left(\mu_{J-2} - \mathbf{z}_i'\boldsymbol{\gamma}\right)\right] \end{cases} \tag{B.3}$$

whereas for correlated errors we have that

$$\Pr(y_i) = \begin{cases} \Pr\left(y = 0 \left| \mathbf{x}_i, \mathbf{z}_i \right.\right) = \Phi_2\left(\mu_0 - \mathbf{z}_i'\boldsymbol{\gamma}, \mathbf{x}_\mathbf{i}'\boldsymbol{\beta}; -\rho\right) \\ \Pr\left(y = j \left| \mathbf{x}_i, \mathbf{z}_i \right.\right) = \Phi_2\left(\mu_1 - \mathbf{z}_i'\boldsymbol{\gamma}, \mathbf{x}_i'\boldsymbol{\beta}; -\rho\right) - \Phi_2\left(\mu_0 - \mathbf{z}_i'\boldsymbol{\gamma}, \mathbf{x}_i'\boldsymbol{\beta}; -\rho\right) + M \\ \Pr\left(y = J-1 \left| \mathbf{x}_i, \mathbf{z}_i \right.\right) = \Phi_2\left(\mathbf{x}_i'\boldsymbol{\beta}, \mathbf{z}_i'\boldsymbol{\gamma} - \mu_{J-2}; \rho\right) \end{cases} \tag{B.4}$$

where $M = 0$ if $y \neq m$ and

$$M = \Phi\left(-\mathbf{x}_i'\boldsymbol{\beta}\right)$$

iff $y = m$. This implies that for the model with independent errors,

$$\Pr\left(y = m \left| \mathbf{x}_i, \mathbf{z}_i \right.\right) = \Phi\left(\mathbf{x}_i'\boldsymbol{\beta}\right) \times \left[\Phi\left(\mu_1 - \mathbf{z}_i'\boldsymbol{\gamma}\right) - \Phi\left(\mu_0 - \mathbf{z}_i'\boldsymbol{\gamma}\right)\right] + 1 - \Phi\left(\mathbf{x}_i'\boldsymbol{\beta}\right) \tag{B.5}$$

and for the case of correlated errors

$$\Pr\left(y = m \left| \mathbf{x}_i, \mathbf{z}_i \right.\right) = \Phi_2\left(\mu_1 - \mathbf{z}_i'\boldsymbol{\gamma}, \mathbf{x}_\mathbf{i}'\boldsymbol{\beta}; -\rho\right) - \Phi_2\left(\mu_0 - \mathbf{z}_i'\boldsymbol{\gamma}, \mathbf{x}_\mathbf{i}'\boldsymbol{\beta}; -\rho\right) + 1 - \Phi\left(\mathbf{x}_i'\boldsymbol{\beta}\right) \tag{B.6}$$

In such a way, the probability of a single, middle category has again been inflated. Diagrammatically, this is depicted on the left hand side of Figure 2, where we again emphasize that m can assume any of the values in the set $j \in \{1, 2 ... J-2\}$. As in the case of the $ZIOP$, we reiterate that the model is estimated simultaneously.

Following logic analogous to that used in Section II, we generalize the inflation process for m. This is illustrated in the right-hand panel of Figure 2: For any given propensity towards a given category $j \neq m$ in the outcome equation, there is a movement towards an inflated middle category, m.

Let these propensities towards m be determined, respectively, by $J-1$ splitting equations - each corresponding to a non-inflated category, namely

$$r^*_{j \neq m} = \mathbf{x}'\boldsymbol{\beta}_j + \varepsilon_j \tag{B.7}$$

such that the probability of a movement towards the inflated middle category, m, is given by

$$\Pr(r_{j \neq m} = 0) = \Phi\left(-\mathbf{x}'\boldsymbol{\beta}_j\right) \tag{B.8}$$

Under independence and standard normality, the overall probabilities for non-inflated outcomes are

$$\Pr(y_i) = \begin{cases} \Pr\left(y = 0 \left| \mathbf{x}_i, \mathbf{z}_i \right.\right) = \Phi\left(\mu_0 - \mathbf{z}_i'\boldsymbol{\gamma}\right) \times \Phi\left(\mathbf{x}_i'\boldsymbol{\beta}_0\right) \\ \Pr\left(y = \widetilde{j} \left| \mathbf{x}_i, \mathbf{z}_i \right.\right) = \left[\Phi\left(\mu_{\widetilde{j}} - \mathbf{z}_i'\boldsymbol{\gamma}\right) - \Phi\left(\mu_{\widetilde{j}-1} - \mathbf{z}_i'\boldsymbol{\gamma}\right)\right] \times \Phi\left(\mathbf{x}_i'\boldsymbol{\beta}_{\widetilde{j}}\right) \\ \Pr\left(y = m \left| \mathbf{x}_i, \mathbf{z}_i \right.\right) = \begin{cases} \left[\Phi\left(\mu_m - \mathbf{z}_i'\boldsymbol{\gamma}\right) - \Phi\left(\mu_{m-1} - \mathbf{z}_i'\boldsymbol{\gamma}\right)\right] \\ + \underbrace{\Phi\left(\mu_0 - \mathbf{z}_i'\boldsymbol{\gamma}\right) \times \Phi\left(-\mathbf{x}_i'\boldsymbol{\beta}_0\right)}_{a} \\ \underbrace{+\sum_{\widetilde{j}} \left[\Phi\left(\mu_{\widetilde{j}} - \mathbf{z}_i'\boldsymbol{\gamma}\right) - \Phi\left(\mu_{\widetilde{j}-1} - \mathbf{z}_i'\boldsymbol{\gamma}\right)\right] \times \Phi\left(-\mathbf{x}_i'\boldsymbol{\beta}_{\widetilde{j}}\right)}_{b} \\ + \underbrace{\left[1 - \Phi\left(\mu_{J-2} - \mathbf{z}_i'\boldsymbol{\gamma}\right)\right] \times \Phi\left(-\mathbf{x}_i'\boldsymbol{\beta}_{J-1}\right)}_{c} \end{cases} \\ \Pr\left(y = J-1 \left| \mathbf{x}_i, \mathbf{z}_i \right.\right) = \left[1 - \Phi\left(\mu_{J-2} - \mathbf{z}_i'\boldsymbol{\gamma}\right)\right] \times \Phi\left(\mathbf{x}_i'\boldsymbol{\beta}_{J-1}\right) \end{cases} . \tag{B.9}$$

where $\widetilde{j}$ includes all middle categories excluding the inflated one. Inflation in category m

is still allowed for by the additional terms of a, b and c in equation (B.9). Expression (B.9) is the *generalized middle-inflated ordered probit* ($GMIOP$) model. Relaxing the assumption of independent errors leads to the correlated *generalized middle-inflated ordered probit* ($GMIOPC$) model of

$$\Pr(y_i) = \begin{cases} \Pr\left(y = 0\,|\mathbf{x}_i, \mathbf{z}_i\right) = \Phi_2(\mu_0 - \mathbf{z}_i'\boldsymbol{\gamma}, \mathbf{x}_i'\boldsymbol{\beta}_0; -\rho_0) \\ \Pr\left(y = \widetilde{j}\,|\mathbf{x}_i, \mathbf{z}_i\right) = \Phi_2\left(\mu_{\widetilde{j}} - \mathbf{z}_i'\boldsymbol{\gamma}, \mathbf{x}_i'\boldsymbol{\beta}_{\widetilde{j}}; -\rho_{\widetilde{j}}\right) - \Phi_2\left(\mu_{\widetilde{j}-1} - \mathbf{z}_i'\boldsymbol{\gamma}, \mathbf{x}_i'\boldsymbol{\beta}_{\widetilde{j}}; -\rho_{\widetilde{j}}\right) \\ \Pr\left(y = m\,|\mathbf{x}_i, \mathbf{z}_i\right) = \begin{cases} \left[\Phi\left(\mu_m - \mathbf{z}_i'\boldsymbol{\gamma}\right) - \Phi\left(\mu_{m-1} - \mathbf{z}_i'\boldsymbol{\gamma}\right)\right] \\ +\underbrace{\Phi_2\left(\mu_0 - \mathbf{z}_i'\boldsymbol{\gamma}, -\mathbf{x}_i'\boldsymbol{\beta}_0; \rho_0\right)}_{a} \\ \underbrace{+\sum_{\widetilde{j}=1}^{J-2}\left[\begin{array}{c} \Phi_2\left(\mu_{\widetilde{j}} - \mathbf{z}_i'\boldsymbol{\gamma}, -\mathbf{x}_i'\boldsymbol{\beta}_{\widetilde{j}}; \rho_{\widetilde{j}}\right) \\ -\Phi_2\left(\mu_{\widetilde{j}-1} - \mathbf{z}_i'\boldsymbol{\gamma}, -\mathbf{x}_i'\boldsymbol{\beta}_{\widetilde{j}}; \rho_{\widetilde{j}}\right) \end{array}\right]}_{b} \\ +\underbrace{\Phi_2\left(\mathbf{z}_i'\boldsymbol{\gamma} - \mu_{J-2}, -\mathbf{x}_i'\boldsymbol{\beta}_{J-1}; -\rho_{J-1}\right)}_{c} \end{cases} \\ \Pr\left(y = J-1\,|\mathbf{x}_i, \mathbf{z}_i\right) = \Phi_2\left(\mathbf{z}_i'\boldsymbol{\gamma} - \mu_{J-2}, \mathbf{x}_i'\boldsymbol{\beta}_{J-1}; \rho_{J-1}\right) \end{cases} \tag{B.10}$$

As in (B.9), inflation arises in category m due to the additional terms of a, b and c. The model is characterized by $J-1$ correlation coefficients $\rho_j\ \forall\ j \neq m$, which correspond to all categories apart from the middle-inflated one. Specifically, these encompass the categories at each end of the choice spectrum, for which we have ρ_0 and ρ_{J-1}; and all of the $\widetilde{j}$ non-inflated middle categories, namely $\rho_{\widetilde{j}}\ \forall\ \widetilde{j}$.

As in Section (II), consider imposing the linear set of restrictions that $\boldsymbol{\beta}_0 = \boldsymbol{\beta}_{\widetilde{j}} = \boldsymbol{\beta}_{J-1} = \boldsymbol{\beta}$ and $\rho_0 = \rho_{\widetilde{j}} = \rho_{J-1} = \rho$ on equation (B.10); setting $\boldsymbol{\beta}_0 = \boldsymbol{\beta}_{\widetilde{j}} = \boldsymbol{\beta}_{J-1} = \boldsymbol{\beta}$ implies that the tempering propensities for all of the $J-1$ splitting equations are identical. This yields

$$
\Pr(y_i) = \begin{cases}
\Pr\left(y=0\,|\mathbf{x}_i,\mathbf{z}_i\right) = \Phi_2(\mu_0 - \mathbf{z}_i'\boldsymbol{\gamma}, \mathbf{x}_i'\boldsymbol{\beta}; -\rho) \\
\Pr\left(y=\widetilde{j}\,|\mathbf{x}_i,\mathbf{z}_i\right) = \Phi_2\left(\mu_{\widetilde{j}} - \mathbf{z}_i'\boldsymbol{\gamma}, \mathbf{x}_i'\boldsymbol{\beta}; -\rho\right) - \Phi_2\left(\mu_{\widetilde{j}-1} - \mathbf{z}_i'\boldsymbol{\gamma}, \mathbf{x}_i'\boldsymbol{\beta}; -\rho\right) \\
\Pr\left(y=m\,|\mathbf{x}_i,\mathbf{z}_i\right) = \begin{cases}
\left[\Phi\left(\mu_m - \mathbf{z}_i'\boldsymbol{\gamma}\right) - \Phi\left(\mu_{m-1} - \mathbf{z}_i'\boldsymbol{\gamma}\right)\right] \\
+\underbrace{\Phi_2\left(\mu_0 - \mathbf{z}_i'\boldsymbol{\gamma}, -\mathbf{x}_i'\boldsymbol{\beta}; \rho\right)}_{a} \\
+\underbrace{\sum_{\widetilde{j}}\left[\begin{array}{c} \Phi_2\left(\mu_{\widetilde{j}} - \mathbf{z}_i'\boldsymbol{\gamma}, -\mathbf{x}_i'\boldsymbol{\beta}; \rho\right) \\ -\Phi_2\left(\mu_{\widetilde{j}-1} - \mathbf{z}_i'\boldsymbol{\gamma}, -\mathbf{x}_i'\boldsymbol{\beta}; \rho\right) \end{array}\right]}_{b} \\
+\underbrace{\Phi_2\left(\mathbf{z}_i'\boldsymbol{\gamma} - \mu_{J-2}, -\mathbf{x}_i'\boldsymbol{\beta}; -\rho\right)}_{c}
\end{cases} \\
\Pr\left(y=J-1\,|\mathbf{x}_i,\mathbf{z}_i\right) = \Phi_2\left(\mathbf{z}_i'\boldsymbol{\gamma} - \mu_{J-2}, \mathbf{x}_i'\boldsymbol{\beta}; \rho\right)
\end{cases} \tag{B.11}
$$

where the expressions for $\Pr\left(y=0\,|\mathbf{z},\mathbf{x}\right)$, $\Pr\left(y=\widetilde{j}\,|\mathbf{z},\mathbf{x}\right)$ and $\Pr\left(y=J-1\,|\mathbf{z},\mathbf{x}\right)$ immediately collapse to those in the $MIOPC$, given in expression (B.4). We stress here that the $\Pr\left(y=\widetilde{j}\,|\mathbf{z},\mathbf{x}\right)$ are equivalent to cases of $\Pr\left(y=j\,|\mathbf{z},\mathbf{x}\right)$ $\forall j=1,2,...J-2$ where $M=0$. Using (B.4), subtracting these terms from one yields

$$
\Pr\left(y=m\right) = \Phi_2\left(\mu_m - \mathbf{z}_i'\boldsymbol{\gamma}, \mathbf{x}'\boldsymbol{\beta}; -\rho\right) - \Phi_2\left(\mu_{m-1} - \mathbf{z}_i'\boldsymbol{\gamma}, \mathbf{x}'\boldsymbol{\beta}; -\rho\right) + 1 - \Phi\left(\mathbf{x}'\boldsymbol{\beta}\right) \tag{B.12}
$$

That is, the $GMIOPC$ collapses to and therefore nests the $MIOPC$. Further, setting $\rho=0$ in (B.12) yields probabilities that are identical to the $MIOP$ probabilities in expression (B.6), *viz.*

$$
\Pr\left(y=m\,|\mathbf{x}_i,\mathbf{z}_i\right) = \Phi\left(\mathbf{x}_i'\boldsymbol{\beta}\right) \times \left[\Phi\left(\mu_1 - \mathbf{z}_i'\boldsymbol{\gamma}\right) - \Phi\left(\mu_0 - \mathbf{z}_i'\boldsymbol{\gamma}\right)\right] + 1 - \Phi\left(\mathbf{x}_i'\boldsymbol{\beta}\right) \tag{B.13}
$$

The $GMIOPC$ also collapses to the $MIOP$, albeit under the alternative set of parameter restrictions $\boldsymbol{\beta}_1 = \boldsymbol{\beta}_2 = \boldsymbol{\beta}_3 ... = \boldsymbol{\beta}_{J-1}$ and $\rho_j = 0\ \forall\ j = 0, \widetilde{j}, J-1$. Applying the latter set of restrictions implicitly reduces the $GMIOPC$ model to the $GMIOP$. Equivalently, imposing the parameter restrictions $\boldsymbol{\beta}_0 = \boldsymbol{\beta}_{\widetilde{j}} = \boldsymbol{\beta}_{J-1}$ on the $GMIOP$ model leads it to nest the $MIOP$ resulting in $GMIOP$ probabilities that are identical to the $MIOP$ probabilities in (B.3). Diagrammatically, this means that the model depicted on the right of Figure 2 nests the model depicted on the left. Testing the parameter restrictions associated with these model variants entails testing (i) the more flexible functional form of the $GMIOPC$ model versus the simpler nested forms of the $MIOPC$ and $MIOP$ models and (ii) the $GMIOP$ versus the $MIOP$ model.

Diagrammatically, this is depicted on the left hand side of Figure 2, where we again emphasize that m can assume any of the values in the set $j \in \{1, 2...J-2\}$. As in the case of the $ZIOP$, we reiterate that the model is estimated simultaneously. Diagrammatically, this means that the model depicted on the right of Figure 2 nests the model depicted on the left. Testing the parameter restrictions associated with these model variants entails testing (i) the more flexible functional form of the $GMIOPC$ model versus the simpler nested forms of the $MIOPC$ and $MIOP$ models and (ii) the $GMIOP$ versus the $MIOP$ model. As with the $GZIOP$ model, the $GMIOP$ is still an inflated ordered probit model. The ordering of outcomes is still preserved, and middle-inflation arises due to $J-1$ distinct DGPs, as opposed to just one.

As with the $GZIOP$ model, the $GMIOP$ is still an inflated ordered probit model. The ordering of outcomes is still preserved, and middle-inflation arises due to $J-1$ distinct DGPs, as opposed to just one. Further, as with the $GZIOP$, a straightforward test of hypotheses can be undertaken using a standard LR test or LM tests. As the $GMIOPC$ score vector closely follows that for the $GZIOPC$, it is presented in Appendix B.1.

B.1 $MIOP$ score vector

To aid notation and to coincide with our empirical application in section V.1, we assume that $J = 3$, and label the ordered choices as $j = 0, 1, 2$ (negative, indifferent, positive), where $j = 1$ is the hypothesized inflated category. Here the explicit form of the $GMIOPC$

probabilities will be

$$\Pr\left(y_i\right)=\begin{cases} 0 & = & \Phi_2\left(\mu_0-\mathbf{z}'\boldsymbol{\gamma},\mathbf{x}'\boldsymbol{\beta}_0;-\rho_0\right) \\ 1 & = & \begin{cases} \Phi\left(\mu_0+\exp\left(\xi_1\right)-\mathbf{z}'\boldsymbol{\gamma}\right)-\Phi\left(\mu_0-\mathbf{z}'\boldsymbol{\gamma}\right) \\ +\Phi_2\left(\mu_0-\mathbf{z}'\boldsymbol{\gamma},-\mathbf{x}'\boldsymbol{\beta}_0;\rho_0\right) \\ +\Phi_2\left(\mathbf{z}'\boldsymbol{\gamma}-\mu_0-\exp\left(\xi_1\right),-\mathbf{x}'\boldsymbol{\beta}_2;-\rho_2\right) \end{cases} \\ 2 & = & \Phi_2\left(\mathbf{z}'\boldsymbol{\gamma}-\mu_0-\exp\left(\xi_1\right),\mathbf{x}'\boldsymbol{\beta}_2;\rho_2\right) \end{cases}. \tag{B.14}$$

The score with respect to $\boldsymbol{\gamma}$ ($\boldsymbol{\nabla\gamma}$) will be

$$\frac{\partial\ell\left(\boldsymbol{\theta}\right)}{\partial\boldsymbol{\gamma}}=\left[\begin{array}{c} \sum\limits_{y_i=0}\left[-\mathbf{z}\phi\left(\mu_0-\mathbf{z}'\boldsymbol{\gamma}\right)\times\Phi\left(\frac{\mathbf{x}'\boldsymbol{\beta}_0+\rho_0\left(\mu_0-\mathbf{z}'\boldsymbol{\gamma}\right)}{\sqrt{1-\rho_0^2}}\right)\right]+ \\ \sum\limits_{y_i=1}\left[\begin{array}{c} \left(-\mathbf{z}\phi\left(\mu_0+\exp\left(\xi_1\right)-\mathbf{z}'\boldsymbol{\gamma}\right)+\mathbf{z}\phi\left(\mu_0-\mathbf{z}'\boldsymbol{\gamma}\right)\right)+ \\ -\mathbf{z}\phi\left(\mu_0-\mathbf{z}'\boldsymbol{\gamma}\right)\times\Phi\left(\frac{\left(-\mathbf{x}'\boldsymbol{\beta}_0\right)-\rho_0\left(\mu_0-\mathbf{z}'\boldsymbol{\gamma}\right)}{\sqrt{1-\rho_0^2}}\right)+ \\ \mathbf{z}\phi\left(\mathbf{z}'\boldsymbol{\gamma}-\mu_0-\exp\left(\xi_1\right)\right)\times\Phi\left(\frac{\left(-\mathbf{x}'\boldsymbol{\beta}_2\right)+\rho_2\left(\mathbf{z}'\boldsymbol{\gamma}-\mu_0-\exp\left(\xi_1\right)\right)}{\sqrt{1-\rho_2^2}}\right) \end{array}\right]+ \\ \sum\limits_{y_i=J-1}\left[\mathbf{z}\phi\left(\mathbf{z}'\boldsymbol{\gamma}-\mu_0-\exp\left(\xi_1\right)\right)\times\Phi\left(\frac{\mathbf{x}'\boldsymbol{\beta}_2-\rho_2\left(\mathbf{z}'\boldsymbol{\gamma}-\mu_0-\exp\left(\xi_1\right)\right)}{\sqrt{1-\rho_2^2}}\right)\right] \end{array}\right]\div P_{j=y_i}^{GMIOPC}. \tag{B.15}$$

And for the boundary parameters, $\nabla\mu_0,\nabla\xi_1$

$$\nabla\mu_0=\left[\begin{array}{c} \sum\limits_{y_i=0}\left[\phi\left(\mu_0-\mathbf{z}'\boldsymbol{\gamma}\right)\times\Phi\left(\frac{\mathbf{x}'\boldsymbol{\beta}_0+\rho_0\left(\mu_0-\mathbf{z}'\boldsymbol{\gamma}\right)}{\sqrt{1-\rho_0^2}}\right)\right]+ \\ \sum\limits_{y_i=1}\left[\begin{array}{c} \phi\left(\mu_0+\exp\left(\xi_1\right)-\mathbf{z}'\boldsymbol{\gamma}\right)-\phi\left(\mu_0-\mathbf{z}'\boldsymbol{\gamma}\right)+ \\ \phi\left(\mu_0-\mathbf{z}'\boldsymbol{\gamma}\right)\times\Phi\left(\frac{\left(-\mathbf{x}'\boldsymbol{\beta}_0\right)-\rho_0\left(\mu_0-\mathbf{z}'\boldsymbol{\gamma}\right)}{\sqrt{1-\rho_0^2}}\right)+ \\ \phi\left(\mathbf{z}'\boldsymbol{\gamma}-\mu_0-\exp\left(\xi_1\right)\right)\times\Phi\left(\frac{\left(-\mathbf{x}'\boldsymbol{\beta}_2\right)+\rho_2\left(\mathbf{z}'\boldsymbol{\gamma}-\mu_0-\exp\left(\xi_1\right)\right)}{\sqrt{1-\rho_2^2}}\right) \end{array}\right]+ \\ \sum\limits_{y_i=J-1}\left[\phi\left(\mathbf{z}'\boldsymbol{\gamma}-\mu_0-\exp\left(\xi_1\right)\right)\times\Phi\left(\frac{\mathbf{x}'\boldsymbol{\beta}_2-\rho_2\left(\mathbf{z}'\boldsymbol{\gamma}-\mu_0-\exp\left(\xi_1\right)\right)}{\sqrt{1-\rho_2^2}}\right)\right] \end{array}\right]\div P_{j=y_i}^{GMIOPC}. \tag{B.16}$$

and

$$\nabla\xi_1 = \left[\begin{array}{c} \sum\limits_{y_i=1} \left[\begin{array}{c} \exp(\xi_1)\,\phi(\mu_0+\exp(\xi_1)-\mathbf{z}'\boldsymbol{\gamma})+ \\ (-\exp(\xi_1))\,\phi(\mathbf{z}'\boldsymbol{\gamma}-\mu_0-\exp(\xi_1))\times\Phi\left(\frac{(-\mathbf{x}'\boldsymbol{\beta}_2)+\rho_2(\mathbf{z}'\boldsymbol{\gamma}-\mu_0-\exp(\xi_1))}{\sqrt{1-\rho_2^2}}\right) \end{array} \right] + \\ \sum\limits_{y_i=J-1} \left[(-\exp(\xi_1))\,\phi(\mathbf{z}'\boldsymbol{\gamma}-\mu_0-\exp(\xi_1))\times\Phi\left(\frac{\mathbf{x}'\boldsymbol{\beta}_2-\rho_2(\mathbf{z}'\boldsymbol{\gamma}-\mu_0-\exp(\xi_1))}{\sqrt{1-\rho_2^2}}\right)\right] \end{array} \right] \div P_{j=y_i}^{GMIOPC}. \tag{B.17}$$

The score with respect to $\boldsymbol{\beta}_0$ ($\boldsymbol{\nabla\beta_0}$) and $\boldsymbol{\beta}_2$ ($\boldsymbol{\nabla\beta_2}$) will respectively be

$$\boldsymbol{\nabla\beta_0} = \left[\begin{array}{c} \sum\limits_{y_i=0} \left[\mathbf{x}\phi(\mathbf{x}'\boldsymbol{\beta}_0)\times\Phi\left(\frac{(\mu_0-\mathbf{z}'\boldsymbol{\gamma})+\rho_0(\mathbf{x}'\boldsymbol{\beta}_0)}{\sqrt{1-\rho_0^2}}\right)\right] + \\ \sum\limits_{y_i=1} \left[-\mathbf{x}\phi(-\mathbf{x}'\boldsymbol{\beta}_0)\times\Phi\left(\frac{(\mu_0-\mathbf{z}'\boldsymbol{\gamma})-\rho_0(-\mathbf{x}'\boldsymbol{\beta}_0)}{\sqrt{1-\rho_0^2}}\right)+\right] \end{array} \right] \div P_{j=y_i}^{GMIOPC} \tag{B.18}$$

and

$$\boldsymbol{\nabla\beta_1} = \left[\begin{array}{c} \sum\limits_{y_i=1} \left[-\mathbf{x}\phi(-\mathbf{x}'\boldsymbol{\beta}_2)\times\Phi\left(\frac{(\mathbf{z}'\boldsymbol{\gamma}-\mu_1)+\rho_2(-\mathbf{x}'\boldsymbol{\beta}_2)}{\sqrt{1-\rho_2^2}}\right)\right] \\ \sum\limits_{y_i=J-1} \left[\mathbf{x}\phi(-\mathbf{x}'\boldsymbol{\beta}_2)\times\Phi\left(\frac{(\mathbf{z}'\boldsymbol{\gamma}-\mu_1)-\rho_2(\mathbf{x}'\boldsymbol{\beta}_2)}{\sqrt{1-\rho_2^2}}\right)\right] \end{array} \right] \div P_{j=y_i}^{GMIOPC} \tag{B.19}$$

Deriving the score vector for the LM test is again, straightforward. Define: P_j^{OP} as the standard OP probabilities implied by equation (3); P_j^{MIOP} as those for the $MIOP$ in expression (B.3); P_j^{GMIOP} as those for the $GMIOP$ model of expression (B.9); and finally, P^0 as the tempering equation probability of $\Phi(\mathbf{x}'\boldsymbol{\beta}_0)$, P^{J-1} as the tempering equation probability of $\Phi(\mathbf{x}'\boldsymbol{\beta}_{J-1})$, and $P^{\widetilde{j}}$ as the tempering equation probabilities of $\Phi\left(\mathbf{x}'\boldsymbol{\beta}_{\widetilde{j}}\right)$, where $\widetilde{j}$ captures all middle outcomes that are *not* inflated.

As with the case of the $GZIOP$, we maintain the necessary ordering of the boundary parameters by specifying them as $\mu_j = \mu_{j-1} + \exp(\xi_j)$, where μ_0 is freely estimated, and where for ease of notation, we assume that $J = 3$. The elements of the score vector are given by

$$\frac{\partial \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\gamma}} = \begin{bmatrix} \sum\limits_{y_i=0} -\mathbf{z}\widetilde{\mu}_{-1}P^0 \\ +\sum\limits_{y_i=1}\left(-\mathbf{z}\widetilde{\mu}_0 - \mathbf{z}\widetilde{\mu}_{-1}(1-P^0) + \mathbf{z}\widetilde{\mu}_1(1-P^2)\right) \\ +\sum\limits_{y_i=2}\mathbf{z}\widetilde{\mu}_1 P^2 \end{bmatrix} \div P_{j=y_i}^{GMIOP} \tag{B.20}$$

$$\begin{aligned} \frac{\partial \ell(\boldsymbol{\theta})}{\partial \mu_0} &= \left[\sum_{y_i=0}\widetilde{\mu}_{-1}P^0\right] \div P_{j=0}^{GMIOP} + \\ &\left[\sum_{y_i=1}\widetilde{\mu}_0 + \widetilde{\mu}_0\left(1-P^0\right) - \widetilde{\mu}_1\left(1-P^2\right)\right] \div P_{j=0}^{GMIOP} + \\ &\left[\sum_{y_i=1} -\widetilde{\mu}_1 P^2\right] \div P_{j=2}^{GMIOP} \end{aligned} \tag{B.21}$$

$$\begin{aligned} \frac{\partial \ell(\boldsymbol{\theta})}{\partial \xi} &= \left[\sum_{y_i=1}\exp(\xi)\widetilde{\mu}_1 - \exp(\xi)\widetilde{\mu}_1\left(1-P^2\right)\right] \div P_{j=1}^{GMIOP} + \\ &\left[\sum_{y_i=2} -\exp(\xi)\widetilde{\mu}_1 P^2\right] \div P_{j=2}^{GMIOP} \end{aligned} \tag{B.22}$$

$$\begin{aligned} \frac{\partial \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\beta}_0} &= \left[\sum_{y_i=0}\mathbf{x}\phi\left(\mathbf{x}'\boldsymbol{\beta}_0\right)P_{j=0}^{OP}\right] \div P_{j=0}^{GMIOP} + \\ &\left[\sum_{y_i=1} -\mathbf{x}\phi\left(\mathbf{x}'\boldsymbol{\beta}_0\right) \times P_{j=0}^{OP}\right] \div P_{j=1}^{GMIOP} \end{aligned} \tag{B.23}$$

$$\begin{aligned} \frac{\partial \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\beta}_2} &= \left[\sum_{y_i=1} -\mathbf{x}\phi\left(\mathbf{x}'\boldsymbol{\beta}_0\right)P_{j=2}^{OP}\right] \div P_{j=1}^{GMIOP} + \\ &\left[\sum_{y_i=2}\mathbf{x}\phi\left(\mathbf{x}'\boldsymbol{\beta}_2\right) \times P_{j=2}^{OP}\right] \div P_{j=2}^{GMIOP} \end{aligned} \tag{B.24}$$

As with the $GZIOP$, in estimation we ensure a well-defined ρ_j, $j = 1, \ldots, J-1$, such that $\rho_j \in (-1, 1)$ where we use the hyperbolic tangent function transformation, $\rho_j = \tanh \rho_j^*$, where ρ_j^* is freely estimated. Following such a transformation the above derivatives for $\boldsymbol{\rho}$ require multiplication by $\partial \tanh \rho_j^* / \rho_j^* = 1 - \tanh^2 \rho_j^*$. Using all of the above quantities, the LM statistic is given by

$$LM_{correlated}^{MIOP} = (\boldsymbol{\nabla\beta}, \boldsymbol{\nabla\gamma}, \nabla\mu_0, \boldsymbol{\nabla\xi}, \boldsymbol{\nabla\rho})' \left[\mathbf{I}\left(\hat{\theta}_R\right)\right]^{-1} (\boldsymbol{\nabla\beta}, \boldsymbol{\nabla\gamma}, \nabla\mu_0, \boldsymbol{\nabla\xi}, \boldsymbol{\nabla\rho}) \qquad \text{(B.25)}$$

which is evaluated at the relevant parameter restrictions as defined by the appropriate null hypothesis. Under H_0, $LM_{correlated}^{MIOP} \sim \chi_q^2$, where q is the number of parameter restrictions under the appropriate H_0. Again, $\left[\mathbf{I}\left(\hat{\theta}_R\right)\right]^{-1}$ is estimated as before, and one would remove $\boldsymbol{\nabla\rho}$ where the uncorrelated generalised variant is the alternative model.

C Model coherency and identification

Accordingly, it is important to ascertain whether the proposed discrete choice model generalisations are, what is often termed in the literature, "coherent" or "logically consistent" (see for instance Maddala 1983, Ch.5). This entails ensuring that the model's parameters are uniquely identified and the associated probabilities are well-defined and sum to unity. For expositional clarity we demonstrate this using the $GZIOP$ model with uncorrelated errors, noting that extensions to the $GMIOP$ and with correlated errors can also be demonstrated. Lastly, we demonstrate that the *generalised ordered probit* ('GOP') models of Terza (1985) and Pudney and Shields (2000), which as arguably characterised by incoherency (Greene, Harris, Hollingsworth, and Weterings 2014), neither nest, nor are nested by the GOP.

C.1 Unique identification

Ensuring that the parameters are uniquely identified is akin to ensuring that the model cannot simultaneously generate more than one value of y simultaneously. In this respect, if one can simulate the dependent variable, then this suggests that the model is, indeed, coherent (implying that the parameters are uniquely identified). Here, consider simulating along the lines of the sequencing suggested in the model descriptions above:

1. Consider the $\widetilde{y}^* = \mathbf{z}'\boldsymbol{\gamma} + u$ equation. With known $\boldsymbol{\gamma}$ and boundary parameters $\boldsymbol{\mu}$, "first stage" $\widetilde{y}$ values can be straightforwardly simulated by simply simulating u from an assumed $N(0,1)$ distribution by the usual relationship between the simulated $\widetilde{y}^*$ and $\boldsymbol{\mu}$.

2. This uniquely places an individual in one, and only one, of the $j = 0, ..., J-1$ $\widetilde{y}$ outcomes.

3. Individuals in the $\widetilde{y} = 0$ category are allocated to observed $y = 0$.

4. For individuals falling uniquely into the $\widetilde{y} = 1$ category one can simulate their observed outcome by consideration of $r^*_{j=1} = \mathbf{x}'\boldsymbol{\beta}_{j=1} + \varepsilon_{j=1}$:

    (a) With known $\boldsymbol{\beta}_{j=1}$ it is straightforward to simulate $r^*_{j=1}$ by simulating $\varepsilon_{j=1}$, again from an assumed $N(0,1)$ distribution.

    (b) The position of the simulated index $r^*_{j=1}$ with respect to 0, uniquely simulates $r_{j=1}$; $r_{j=1} = 1\left(r^*_{j=1} > 0\right)$.

    (c) With $\widetilde{y} = 1$ and $r_{j=1}$ in hand, $y_{j=1}$ is uniquely determined by the observability criteria defined above, here explicitly, $y_{j=1} = \widetilde{y}r_{j=1}$.

5. Similarly, for all individuals uniquely falling into the $\widetilde{y} = 2$ category, observed $y_{j=2} = \widetilde{y}r_{j=2}$, with $r_{j=2}$ being determined as above by $1\left(r^*_{j=2} > 0\right)$.

6. And so on, for all other $j \geq 3$.

7. Equivalently, y_0 can be also be simulated as

$$1\left(\mathbf{z}'\boldsymbol{\gamma} + u < \mu_0\right) + \sum_{j=1}^{J-1} 1\left(\left[\mu_{j-1} < \mathbf{z}'\boldsymbol{\gamma} + u < \mu_j\right]\left[\mathbf{x}'\boldsymbol{\beta}_j + \varepsilon_j < 0\right]\right)$$

   with the usual convention of $\mu_{J-1} = \infty$. As all components of this are mutually exclusive, this uniquely maps on to a single value for all observed y (similar expressions apply for the remaining j).

Thus although there is nothing in the model to prevent the "existence" of several of the r_j variables "being equal to one", apart from the one corresponding to the uniquely determined $\widetilde{y}_j$ value, all others are redundant. The reason for this, follows from the more general latent class models (of which, our approach, is a special case, as described in the text), in which individuals can only be in any one particular class (at a given point in time), therefore behaviours in all other classes simply do not exist. In this way, our approach mirrors that of the standard latent class approach.

C.2 Well-defined probabilities

We now explore if our proposed models have well-defined probabilities. Significantly, it is straightforward to show that model probabilities all lie within the unit circle and sum to unity. For ease of exposition consider the $GZIOP$ with $J = 3$. Here we have that

$$\begin{aligned}
P_0 &= \Phi\left(\mu_0 - \mathbf{z}'\boldsymbol{\gamma}\right) + \left[\Phi\left(\mu_1 - \mathbf{z}'\boldsymbol{\gamma}\right) - \Phi\left(\mu_0 - \mathbf{z}'\boldsymbol{\gamma}\right)\right]\Phi\left(-\mathbf{x}'\boldsymbol{\beta}_1\right) + \Phi\left(\mathbf{z}'\boldsymbol{\gamma} - \mu_1\right)\Phi\left(-\mathbf{x}'\boldsymbol{\beta}_2\right) \\
P_1 &= \left[\Phi\left(\mu_1 - \mathbf{z}'\boldsymbol{\gamma}\right) - \Phi\left(\mu_0 - \mathbf{z}'\boldsymbol{\gamma}\right)\right]\Phi\left(\mathbf{x}'\boldsymbol{\beta}_1\right) \\
P_2 &= \Phi\left(\mathbf{z}'\boldsymbol{\gamma} - \mu_1\right)\Phi\left(\mathbf{x}'\boldsymbol{\beta}_2\right)
\end{aligned}$$

So $\sum_j P_j$ is

$$\begin{aligned}
&= \Phi\left(\mu_0 - \mathbf{z}'\boldsymbol{\gamma}\right) + \left[\Phi\left(\mu_1 - \mathbf{z}'\boldsymbol{\gamma}\right) - \Phi\left(\mu_0 - \mathbf{z}'\boldsymbol{\gamma}\right)\right]\Phi\left(-\mathbf{x}'\boldsymbol{\beta}_1\right) + \Phi\left(\mathbf{z}'\boldsymbol{\gamma} - \mu_1\right)\Phi\left(-\mathbf{x}'\boldsymbol{\beta}_2\right) \\
&\quad + \left[\Phi\left(\mu_1 - \mathbf{z}'\boldsymbol{\gamma}\right) - \Phi\left(\mu_0 - \mathbf{z}'\boldsymbol{\gamma}\right)\right]\Phi\left(\mathbf{x}'\boldsymbol{\beta}_1\right) \\
&\quad + \Phi\left(\mathbf{z}'\boldsymbol{\gamma} - \mu_1\right)\Phi\left(\mathbf{x}'\boldsymbol{\beta}_2\right) \\
&= \Phi\left(\mu_0 - \mathbf{z}'\boldsymbol{\gamma}\right) \\
&\quad + \left[\Phi\left(\mu_1 - \mathbf{z}'\boldsymbol{\gamma}\right) - \Phi\left(\mu_0 - \mathbf{z}'\boldsymbol{\gamma}\right)\right]\left[1 - \Phi\left(\mathbf{x}'\boldsymbol{\beta}_1\right)\right] + \left[\Phi\left(\mu_1 - \mathbf{z}'\boldsymbol{\gamma}\right) - \Phi\left(\mu_0 - \mathbf{z}'\boldsymbol{\gamma}\right)\right]\Phi\left(\mathbf{x}'\boldsymbol{\beta}_1\right) \\
&\quad + \Phi\left(\mathbf{z}'\boldsymbol{\gamma} - \mu_1\right)\left[1 - \Phi\left(\mathbf{x}'\boldsymbol{\beta}_2\right)\right] + \Phi\left(\mathbf{z}'\boldsymbol{\gamma} - \mu_1\right)\Phi\left(\mathbf{x}'\boldsymbol{\beta}_2\right) \\
&= \Phi\left(\mu_0 - \mathbf{z}'\boldsymbol{\gamma}\right) + \\
&\quad \left[\Phi\left(\mu_1 - \mathbf{z}'\boldsymbol{\gamma}\right) - \Phi\left(\mu_0 - \mathbf{z}'\boldsymbol{\gamma}\right)\right] - \left[\Phi\left(\mu_1 - \mathbf{z}'\boldsymbol{\gamma}\right) - \Phi\left(\mu_0 - \mathbf{z}'\boldsymbol{\gamma}\right)\right]\Phi\left(\mathbf{x}'\boldsymbol{\beta}_1\right) \\
&\qquad + \left[\Phi\left(\mu_1 - \mathbf{z}'\boldsymbol{\gamma}\right) - \Phi\left(\mu_0 - \mathbf{z}'\boldsymbol{\gamma}\right)\right]\Phi\left(\mathbf{x}'\boldsymbol{\beta}_1\right) \\
&\quad + \Phi\left(\mathbf{z}'\boldsymbol{\gamma} - \mu_1\right) - \Phi\left(\mathbf{z}'\boldsymbol{\gamma} - \mu_1\right)\Phi\left(\mathbf{x}'\boldsymbol{\beta}_2\right) + \Phi\left(\mathbf{z}'\boldsymbol{\gamma} - \mu_1\right)\Phi\left(\mathbf{x}'\boldsymbol{\beta}_2\right) \\
&= \Phi\left(\mu_0 - \mathbf{z}'\boldsymbol{\gamma}\right) \\
&\quad + \left[\Phi\left(\mu_1 - \mathbf{z}'\boldsymbol{\gamma}\right) - \Phi\left(\mu_0 - \mathbf{z}'\boldsymbol{\gamma}\right)\right] \\
&\quad + \Phi\left(\mathbf{z}'\boldsymbol{\gamma} - \mu_1\right) \\
&= 1
\end{aligned}$$

Finally, it is evident that all individual outcome probabilities must lie in the unit circle: they are all composed of positive, or sums of positive, components (due to the $\Phi(.)$ transformation) and therefore are all positive. And as the sum across j has been above shown to sum to unity, then the individual ones are definitionally in the $(0, 1)$ space, and are accordingly well-defined.

C.3 Comparison/equivalence with a standard Generalised Ordered Probit (GOP) model

The literature on discrete choice is characterised by a number of contributions which propose generalisations of the ordered probit model. A well-known and popular approach is found in the *generalised ordered probit* ('GOP') models of Terza (1985) and Pudney and Shields (2000), in which the threshold parameters are allowed to vary. Here, Greene, Harris, Hollingsworth, and Weterings (2014) argue that because the ordering of the thresholds are not enforced in these models, the predicted probabilities can lie outside of the range of zero and one. As demonstrated above, our proposed generalisations do not suffer from this form of incoherency. However, of related interest is whether under certain parameter restrictions, our model either nests, or is nested, by the GOP. Put another way, it be the case that our proposed extensions to the $ZIOP$ and $MIOP$ models, and their generalizations are simply re-parameterizations of the more usual generalised ordered probit (GOP) model. We now explose, and subsequently discount this possibility using the example of a $GZIOP$ model.

In its most usual form, the boundary parameters in a GOP model would be specified as (REFS)

$$\begin{aligned} \mu_{i0} &= \mathbf{x}_i' \boldsymbol{\delta}_0 \\ \mu_{i1} &= \mu_{i0} + \exp\left(\mathbf{x}_i' \boldsymbol{\delta}_1\right) \\ &\vdots \end{aligned}$$

so that, in particular, P_0 in a GOP would be

$$P_{i0} = \Phi\left(\mathbf{x}_i' \boldsymbol{\delta}_0 - \mathbf{z}_i' \boldsymbol{\gamma}\right)$$

so compared to the same for the $GZIOP$ means that equity would imply that

$$\begin{aligned} \Phi\left(\mathbf{x}_i' \boldsymbol{\delta}_0 - \mathbf{z}_i' \boldsymbol{\gamma}\right) &= \Phi\left(\mu_0 - \mathbf{z}' \boldsymbol{\gamma}\right) + \\ &\quad \left[\Phi\left(\mu_1 - \mathbf{z}' \boldsymbol{\gamma}\right) - \Phi\left(\mu_0 - \mathbf{z}' \boldsymbol{\gamma}\right)\right] \Phi\left(-\mathbf{x}' \boldsymbol{\beta}_1\right) + \\ &\quad \Phi\left(\mathbf{z}' \boldsymbol{\gamma} - \mu_1\right) \Phi\left(-\mathbf{x}' \boldsymbol{\beta}_2\right) \end{aligned}$$

clearly there are no obvious restrictions under which this condition would hold. On this basis, one would conclude that the proposed new models are not simple re-parameterisations of a GOP model.

D Hit-and-miss tables

To evaluate the predictive performance of our models we construct hit–and–miss tables, which provide information about the proportion of correct predictions. This involves cross-tabulating the predictions of a given model obtained using the maximum probability rule, *viz.*,

$$\widehat{y}_i = m \text{ if } \widehat{P}_{im} = \max(\widehat{P}_{i0}, \widehat{P}_{i1}, \widehat{P}_{i2}, ..., \widehat{P}_{iJ-1}) \tag{D.1}$$

against the observed outcomes in a $J \times J$ contingency table, where $\widehat{P}_{ij}$ denotes the predicted probability of outcome j arising for respondent i. The proportion of correct predictions will be given by the sum of all J diagonal elements divided by the total number of observations N, that is

$$CP = \frac{1}{N}\sum_{i=1}^{N} 1\left(\widehat{y}_i = y_i\right) \tag{D.2}$$

Analgously, for each $j = 0, 1, 2, ..., J–1$ this will be obtained by dividing the number of correct predictions within each category by the total number of predictions for that category. We note here that both of our empirical applications are characterised by one outcome dominating all others: in the $ZIOP$ application 76% of observations are non-smokers; for the $MIOP$ application, 56% of respondents answered that joining the EU is a '*good thing*'. Expression (D.2) is therefore adjusted to accommodate the possibility that a high percentage of correct predictions given by (D.2) does not necessarily mean that a statistical model has a good prediction performance. This methodological approach is due to Merton (1981) and Henriksson and Merton (1981), and mitigates the problem of what the authors refer to as a 'stopped-clock' strategy when evaluating forecasts. In our example, this translates to the traditional 'hit and miss' approach arguably placing too much weight on the most heavily chosen outcome.

Here, we follow Kim, Mizen, and Chevapatrakul (2008) and Rosa (2009), who adapt the above approach to give a more reliable criterion of predictive ability when one categorical outcome dominates all others in a discrete ordered setting. Following the maximum probability rule in (D.1), let be the proportion of the correct predictions made by $\widehat{y}_i$ when the true state is given by $y_i = j$ be calculated using

$$CP_j^* = \frac{\frac{1}{N}\sum_{i=1}^{N} 1\left(\widehat{y}_i = j\right) 1\left(y_i = j\right)}{\frac{1}{N}\sum_{i=1}^{N} 1\left(y_i = j\right)}. \tag{D.3}$$

The more reliable criterion is given by

$$CP^* = \frac{1}{J-1}\left[\sum_{j=0}^{J-1} CP_j - 1\right] \tag{D.4}$$

where following Section II, $J = 0, 1, 2....J-1$ is the number of categorical outcomes. The measure lies between $-1/J-1$ and 1: a value of $-1/J-1 \leq CP < 0$ implies a forecasting performance worse than the stopped–clock strategy; ; a value of $CP = 1$ suggests zero predictability, which is consistent with the 'stopped clock' strategy; a value of $CP = 1$ implies a perfect forecasting model, which is consistent with $CP_j = 1$, $\forall\ j = 0, 1, 2, ...J-1$.

Tables D.1 and D.2 present summary measures for the respective $ZIOP$ and $MIOP$ applications from hit–and–miss tables both within sample and for a 10% 'hold-out' sample. In all both cases, the results suggest the generalised models out-perform the other models, although there is some disagreement between the correlated and independent variants.

Table D.1: In-sample and out-of-sample contingency tables for *ZIOP* applications

	Predicted ($\widehat{y}_i$): In-sample																				
Specification	*OP*				*ZIOP*				*ZIOPC*				*GZIOP*				*GZIOPC*				
Actual (y_i)	**0**	**1**	**2**	**3**	**0**	**1**	**2**	**3**	**0**	**1**	**2**	**3**	**0**	**1**	**2**	**3**	**0**	**1**	**2**	**3**	Total
0	21535	0	0	4	21508	0	25	6	21518	0	15	6	21484	0	46	9	21492	0	43	4	21539
1	1244	0	0	0	1239	0	5	0	1241	0	3	0	1237	0	7	0	1236	0	8	0	1244
2	4167	0	0	5	4136	0	32	4	4148	0	19	5	4105	0	61	6	4107	0	63	2	4172
3	1854	0	0	4	1838	0	17	3	1847	0	8	3	1830	0	22	6	1829	0	28	1	1858
Total	28800	0	0	13	28721	0	79	13	28754	0	45	14	28656	0	136	21	28664	0	142	7	28813
CP	0.7475				0.7477				0.7476				0.7480				**0.7481**				
CP^*	0.0006557				0.00262				0.001731				**0.005099**				0.004486				
	Predicted ($\widehat{y}_i$): Out-of-sample – 10% 'hold-out' sample																				
Actual (y_i)	**0**	**1**	**2**	**3**	**0**	**1**	**2**	**3**	**0**	**1**	**2**	**3**	**0**	**1**	**2**	**3**	**0**	**1**	**2**	**3**	Total
0	2200	0	0	0	2199	0	1	0	2199	0	1	0	2196	0	3	1	2196	0	4	0	2200
1	135	0	0	0	135	0	0	0	135	0	0	0	135	0	0	0	135	0	0	0	135
2	419	0	0	0	415	0	2	2	416	0	1	2	413	0	4	2	414	0	5	0	419
3	180	0	0	0	179	0	1	0	179	0	1	0	177	0	3	0	177	0	3	0	180
Total	2934	0	0	0	2928	0	4	2	2929	0	3	2	2921	0	10	3	2922	0	12	0	2934
CP	0.7498				0.7502				0.7498				0.7498				**0.7502**				
CP^*	0				0.0014395747				0.0006440				0.002576				**0.003372**				

Table D.2: In-sample and out-of-sample contingency tables for *MIOP* applications

	Predicted ($\widehat{y}_i$): In-sample															
Specification	*OP*			*MIOP*			*MIOPC*			*GMIOP*			*GMIOPC*			Total
Actual (y_i)	**0**	**1**	**2**	**0**	**1**	**2**	**0**	**1**	**2**	**0**	**1**	**2**	**0**	**1**	**2**	Total
0	2	318	667	14	290	683	10	307	670	6	328	653	8	311	668	987
1	2	640	2372	18	803	2193	9	832	2173	5	916	2093	6	860	2148	3014
2	2	470	4640	11	544	4557	3	574	4535	3	633	4476	4	589	4519	5112
Total	6	1428	7679	43	1637	7433	22	1713	7378	14	1877	7222	18	1760	7335	9113
CP	0.5796			0.5897			0.5900			**0.5923**			0.5911			
CP^*	0.06102			0.08602			0.08665			**0.09279**			0.08872			
	Predicted ($\widehat{y}_i$): Out-of-sample – 10% 'hold-out' sample															
Actual (y_i)	**0**	**1**	**2**	**0**	**1**	**2**	**0**	**1**	**2**	**0**	**1**	**2**	**0**	**1**	**2**	Total
0	0	35	75	1	33	76	1	33	76	0	32	78	1	30	79	110
1	0	59	221	2	71	207	1	78	201	0	86	194	0	75	205	280
2	0	48	487	1	60	474	0	64	471	0	69	466	0	61	474	535
Total	0	142	783	4	164	757	2	175	748	0	187	738	1	166	758	925
CP	0.5903			0.5903			0.5946			**0.5968**			0.5946			
CP^*	0.06050			0.07432			0.08402			**0.08909**			0.08146			