Estimation, specification and testing in middle- and zero-inflated ordered probit models*

Sarah Brown† Mark N. Harris‡ Christopher Spencer§

June 2018

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 our generalisations in modelling zero- and middle-inflation processes.

JEL Classification: C12, C35

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

---

*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). We are also grateful to the Australian Research Council for funding.

†Department of Economics, University of Sheffield, UK. sarah.brown@sheffield.ac.uk.

‡Curtin Business School, Curtin University, Perth, Australia. mark.harris@curtin.edu.au.

§School of Business and Economics, Loughborough University, UK. c.g.spencer@lboro.ac.uk.

1 Introduction and Background

Recent advances in discrete choice modelling have witnessed the development of so-called *inflated ordered probit* models. These build on 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) - and 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.[1] This modelling framework 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. 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), which at the moment, is not available in standard econometric software. Using empirical applications from

[1] The popularity of the $ZIOP$ model is reflected in its recent incorporation into mainstream statistical software (e.g., STATA 15, Limdep/NLogit). Our contribution is thus particularly timely.

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.

To evaluate the proposed LM tests, two studies of category inflation from the literature are exploited. 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.

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 simple 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 second 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.

2 Econometric Framework

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

[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

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.

2.1 The Generalized Zero-Inflated Ordered Probit Model ($GZIOP$)

Consider a discrete random variable y that assumes the discrete ordered values of $y \in 0, 1, ..., J-1$.[3] 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

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

where $\Phi(.)$ is the cumulative distribution function (CDF) of the univariate standard normal distribution. Outcomes in regime 1 are represented by a discrete variable $\widetilde{y}$ ($\widetilde{y} = 0, 1, ..., J-1$)

[3] For ease of comparison, our notation throughout is consistent with that used in Harris and Zhao (2007).

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

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

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.*,

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

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[4]

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

where the indicator function h_{ij} is

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

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 model is an example of a partial observability one:[5] Diagrammatically, this model is illustrated in the left hand panel of Figure 1.

Now consider the model set-up depicted in the right hand panel of Figure 1, in which for all $j > 0$ categories, the individual is tempered towards zero by $J-1$ distinct binary splitting equations. Essentially, this simply reverses the "implicit sequencing" inherent the original $ZIOP$ model.[6] We refer to this approach, and the resulting econometric model derived below, as the "generalized $ZIOP$" (hereafter $GZIOP$) model. As with the $ZIOP$ approach, all equations are estimated simultaneously; however, as we now have $J > 2$ equations, the joint outcome of J DGPs ($J-1$ binary equations and a single OP one) as compared to two, is observed. Importantly, as we shall see below, the model still embodies the important attribute of zero-inflation. In what follows, we demonstrate that the $GZIOP$ collapses to the $ZIOP$ under a certain set of parameter restrictions.

The $GZIOP$ model is considerably more flexible than the $ZIOP$ one. For example, now

[4] In our empirical applications, the common sandwich estimator (White 1982) is used to compute robust standard errors of parameters. Standard errors of secondary estimated quantities, 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 $\boldsymbol{\theta}$.

[5] Also see Poirier (1980) where this concept is applied in the context of a bivariate probit model.

[6] We provide some intuition, and justification, as to the appropriateness of such a generalized approach in our empirical applications below.

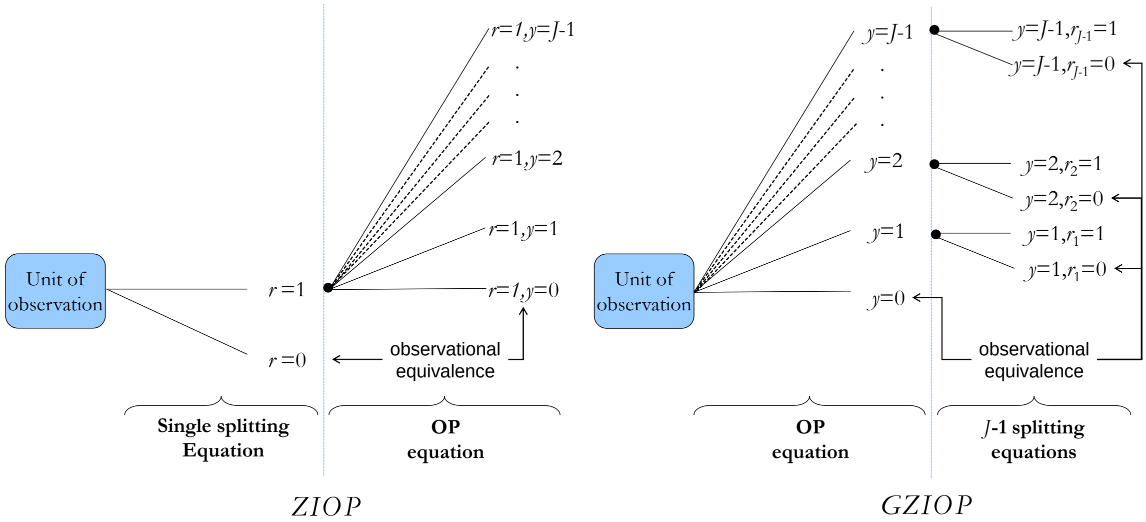

Figure 1: The Zero-Inflated Ordered Probit model ($ZIOP$) and its generalisation ($GZIOP$)

a given variable can have a potential impact via the outcome equation and/or via the set of $J-1$ splitting equations, with differential tempering effects corresponding to each OP propensity permitted via each of the $J-1$ splitting equations. The functional form of the $GZIOP$ model is obtained as follows. Unlike the $ZIOP$ model, which is characterized by a single splitting equation (1), the $GZIOP$ has $j = 1, 2, ..., J-1$ splitting equations of 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

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

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

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

such that

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

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

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

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 $\rho_j\ \forall\ j=1,2,3...J-1$.[7] 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. Diagrammatically, this implies that the model depicted 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. A notable implication of this model characteristic is that compared to the non-generalised model variants, the choice of exclusion restrictions assumes a more

[7] 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 tempering 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.

prominent role, as several splitting equations require identification instead of one.[8]

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

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

[8] In both of our empirical applications, no evidence of identification issues were found to be present. Stronger identification could also be achieved by having differing variable sets in the various splitting equations, although the original *ZIOP*/*MIOP* models would no longer be nested.

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

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

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.*

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

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.[9]

To test the hypotheses associated with the various sets of parameter restrictions described above, this paper uses two approaches in our empirical applications. First, we use the standard LR test. We also propose an LM test. This represents an appealing specification test for the $ZIOPC$ models *versus* their generalized alternatives, as it only requires estimation of the simpler nested models (which can be currently estimated in standard software, unlike the generalized counterparts). It involves evaluation of the score vector of the more general model evaluated at parameter values under the null hypothesis (*i.e.*, at the $ZIOPC$ ones); full details of the LM test can be found in Appendix A. As shown below, the results of the LR and LM tests are very similar in our empirical application, suggesting that the log-likelihood function is well-behaved, and further, that standard asymptotic theory performs well.[10] This finding carries through to the models of middle-inflation, which are discussed in the next section.

[9] 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 property also holds for the $MIOP$ and $MIOPC$ models.

[10] It is also possible to consider subsets of the generalised $ZIOP$ 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 would invariably be based on mis-specified alternative models that would likely adversely affect the test performance.

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

Building on the $ZIOP$ model, two notable contributions - Bagozzi and Mukherjee (2012) and Brooks et al. (2012) - independently suggested the *middle-inflated ordered probit* ($MIOP$) model to allow for inflation in any arbitrary category.[11] In keeping with Section 2.1, we develop the $GMIOP$ framework in the context of J outcomes.[12] Consider again an OP model as a starting point, where each individual i has an unobserved underlying propensity

$$y^* = \mathbf{z}'\boldsymbol{\gamma} + \eta \tag{21}$$

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

$$r^* = \mathbf{x}'\boldsymbol{\beta} + \varepsilon. \tag{22}$$

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

$$\Pr(y_i) = \begin{cases} \Pr\left(y = 0 \left| \mathbf{x}_i, \mathbf{z}_i \right.\right) = \Phi\left(\mathbf{x}_i'\boldsymbol{\beta}\right) \times \Phi\left(\mu_0 - \mathbf{z}_i'\boldsymbol{\gamma}\right) \\ \Pr\left(y = j \left| \mathbf{x}_i, \mathbf{z}_i \right.\right) = \Phi\left(\mathbf{x}_i'\boldsymbol{\beta}\right) \times \left[\Phi\left(\mu_1 - \mathbf{z}_i'\boldsymbol{\gamma}\right) - \Phi\left(\mu_0 - \mathbf{z}_i'\boldsymbol{\gamma}\right)\right] + M \\ \Pr\left(y = J-1 \left| \mathbf{x}_i, \mathbf{z}_i \right.\right) = \Phi\left(\mathbf{x}_i'\boldsymbol{\beta}\right) \times \left[1 - \Phi\left(\mu_{J-2} - \mathbf{z}_i'\boldsymbol{\gamma}\right)\right] \end{cases} \tag{23}$$

[11] 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.

[12] In both original contributions the empirical analysis is restricted to three outcomes. However, the model developed in this section naturally applies to instances where $J > 3$.

whereas for correlated errors we have that

$$\Pr(y_i) = \begin{cases} \Pr\left(y = 0 \left| \mathbf{x}_i, \mathbf{z}_i \right.\right) = \Phi_2\left(\mu_0 - \mathbf{z}_i'\boldsymbol{\gamma}, \mathbf{x}_\mathbf{i}'\boldsymbol{\beta}; -\rho\right) \\ \Pr\left(y = j \left| \mathbf{x}_i, \mathbf{z}_i \right.\right) = \Phi_2\left(\mu_1 - \mathbf{z}_i'\boldsymbol{\gamma}, \mathbf{x}_i'\boldsymbol{\beta}; -\rho\right) - \Phi_2\left(\mu_0 - \mathbf{z}_i'\boldsymbol{\gamma}, \mathbf{x}_i'\boldsymbol{\beta}; -\rho\right) + M \\ \Pr\left(y = J-1 \left| \mathbf{x}_i, \mathbf{z}_i \right.\right) = \Phi_2\left(\mathbf{x}_i'\boldsymbol{\beta}, \mathbf{z}_i'\boldsymbol{\gamma} - \mu_{J-2}; \rho\right) \end{cases} \tag{24}$$

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,

$$\Pr\left(y = m \left| \mathbf{x}_i, \mathbf{z}_i \right.\right) = \Phi\left(\mathbf{x}_i'\boldsymbol{\beta}\right) \times \left[\Phi\left(\mu_1 - \mathbf{z}_i'\boldsymbol{\gamma}\right) - \Phi\left(\mu_0 - \mathbf{z}_i'\boldsymbol{\gamma}\right)\right] + 1 - \Phi\left(\mathbf{x}_i'\boldsymbol{\beta}\right) \tag{25}$$

and for the case of correlated errors

$$\Pr\left(y = m \left| \mathbf{x}_i, \mathbf{z}_i \right.\right) = \Phi_2\left(\mu_1 - \mathbf{z}_i'\boldsymbol{\gamma}, \mathbf{x}_\mathbf{i}'\boldsymbol{\beta}; -\rho\right) - \Phi_2\left(\mu_0 - \mathbf{z}_i'\boldsymbol{\gamma}, \mathbf{x}_\mathbf{i}'\boldsymbol{\beta}; -\rho\right) + 1 - \Phi\left(\mathbf{x}_i'\boldsymbol{\beta}\right) \tag{26}$$

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 2.1, 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

$$r_{j\neq m}^{*} = \mathbf{x}'\boldsymbol{\beta}_j + \varepsilon_j \tag{27}$$

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

$$\Pr(r_{j\neq m} = 0) = \Phi\left(-\mathbf{x}'\boldsymbol{\beta}_j\right) \tag{28}$$

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

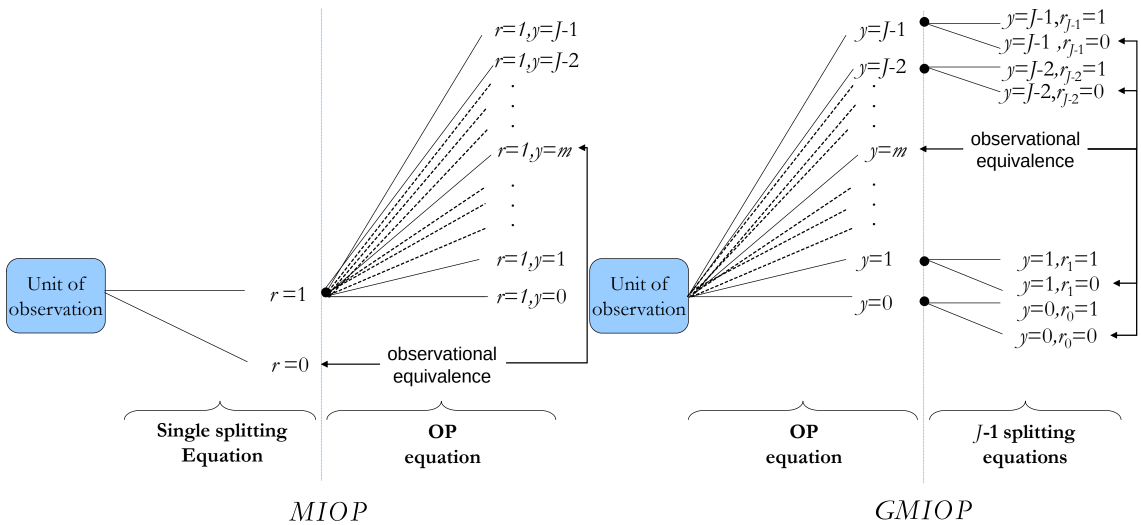

Figure 2: The Middle-Inflated Ordered Probit model ($MIOP$) and its generalisation, ($GMIOP$)

comes are

$$\Pr(y_i) = \begin{cases} \Pr\left(y = 0 \left| \mathbf{x}_i, \mathbf{z}_i \right.\right) = \Phi\left(\mu_0 - \mathbf{z}_i'\boldsymbol{\gamma}\right) \times \Phi\left(\mathbf{x}_i'\boldsymbol{\beta}_0\right) \\ \Pr\left(y = \widetilde{j} \left| \mathbf{x}_i, \mathbf{z}_i \right.\right) = \left[\Phi\left(\mu_{\widetilde{j}} - \mathbf{z}_i'\boldsymbol{\gamma}\right) - \Phi\left(\mu_{\widetilde{j}-1} - \mathbf{z}_i'\boldsymbol{\gamma}\right)\right] \times \Phi\left(\mathbf{x}_i'\boldsymbol{\beta}_{\widetilde{j}}\right) \\ \Pr\left(y = m \left| \mathbf{x}_i, \mathbf{z}_i \right.\right) = \begin{cases} \left[\Phi\left(\mu_m - \mathbf{z}_i'\boldsymbol{\gamma}\right) - \Phi\left(\mu_{m-1} - \mathbf{z}_i'\boldsymbol{\gamma}\right)\right] \\ + \underbrace{\Phi\left(\mu_0 - \mathbf{z}_i'\boldsymbol{\gamma}\right) \times \Phi\left(-\mathbf{x}_i'\boldsymbol{\beta}_0\right)}_{a} \\ \underbrace{+ \sum_{\widetilde{j}} \left[\Phi\left(\mu_{\widetilde{j}} - \mathbf{z}_i'\boldsymbol{\gamma}\right) - \Phi\left(\mu_{\widetilde{j}-1} - \mathbf{z}_i'\boldsymbol{\gamma}\right)\right] \times \Phi\left(-\mathbf{x}_i'\boldsymbol{\beta}_{\widetilde{j}}\right)}_{b} \\ + \underbrace{\left[1 - \Phi\left(\mu_{J-2} - \mathbf{z}_i'\boldsymbol{\gamma}\right)\right] \times \Phi\left(-\mathbf{x}_i'\boldsymbol{\beta}_{J-1}\right)}_{c} \end{cases} \\ \Pr\left(y = J-1 \left| \mathbf{x}_i, \mathbf{z}_i \right.\right) = \left[1 - \Phi\left(\mu_{J-2} - \mathbf{z}_i'\boldsymbol{\gamma}\right)\right] \times \Phi\left(\mathbf{x}_i'\boldsymbol{\beta}_{J-1}\right) \end{cases} . \tag{29}$$

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 (29). Expression (29) 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

$$\Pr(y_i) = \begin{cases} \Pr\left(y=0\,|\mathbf{x}_i,\mathbf{z}_i\right) = \Phi_2(\mu_0 - \mathbf{z}_i'\boldsymbol{\gamma}, \mathbf{x}_i'\boldsymbol{\beta}_0; -\rho_0) \\ \Pr\left(y=\widetilde{j}\,|\mathbf{x}_i,\mathbf{z}_i\right) = \Phi_2\left(\mu_{\widetilde{j}} - \mathbf{z}_i'\boldsymbol{\gamma}, \mathbf{x}_i'\boldsymbol{\beta}_{\widetilde{j}}; -\rho_{\widetilde{j}}\right) - \Phi_2\left(\mu_{\widetilde{j}-1} - \mathbf{z}_i'\boldsymbol{\gamma}, \mathbf{x}_i'\boldsymbol{\beta}_{\widetilde{j}}; -\rho_{\widetilde{j}}\right) \\ \Pr\left(y=m\,|\mathbf{x}_i,\mathbf{z}_i\right) = \begin{cases} \left[\Phi\left(\mu_m - \mathbf{z}_i'\boldsymbol{\gamma}\right) - \Phi\left(\mu_{m-1} - \mathbf{z}_i'\boldsymbol{\gamma}\right)\right] \\ +\underbrace{\Phi_2\left(\mu_0 - \mathbf{z}_i'\boldsymbol{\gamma}, -\mathbf{x}_i'\boldsymbol{\beta}_0; \rho_0\right)}_{a} \\ \underbrace{+\sum_{\widetilde{j}=1}^{J-2}\begin{bmatrix} \Phi_2\left(\mu_{\widetilde{j}} - \mathbf{z}_i'\boldsymbol{\gamma}, -\mathbf{x}_i'\boldsymbol{\beta}_{\widetilde{j}}; \rho_{\widetilde{j}}\right) \\ -\Phi_2\left(\mu_{\widetilde{j}-1} - \mathbf{z}_i'\boldsymbol{\gamma}, -\mathbf{x}_i'\boldsymbol{\beta}_{\widetilde{j}}; \rho_{\widetilde{j}}\right)\end{bmatrix}}_{b} \\ +\underbrace{\Phi_2\left(\mathbf{z}_i'\boldsymbol{\gamma} - \mu_{J-2}, -\mathbf{x}_i'\boldsymbol{\beta}_{J-1}; -\rho_{J-1}\right)}_{c} \end{cases} \\ \Pr\left(y=J-1\,|\mathbf{x}_i,\mathbf{z}_i\right) = \Phi_2\left(\mathbf{z}_i'\boldsymbol{\gamma} - \mu_{J-2}, \mathbf{x}_i'\boldsymbol{\beta}_{J-1}; \rho_{J-1}\right) \end{cases} \tag{30}$$

As in (29), 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 (2.1), 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 (30); 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

$$\Pr(y_i) = \begin{cases} \Pr\left(y=0\,|\mathbf{x}_i,\mathbf{z}_i\right) = \Phi_2(\mu_0 - \mathbf{z}_i'\boldsymbol{\gamma}, \mathbf{x}_i'\boldsymbol{\beta}; -\rho) \\ \Pr\left(y=\widetilde{j}\,|\mathbf{x}_i,\mathbf{z}_i\right) = \Phi_2\left(\mu_{\widetilde{j}} - \mathbf{z}_i'\boldsymbol{\gamma}, \mathbf{x}_i'\boldsymbol{\beta}; -\rho\right) - \Phi_2\left(\mu_{\widetilde{j}-1} - \mathbf{z}_i'\boldsymbol{\gamma}, \mathbf{x}_i'\boldsymbol{\beta}; -\rho\right) \\ \Pr\left(y=m\,|\mathbf{x}_i,\mathbf{z}_i\right) = \begin{cases} \left[\Phi\left(\mu_m - \mathbf{z}_i'\boldsymbol{\gamma}\right) - \Phi\left(\mu_{m-1} - \mathbf{z}_i'\boldsymbol{\gamma}\right)\right] \\ +\underbrace{\Phi_2\left(\mu_0 - \mathbf{z}_i'\boldsymbol{\gamma}, -\mathbf{x}_i'\boldsymbol{\beta}; \rho\right)}_{a} \\ \underbrace{+\sum_{\widetilde{j}} \begin{bmatrix} \Phi_2\left(\mu_{\widetilde{j}} - \mathbf{z}_i'\boldsymbol{\gamma}, -\mathbf{x}_i'\boldsymbol{\beta}; \rho\right) \\ -\Phi_2\left(\mu_{\widetilde{j}-1} - \mathbf{z}_i'\boldsymbol{\gamma}, -\mathbf{x}_i'\boldsymbol{\beta}; \rho\right) \end{bmatrix}}_{b} \\ +\underbrace{\Phi_2\left(\mathbf{z}_i'\boldsymbol{\gamma} - \mu_{J-2}, -\mathbf{x}_i'\boldsymbol{\beta}; -\rho\right)}_{c} \end{cases} \\ \Pr\left(y=J-1\,|\mathbf{x}_i,\mathbf{z}_i\right) = \Phi_2\left(\mathbf{z}_i'\boldsymbol{\gamma} - \mu_{J-2}, \mathbf{x}_i'\boldsymbol{\beta}; \rho\right) \end{cases} \tag{31}$$

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 (24).[13] Using (24), subtracting these terms from one yields

$$\Pr\left(y=m\right) = \Phi_2\left(\mu_m - \mathbf{z}_i'\boldsymbol{\gamma}, \mathbf{x}'\boldsymbol{\beta}; -\rho\right) - \Phi_2\left(\mu_{m-1} - \mathbf{z}_i'\boldsymbol{\gamma}, \mathbf{x}'\boldsymbol{\beta}; -\rho\right) + 1 - \Phi\left(\mathbf{x}'\boldsymbol{\beta}\right) \tag{32}$$

That is, the *GMIOPC* collapses to and therefore nests the *MIOPC*. Further, setting $\rho = 0$ in (32) yields probabilities that are identical to the *MIOP* probabilities in expression (26), *viz.*

$$\Pr\left(y=m\,|\mathbf{x}_i,\mathbf{z}_i\right) = \Phi\left(\mathbf{x}_i'\boldsymbol{\beta}\right) \times \left[\Phi\left(\mu_1 - \mathbf{z}_i'\boldsymbol{\gamma}\right) - \Phi\left(\mu_0 - \mathbf{z}_i'\boldsymbol{\gamma}\right)\right] + 1 - \Phi\left(\mathbf{x}_i'\boldsymbol{\beta}\right) \tag{33}$$

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 (23). 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

[13] 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 $M = 0$.

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. Further, as with the $GZIOP$, a straightforward test of hypotheses can be undertaken using a standard LR test or LM tests.[14]

3 Empirical applications

To explore the performance of the generalizations and testing framework developed above, we consider two key empirical examples from the extant literature. We revisit these examples in order to help us assess the importance of our modelling contribution and to re-evaluate the findings of the original contributions themselves. Each of these is characterized by the application of an inflated ordered probit approach to modelling responses in large-scale survey data sets. Our $GZIOP$ application re-visits the original work of Harris and Zhao (2007) – which introduced the $ZIOP$ – for which our test uses the same data set and specification. The focus of their health economics based application is tobacco consumption. For the $GMIOP$ attention turns to the work of Bagozzi and Mukherjee (2012), who in an innovative study, 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).[15]

3.1 $GZIOP$ application: tobacco consumption

Our empirical example for the $GZIOP$ model uses a data set that has been used to analyze the determinants of how many cigarettes an individual chooses to smoke on a daily basis. A zero-inflated application is deemed to be particularly appropriate, since zeros may be construed as relating to two DGPs: Non-participation due to, for example, health and legal reasons; and further, being at a corner solution whereby such individuals will smoke if the price is low enough, or their income is high enough. In using a $ZIOP$, Harris and

[14] The $GMIOPC$ score vector closely follows that for the $GZIOPC$ one and is presented in Appendix B.

[15] 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.

Zhao (2007) argue that if the underlying processes driving category inflation are modelled incorrectly - for instance, by using a simple ordered probit (OP) approach - any subsequent policy implications may be invalid.

The data set in Harris and Zhao (2007) 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 (outcome) equation, covariates include standard demand-schedule variables such as income and own- and cross-drug prices,[16] 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 $k_z = 18$.

As noted above, such zero-inflated models are examples of *partial observability* ones: observationally equivalent outcomes can arise from distinct DGPs. However, it may help intuition to think of inherent sequencing. For example, in Harris and Zhao (2007) an individual makes a participation decision, and then for participants, a consumption decision is made. The fact that consumption can still be zero for some participants gives rise to zero-inflation.

However, consider reversing the implicit process described above, such that an individual may be predisposed to a given level of cigarette consumption, which may be genetically determined. With respect to the latter point, there is much medical evidence suggesting

[16] In the results presented below, $Ln(P_{A/M/T})$ respectively, refers to the (log of) the price of alcohol / marijuana / tobacco.

that individuals have genetic predispositions to levels of drug use; see, for example, Gelernter and Kranzler (2015). 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 to zero. To accommodate this requires that corresponding to each positive consumption level is a separate binary equation which splits individuals into two types: those remaining at their inherent consumption level, and those who are "pushed" towards zero. The $GZIOPC$ model developed above allows for this possibility. 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 $j \geq 1$.[17] 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.[18]

[17] In our empirical application this is attributable to factors such as health status, medical considerations, income, and wealth.

[18] 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

Table 1: 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	$3.33E-27$
$GZIOP$vs $GZIOPC$	27	3	$5.89E-06$	34	$1.98E-07$

Table 1 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 1 reports the corresponding LR tests, which closely mirror the LM ones.[19] The closeness of the LR and LM test statistics is suggestive 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 2 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 3. For the $GZIOPC$ we witness some very large changes across $j = 1, 2$ and 3 as compared to $ZIOPC$; here, we recall that

$MIOP$) 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.

[19] We note 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). This is also re-visited in the Monte Carlo section in Appendix C.

Table 2: 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.

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).[20] 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 4 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 $GZIOPC$ in-

[20] Similar results, not reported here, 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.

Table 3: 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 2.

dicates 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 5 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.[21] This is achieved 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.[22] 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 5 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.

3.2 $MIOP$ application: ***Eurobarometer*** survey data

Bagozzi and Mukherjee (2012) introduce a $MIOP$ framework to analyze individual responses in a survey data set that explores attitudes towards European Union (EU) membership in EU

[21] Reassuringly, we also note that the overall probabilities for all model variants match the observed sample means in the dataset.

[22] Note that 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.

Table 4: 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 2.

Table 5: 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 2.

accession countries. 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, thereby warranting a $MIOP$ approach.[23] 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 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

[23] Following a number of extant studies which use similar data, a fourth '*do not know*' category is treated as being a "neither good nor bad" response by Bagozzi and Mukherjee (2012). The authors report that their findings remain unchanged when "do not know" responses are dropped from estimations.

Table 6: 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

xenophobic (*xenophobia*).

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 6 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 3.1, 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.[24] 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 8. 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

[24] It is possible that the tests against the $GMIOP$ model are picking-up model mis-specification due to erroneously ignoring the correlation; see Appendix C.

Table 7: 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	

model.

Table 8 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 9. 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

Table 8: 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)**

$j = 1, 2$ 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 8, 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.[25] 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.

[25] In this regard, 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.

Table 9: 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 10: 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%		

Model summary probabilities are given in Table 10. 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.[26] 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 3.1, 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 8 to 9, are associated with movement away from the $j = 2$ tempering equation.[27]

4 Conclusions and discussion

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

[26] Again, for the correlated versions the implied independent OP is used in these calculations.

[27] To ascertain the finite sample behavior of the suggested tests we also conducted a set of Monte Carlo (MC) experiments. All LM tests were correctly sized, and exhibited good power in identifying mis-specified models. See Appendix C for full details.

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.

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). Eurobarameter 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.

Gelernter, J. and H. R. Kranzler (2015). Genetics of addiction. In M. Galanter, H. D. Kleber, and K. T. Brady (Eds.), *The American Psychiatric Publishing Textbook of Substance Abuse Treatment* (5th ed.)., Chapter 2, pp. 25–45. American Psychiatric Association.

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

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

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.

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.

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.

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

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.

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

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

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

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

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

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

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

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

As stated in Section 2.1 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,

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

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

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

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

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

$$\vdots$$

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

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

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

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

whereas for ρ_{J-1} we have

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

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

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

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

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 $MIOP$ score vector

To aid notation and to coincide with our empirical application in section 3.2, 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

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

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

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

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

$$\nabla\mu_0 = \begin{bmatrix} \sum\limits_{y_i=0} \left[\phi\left(\mu_0 - \mathbf{z}'\boldsymbol{\gamma}\right) \times \Phi\left(\frac{\mathbf{x}'\boldsymbol{\beta}_0 + \rho_0(\mu_0 - \mathbf{z}'\boldsymbol{\gamma})}{\sqrt{1-\rho_0^2}}\right)\right] + \\ \sum\limits_{y_i=1} \begin{bmatrix} \phi\left(\mu_0 + \exp\left(\xi_1\right) - \mathbf{z}'\boldsymbol{\gamma}\right) - \phi\left(\mu_0 - \mathbf{z}'\boldsymbol{\gamma}\right) + \\ \phi\left(\mu_0 - \mathbf{z}'\boldsymbol{\gamma}\right) \times \Phi\left(\frac{(-\mathbf{x}'\boldsymbol{\beta}_0) - \rho_0(\mu_0 - \mathbf{z}'\boldsymbol{\gamma})}{\sqrt{1-\rho_0^2}}\right) + \\ \phi\left(\mathbf{z}'\boldsymbol{\gamma} - \mu_0 - \exp\left(\xi_1\right)\right) \times \Phi\left(\frac{(-\mathbf{x}'\boldsymbol{\beta}_2) + \rho_2(\mathbf{z}'\boldsymbol{\gamma} - \mu_0 - \exp(\xi_1))}{\sqrt{1-\rho_2^2}}\right) \end{bmatrix} + \\ \sum\limits_{y_i=J-1} \left[\phi\left(\mathbf{z}'\boldsymbol{\gamma} - \mu_0 - \exp\left(\xi_1\right)\right) \times \Phi\left(\frac{\mathbf{x}'\boldsymbol{\beta}_2 - \rho_2(\mathbf{z}'\boldsymbol{\gamma} - \mu_0 - \exp(\xi_1))}{\sqrt{1-\rho_2^2}}\right)\right] \end{bmatrix} \div P_{j=y_i}^{GMIOPC}. \tag{B.3}$$

and

$$\nabla\xi_1 = \begin{bmatrix} \sum\limits_{y_i=1} \begin{bmatrix} \exp\left(\xi_1\right)\phi\left(\mu_0 + \exp\left(\xi_1\right) - \mathbf{z}'\boldsymbol{\gamma}\right) + \\ \left(-\exp\left(\xi_1\right)\right)\phi\left(\mathbf{z}'\boldsymbol{\gamma} - \mu_0 - \exp\left(\xi_1\right)\right) \times \Phi\left(\frac{(-\mathbf{x}'\boldsymbol{\beta}_2) + \rho_2(\mathbf{z}'\boldsymbol{\gamma} - \mu_0 - \exp(\xi_1))}{\sqrt{1-\rho_2^2}}\right) \end{bmatrix} + \\ \sum\limits_{y_i=J-1} \left[\left(-\exp\left(\xi_1\right)\right)\phi\left(\mathbf{z}'\boldsymbol{\gamma} - \mu_0 - \exp\left(\xi_1\right)\right) \times \Phi\left(\frac{\mathbf{x}'\boldsymbol{\beta}_2 - \rho_2(\mathbf{z}'\boldsymbol{\gamma} - \mu_0 - \exp(\xi_1))}{\sqrt{1-\rho_2^2}}\right)\right] \end{bmatrix} \div P_{j=y_i}^{GMIOPC}. \tag{B.4}$$

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

$$\boldsymbol{\nabla\beta_0} = \begin{bmatrix} \sum\limits_{y_i=0} \left[\mathbf{x}\phi\left(\mathbf{x}'\boldsymbol{\beta}_0\right) \times \Phi\left(\frac{(\mu_0 - \mathbf{z}'\boldsymbol{\gamma}) + \rho_0(\mathbf{x}'\boldsymbol{\beta}_0)}{\sqrt{1-\rho_0^2}}\right)\right] + \\ \sum\limits_{y_i=1} \left[-\mathbf{x}\phi\left(-\mathbf{x}'\boldsymbol{\beta}_0\right) \times \Phi\left(\frac{(\mu_0 - \mathbf{z}'\boldsymbol{\gamma}) - \rho_0(-\mathbf{x}'\boldsymbol{\beta}_0)}{\sqrt{1-\rho_0^2}}\right) + \right] \end{bmatrix} \div P_{j=y_i}^{GMIOPC} \tag{B.5}$$

and

$$\boldsymbol{\nabla\beta_1} = \begin{bmatrix} \sum\limits_{y_i=1} \left[-\mathbf{x}\phi\left(-\mathbf{x}'\boldsymbol{\beta}_2\right) \times \Phi\left(\frac{(\mathbf{z}'\boldsymbol{\gamma} - \mu_1) + \rho_2(-\mathbf{x}'\boldsymbol{\beta}_2)}{\sqrt{1-\rho_2^2}}\right)\right] \\ \sum\limits_{y_i=J-1} \left[\mathbf{x}\phi\left(-\mathbf{x}'\boldsymbol{\beta}_2\right) \times \Phi\left(\frac{(\mathbf{z}'\boldsymbol{\gamma} - \mu_1) - \rho_2(\mathbf{x}'\boldsymbol{\beta}_2)}{\sqrt{1-\rho_2^2}}\right)\right] \end{bmatrix} \div P_{j=y_i}^{GMIOPC} \tag{B.6}$$

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 (23); P_j^{GMIOP} as those for the $GMIOP$ model of expression (29); and finally, P^0 as the tempering equation probability of $\Phi\left(\mathbf{x}'\boldsymbol{\beta}_0\right)$, P^{J-1} as the tempering equation probability of $\Phi\left(\mathbf{x}'\boldsymbol{\beta}_{J-1}\right)$, 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 $\mu_j = \mu_{j-1} + \exp\left(\xi_j\right)$, 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

$$\frac{\partial \ell\left(\boldsymbol{\theta}\right)}{\partial \boldsymbol{\gamma}} = \left[\begin{array}{c} \sum\limits_{y_i=0} -\mathbf{z}\widetilde{\mu}_{-1}P^0 \\ + \sum\limits_{y_i=1} \left(-\mathbf{z}\widetilde{\mu}_0 - \mathbf{z}\widetilde{\mu}_{-1}\left(1-P^0\right) + \mathbf{z}\widetilde{\mu}_1\left(1-P^2\right)\right) \\ + \sum\limits_{y_i=2} \mathbf{z}\widetilde{\mu}_1 P^2 \end{array}\right] \div P^{GMIOP}_{j=y_i} \tag{B.7}$$

$$\begin{aligned} \frac{\partial \ell\left(\boldsymbol{\theta}\right)}{\partial \mu_0} &= \left[\sum_{y_i=0} \widetilde{\mu}_{-1}P^0\right] \div P^{GMIOP}_{j=0} + \\ &\left[\sum_{y_i=1} \widetilde{\mu}_0 + \widetilde{\mu}_0\left(1-P^0\right) - \widetilde{\mu}_1\left(1-P^2\right)\right] \div P^{GMIOP}_{j=0} + \\ &\left[\sum_{y_i=1} -\widetilde{\mu}_1 P^2\right] \div P^{GMIOP}_{j=2} \end{aligned} \tag{B.8}$$

$$\begin{aligned} \frac{\partial \ell\left(\boldsymbol{\theta}\right)}{\partial \xi} &= \left[\sum_{y_i=1} \exp\left(\xi\right)\widetilde{\mu}_1 - \exp\left(\xi\right)\widetilde{\mu}_1\left(1-P^2\right)\right] \div P^{GMIOP}_{j=1} + \\ &\left[\sum_{y_i=2} -\exp\left(\xi\right)\widetilde{\mu}_1 P^2\right] \div P^{GMIOP}_{j=2} \end{aligned} \tag{B.9}$$

$$\begin{aligned} \frac{\partial \ell\left(\boldsymbol{\theta}\right)}{\partial \boldsymbol{\beta}_0} &= \left[\sum_{y_i=0} \mathbf{x}\phi\left(\mathbf{x}'\boldsymbol{\beta}_0\right) P^{OP}_{j=0}\right] \div P^{GMIOP}_{j=0} + \\ &\left[\sum_{y_i=1} -\mathbf{x}\phi\left(\mathbf{x}'\boldsymbol{\beta}_0\right) \times P^{OP}_{j=0}\right] \div P^{GMIOP}_{j=1} \end{aligned} \tag{B.10}$$

$$\begin{aligned} \frac{\partial \ell\left(\boldsymbol{\theta}\right)}{\partial \boldsymbol{\beta}_2} &= \left[\sum_{y_i=1} -\mathbf{x}\phi\left(\mathbf{x}'\boldsymbol{\beta}_0\right) P^{OP}_{j=2}\right] \div P^{GMIOP}_{j=1} + \\ &\left[\sum_{y_i=2} \mathbf{x}\phi\left(\mathbf{x}'\boldsymbol{\beta}_2\right) \times P^{OP}_{j=2}\right] \div P^{GMIOP}_{j=2} \end{aligned} \tag{B.11}$$

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

$$LM_{correlated}^{MIOP} = (\nabla\boldsymbol{\beta}, \nabla\boldsymbol{\gamma}, \nabla\mu_0, \nabla\boldsymbol{\xi}, \nabla\boldsymbol{\rho})' \left[\mathbf{I}\left(\hat{\theta}_R\right)\right]^{-1} (\nabla\boldsymbol{\beta}, \nabla\boldsymbol{\gamma}, \nabla\mu_0, \nabla\boldsymbol{\xi}, \nabla\boldsymbol{\rho}) \qquad \text{(B.12)}$$

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 $\nabla\boldsymbol{\rho}$ where the uncorrelated generalised variant is the alternative model.

C Finite sample performance

To ascertain the finite sample behavior of the suggested tests we conduct a set of Monte Carlo (MC) experiments. In each experiment the total number of repetitions was set to $2,000$, although any simulation "noise" had settled down after about $1,000$ repetitions. For both the $ZIOPC$ and $MIOPC$ models, we considered two null models: A correlated version and an independent version. As the $MIOPC$ specification appeared to be relatively well-specified, to generate under the two nulls we estimated both $MIOP$ and $MIOPC$ models and took these estimated parameters to be the true ones in generating the data. However, as the $ZIOPC$ specification failed so drastically, this was not considered to be an appropriate specification to generate under the null hypotheses. Accordingly, the model was re-specified as detailed below. The results of these simulation exercises can be found in Panel A (zero-inflated models) and Panel B (middle-inflated models) in Table C.1.

Consider the $ZIOPC$ results. As noted above the model is re-specified given its drastic failure. This is achieved by combining the middle two categories. In the first row of Panel A we generate under the $ZIOP$, such that the numbers in columns 1 and 3 correspond to empirical size. At just over 5% these are almost exactly sized. Interestingly, the second column, where we test for $GZIOPC\ versus\ ZIOPC$, also appears correctly "sized", although the null model here is actually assumed to be $ZIOPC$. These findings may reflect the fact that when we generate under the $ZIOP$ the model estimates of ρ are invariably very small such that the $ZIOP$ and the $ZIOPC$ are essentially the same model. In the second row, where the DGP is the $ZIOPC$, the second column provides the true empirical size of the test, and again appears correct at 5.8%. However, columns 1 and 3 report relatively high rejection probabilities against the (false) null model of $ZIOP$. These tests indicate that the $ZIOP$ should be rejected due to the erroneous omission of ρ. This rejection suggests that if tests for the non-correlated versions fail, it is appropriate to estimate the correlated model before moving onto the generalised variants.

In the row three of Panel A ($ZIOP(df = 26, 28, 29)$), we explore the consequences of increasing the degrees of freedom (df) of the test. This is achieved by reverting to the original $J = 4$ categories (which doubles the df). The model is estimated under this specification, and the resulting parameter estimates are used as the true ones generating the data in the MC repetitions. Once again, all three tests have very similar performance, and the results are all correctly sized at around 6%.[28]

Rejection of the null models might not be due to mis-specification in the direction of the generalised variants, but may reflect other forms of model mis-specification. To ascertain the suitability of the proposed tests as general specification tests, we also generate under an ordered probit DGP. This is based on an equation of the form of expression (3) with a single $\boldsymbol{\gamma}$ vector and a set of boundary parameters. Specifically, we estimate the $ZIOPC$ and $MIOPC$ models on this data and apply the new LM tests. We also consider a parallel regressions type model (Brant 1990) where the data is generated by multiple $\boldsymbol{\gamma}_j$ vectors generated by independent binary models for all observed values of j: For example, $\boldsymbol{\gamma}_0$ is obtained by a binary model of $1(j = 0)$ on $\mathbf{z}$. These two alternatives were chosen as they reflect the most likely forms of serious model mis-specification encountered with data of this sort. As above, the respective parameters chosen to generate under these specifications were obtained by estimating the true null model on the real data, and then subsequently using these estimated parameters as the true ones in the subsequent experiments.

The results of these exercises are also presented in Table C.1, Panels A and B, rows 4 and 5. Concentrating on the $ZIOPC$ results, all tests have good general power ($24\% - 36\%$) against the most likely source of model mis-specification, the OP DGP. Against the parallel regressions type model, all of the tests have power at around 14%. It is important to acknowledge that these favourable results may be specific to the experimental set-ups considered.

Next we consider the results for the LM tests applied to the $MIOPC$ model(s). Where the $MIOP$ model is the true DGP, The tests are correctly sized (columns 1 and 3), with rejection probabilities around the 5% mark (5.7% and 6.2%, respectively). The correlated

[28]Intervals of size $\pm$ 2 standard deviations all effectively contain the nominal 5% value throughout the size experiments.

Table C.1: Monte Carlo rejection probabilities

	Panel A		
Model DGP	*GZIOP vs ZIOP*	*GZIOPC vs ZIOPC*	*GZIOPC vs ZIOP*
ZIOP ($df = 13, 14, 15$)	0.053	0.058	0.056
ZIOPC ($df = 13, 14, 15$)	0.381	0.058	0.489
ZIOP ($df = 26, 28, 29$)	0.059	0.061	0.063
OP ($df = 13, 14, 15$)	0.252	0.358	0.239
Parallel ($df = 13, 14, 15$)	0.141	0.140	0.144
	Panel B		
Model DGP	*GMIOP vs MIOP*	*GMIOPC vs MIOPC*	*GMIOPC vs MIOP*
MIOP ($df = 12, 13, 14$)	0.057	0.061	0.062
MIOPC ($df = 12, 13, 14$)	0.181	0.061	0.241
MIOP ($df = 7, 8, 9$)	0.056	0.055	0.053
OP ($df = 12, 13, 14$)	0.484	0.788	0.657
Parallel ($df = 12, 13, 14$)	0.253	0.677	0.503

version (column 2) does not have power in identifying the mis-specified *MIOPC* model. In row three, Panel B, we investigate the effect of changing the *df* of the test. As we only have 3 outcomes we reduce the *df* from the baseline model by removing from it insignificant variables. These results mirror those of Panel B, row 1, with the exception that the smaller *df* experiments exhibit marginally lower rejection probabilities in line with the findings of the *ZIOP* experiments. The final size experiment, in Panel B, row 2, considers test performance where the DGP is the *MIOPC*. Mirroring the *ZIOP* results, the test appears approximately correctly sized, at just over 6%, and the other tests (columns 1 and 3) seem to have reasonable power at picking-up the mis-specified *MIOP* model (with rejection probabilities around 18% − 25%, respectively). Finally, *quasi*-power experiments are considered, generating under an *OP* model and parallel regressions one (Panel B, rows 4 and 5, respectively). All tests behave well as general specification tests, as evidenced by high rejection probabilities.

For illustrative purposes, we conduct true power experiments for the *ZIOP* (*MIOP*) model(s) – that is, in the direction of the alternative model of the *GZIOP* (*GMIOP*). For brevity, we only report power runs for the non-correlated DGPs and the appropriate *LM* test(s), namely the non-correlated generalised variant *versus* the non-correlated nested ones. For the *ZIOP* model these consisted of taking the estimated value of $\boldsymbol{\beta}$ from the *ZIOP* model, setting $\boldsymbol{\beta}_j = \boldsymbol{\beta}\ \forall\ j$ in the *GZIOP* set-up, and then perturbing a single parameter $\beta_{0,J-1}$ by successively larger increments. Specifically, small amounts are successively added to $\beta_{0,J-1}$. Comparable experiments were then undertaken for the *MIOP* model. We also consider the effect of changing the *df* for the test(s). Two distinct approaches to doing this were considered. In the *ZIOP* model the *df* is increased by having an additional outcome, and accordingly an additional tempering parameter vector in the null model. Alternatively, for the $MIOP$ model, the decrease in the df is obtained by dropping insignificant variables from the tempering equation(s) based on estimation under the null hypothesis. For both cases power curves are obtained by gradually perturbing a single parameter from its null value as before.

For both models, the results of these experiments are shown in Figures C.1 and C.2. All tests have the usual shaped power curves. The power curve corresponding to a higher df for the $ZIOP$ model has uniformly higher power,[29] although in the case of the $MIOP$ model, the difference in df has no discernible effect. This latter finding is to be expected, as the perturbed parameter is still only different from one other in all experiments in both cases. However, one might expect in this particular instance, power gains from a more severe difference in the df, as in smaller df models a single parameter model failure could be interpreted as representing a "drastic" model failure.[30]

To summarize all LM tests appear correctly sized, and typically have good power in identifying mis-specified models. Tests when the true model is correlated correctly reject false uncorrelated variants. However, this is simply a rejection of the non-correlated model, and not the structural form of the model. This would imply that model rejection should lead to consideration of the correlated versions of these models before rejecting these models in favour of their generalised counterparts. Finally, any issues regarding the estimation of the null model and/or issues in conducting the test, would strongly suggest serious model mis-specification. The true power results demonstrate that the ability of the test(s) 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. Moreover, they also respond to changes in the df of the test. Differences in the way that the df are obtained may have effects on the power of the test(s). However, as with all MC experiments, the results may be dependent upon the particular experiments considered.

[29] The increased df scenario was achieved by increasing J from 3 to 4. In the former case the single perturbed parameter differs from 2 others, whereas in the small df case it differs from only one.

[30] 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.

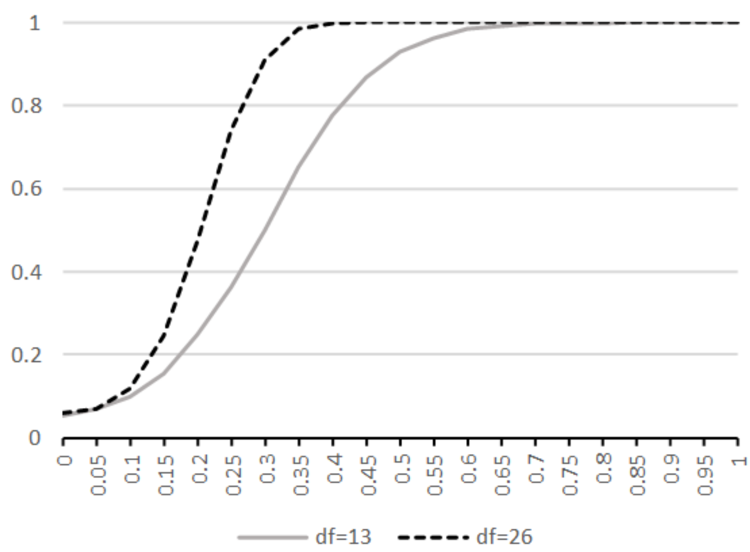

Figure C.1: ZIOP empirical power curves: large and small degrees of freedom

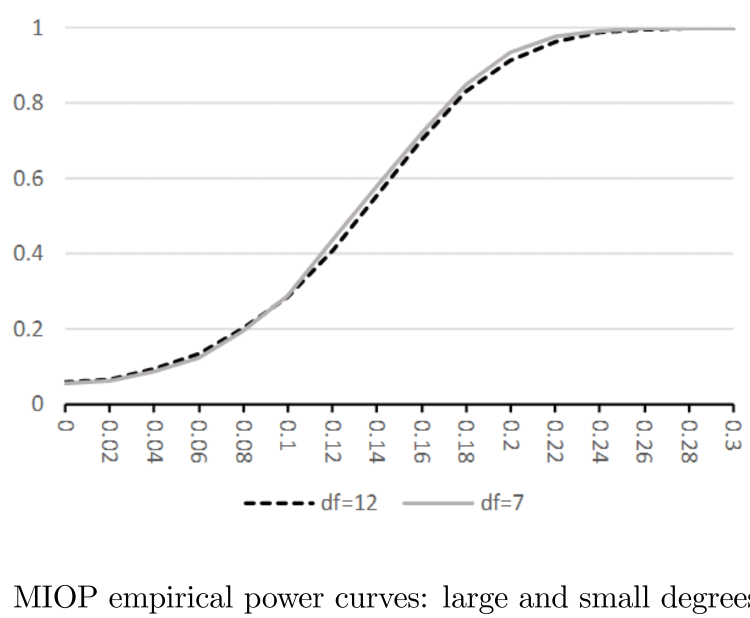

Figure C.2: MIOP empirical power curves: large and small degrees of freedom