Oberhofer, Harald; Pfaffermayr, Michael

Article

Two-part models for fractional responses defined as ratios of integers

Econometrics

Provided in Cooperation with:
MDPI – Multidisciplinary Digital Publishing Institute, Basel

Suggested Citation: Oberhofer, Harald; Pfaffermayr, Michael (2014) : Two-part models for fractional responses defined as ratios of integers, Econometrics, ISSN 2225-1146, MDPI, Basel, Vol. 2, Iss. 3, pp. 123-144, https://doi.org/10.3390/econometrics2030123

Econometrics **2014**, *2*, 123-144; doi:10.3390/econometrics2030123

OPEN ACCESS

econometrics

ISSN 2225-1146

www.mdpi.com/journal/econometrics

Article

Two-Part Models for Fractional Responses Defined as Ratios of Integers

Harald Oberhofer [1,2,*] and Michael Pfaffermayr [3,4]

[1] Department of Economics and Social Sciences, University of Salzburg, Residenzplatz 9, 5010 Salzburg, Austria

[2] The Austrian Center for Labor Economics and the Analysis of the Welfare State,c/o Rudolf Winter-Ebmer, Altenbergerstr. 69, 4040 Linz, Austria

[3] Department of Economics, University of Innsbruck, Universitaetsstrasse 15, 6020 Innsbruck, Austria; E-Mail: Michael.Pfaffermayr@uibk.ac.at

[4] Austrian Institute of Economic Research (WIFO), Objekt 20, Arsenal, 1030 Vienna, Austria

* Author to whom correspondence should be addressed; E-Mail: Harald.Oberhofer@sbg.ac.at; Tel.: +43-662-8044-623.

Received: 7 July 2014; in revised form: 28 August 2014 / Accepted: 10 September 2014
Published: 19 September 2014

Abstract: This paper discusses two alternative two-part models for fractional response variables that are defined as ratios of integers. The first two-part model assumes a Binomial distribution and known group size. It nests the one-part fractional response model proposed by Papke and Wooldridge (1996) and, thus, allows one to apply Wald, LM and/or LR tests in order to discriminate between the two models. The second model extends the first one by allowing for overdispersion in the data. We demonstrate the usefulness of the proposed two-part models for data on the 401(k) pension plan participation rates used in Papke and Wooldridge (1996).

Keywords: fractional response models for ratios of integers; one-part *versus* two-part models; Wald test; LM test; LR test

JEL classifications: C12, C15, C25, C52

---

1. Introduction

Many empirical studies deal with fractional response data that are bounded in the [0,1] interval and, in addition, contain a significant amount of observations at the boundary values of zero or one. In their seminal paper, Papke and Wooldridge (1996) [1] propose a one-part fractional response model that extends the generalized linear model (GLM) literature from statistics.[1] In particular, they introduce a quasi-maximum likelihood (QMLE) approach that only requires the correct specification of the conditional mean to consistently estimate one-part fractional response models. In this framework, there is no need for an *ad hoc* transformation of the boundary values of zero or one.

As an alternative to this QMLE approach, scholars have also proposed to assume a beta distribution and estimate the resulting model via maximum likelihood (see, e.g., Paolino 2001 [4], Kieschnick and McCullough 2003 [5], Ferrari and Cribari-Neto 2004 [6]). The beta distribution is defined in the [0,1] interval, and its two (positive) shape parameters allow for very flexible functional forms.

If the data at hand contain a large share of boundary values, the econometric literature alternatively offers two-part models that assume a different data generating process for the zeros or ones, respectively (see, e.g., Mullahy 1986 [7], Lambert 1992 [8], Cameron and Trivedi 2005 [9], Wooldridge 2002 [10], Problem 19.8, Ramalho and da Silva 2009 [11], 2013 [12], Ramalho *et al.* 2011 [3] and Oberhofer and Pfaffermayr 2012 [13]). Based on Papke and Wooldridge's (1996) [1] binomial log-likelihood approach, these two-part models typically combine a logit or probit model for the boundary value with the QMLE for the remaining non-boundary observations. In the beta regression framework, two-part models are proposed by, e.g., Cook, Kieschnick and McCullough (2008) [14] and Ospina and Ferrari (2012) [15]. Cook *et al.* (2008) [14] model the probability of zeros with the cdf of the logistic function and formulate a beta regression model for the continuous portion of their distributions. Ospina and Ferrari (2012) [15] discuss various different link functions, such as, e.g., the logit and probit links, respectively, that allow one to model the discrete part of a general class of mixed continuous-discrete distributions, which apply the beta distribution to the continuous part.

The selection of one of the competing (one-part or two-part) models for the data at hand is crucial for the reliability of the obtained results. If one, for example, neglects the existence of mass points in the data and simply estimates one-part models in the spirit of Papke and Woolridge (1996) [1], one may obtain misspecified conditional mean functions. In such a situation, the estimates of the one-part model would be biased and inconsistent. So far, the literature typically applies a P test for non-nested hypotheses, as described in Davidson and MacKinnon (1981) [16] and Ramalho, Ramalho and Murteira (2011) [3], in order to discriminate between the competing one-part and two-part fractional response models.

In many empirical applications, the fractional response variable is defined as the ratio of integers, and hence, the group size is known. The share of employees participating in a voluntary pension plan constitutes such a case (see Papke and Wooldridge 1996) [1]. In their working paper version, Papke and Wooldridge (1993) [17] explicitly discuss the case of known group size in the context of their QMLE

---

[1] In a more recent paper, Papke and Wooldridge (2008) [2] discuss fractional response models for panel data. Ramalho *et al.* (2011) [3] provide a comprehensive up-to-date overview on the econometrics of fractional response models.

approach. Accordingly, this additional group size information can explicitly be used for the empirical analysis. Starting from this observation, we propose two-part models that exploit information on the group size and, additionally, nest their one-part alternatives following the approach suggested by Lin and Schmidt (1984) [18] and Mullahy (1986) [7]. This nesting approach allows one to directly test between the competing one- and two-part models for fractional response data that are defined as ratios of integers. In this regard, the paper presents alternative testing procedures to the already available non-nested *P* test.

In line with the strands of the literature discussed above, the first two-part model is based on the binomial likelihood framework, while the second one additionally allows for overdispersion in the data by assuming a beta-binomial likelihood function. This latter model, for example, is able to account for correlated individual zero and one decisions that could be triggered by, e.g., group-specific random effects. Applying the maximum likelihood framework to both models, this paper derives explicit formulas for the Wald and the LM tests, respectively. We apply the different estimators to firm-level data on 401(k) pension plan participation rates as used in Papke and Wooldridge (1996) [1] and document that participation decisions are highly correlated within firms.

The remainder of the paper is organized as follows: Section 2 presents the nested two-part models for fractional response variables that are defined as a ratio of integers. Section 3 offers an empirical application for 401(k) plan participation rates, and in Section 4 we provide some concluding remarks. In Appendix A and B, we derive the proposed LM and Wald tests, respectively, while Appendix C provides the analytical expressions for the calculation of marginal effects in the proposed two-part models.

2. Two-Part Fractional Response Models for Responses Defined as Ratios of Integers

2.1. The Binomial Two-Part Model

The typical fractional response model is based on the Bernoulli or binomial distribution. Assume there are $i = 1, \ldots, N$ groups (e.g., firms) in which $j = 1, \ldots, n_i$ units (workers) are confronted with a $0/1$ decision (e.g., to participate in a voluntary pension plan). We focus on situations where the number of units, n_i, is observed as in Papke and Wooldridge (1993 [17], 1996 [1]) and assume that n_i is exogenously given, so that it is appropriate to condition on it. The probability that unit j in group i opts for one (e.g., to participate in a voluntary pension plan) is denoted by θ_i, which is assumed to be group-, but not unit-, specific. The number of units within a group choosing one is denoted by k_i, and the corresponding observed share (at the group level) is given by $y_i = \frac{k_i}{n_i}$ with $0 \leq y_i \leq 1$ or $k_i = n_i y_i$, respectively.[2] Following Papke and Wooldridge (1996) [1], for such a set-up, the conditional expectation of the fractional response variable y_i is group-specific and can be specified as:

$$E(y_i|x_i, n_i) = G(x_i\beta), \quad i = 1, \ldots, N \tag{1}$$

where (the $1 \times k$ vector) x_i refers to a set of i-specific explanatory variables with the corresponding parameter vector β. Typically, $G(.)$ is a cumulative distribution function (cdf), such as the logistic

[2] Note, in comparison to the fractional response model analyzed in Papke and Wooldridge (1996) [1], the individual contributions to the likelihood, the estimated score and the estimated information matrix are all multiplied by n_i in this model (see also Papke and Wooldridge, 1993, [17], pp. 10–11).

function $G(z) = \exp(z)/(1+\exp(z))$, which maps z to the $(0,1)$ interval.[3] In this case, the group-specific contributions to the log likelihood can be written as:

$$\ln(f(\beta; y_i, x_i)) = n_i(y_i \ln(G(x_i\beta)) + (1-y_i)\ln(1-G(x_i\beta))) + const \tag{2}$$

Following Wooldridge (2002, [10], Problem 19.8), Cameron and Trivedi (2005, [9], p. 680), Ramalho and da Silva (2009 [11], 2013 [12]), Ramalho *et al.* (2011) [3] and Oberhofer and Pfaffermayr (2012) [13], one may consider a two-part model to explicitly account for an excessive number of boundary values. Here, we concentrate on the case of boundary values at one, but similar arguments apply to the case of an excessive number of zeros. In contrast to the one-part model, the two-part alternative assumes a different data generating process for the boundary values. For notational simplicity, the explanatory variables in the first and second part of the model are assumed to be the same, but in general, they could differ. Formally, this two-part model can be defined as in Cameron and Trivedi (2005 [9], pp. 545, 680) and is given by:

$$f(y_i|x_i,n_i) = \begin{cases} P_1(n_i,x_i) & \text{if} \quad y_i = 1 \quad \text{or} \quad k_i = n_i \\ (1-P_1(n_i,x_i))\frac{P_2(k_i,x_i)}{1-P_2(n_i,x_i)} & \text{if} \quad y_i < 1 \quad \text{or} \quad k_i < n_i \end{cases} \tag{3}$$

where $P_1(n_i,x_i) = P_1(K_i = n_i, x_i)$ refers to the first part of the model and $P_2(k_i,x_i) = P_2(K_i = k_i, x_i)$, $k_i = 1, ..., n_i$ to its second part. Under independent unit decisions, K_i is assumed to be distributed as binomial with conditional probabilities:

$$\begin{aligned} P_1(n_i,x_i) &= \theta_{i1}^{n_i} \\ P_2(n_i y_i, x_i) &= \binom{n_i}{n_i y_i} \theta_{i2}^{n_i y_i}(1-\theta_{i2})^{n_i - n_i y_i} \end{aligned} \tag{4}$$

where for $0 < \theta_{i1} < 1$ the probability of y_i amounting exactly to one is given by $\theta_{i1}^{n_i}$ under $P_1(n_i,x_i)$ and $\theta_{i2}^{n_i}$ under $P_2(n_i,x_i)$. For the nested one-part model, it holds that $P_2(n_i,x_i) = P_1(n_i,x_i)$ or $\theta_{i1} = \theta_{i2}$, and $f(y_i|x_i,n_i)$ reduces to $P_2(k_i,x_i)$.

Under this two-part model, we specify the probability of observing a share of one by $P_1(n_i,x_i) = \theta_{i1}^{n_i}$ assuming $\theta_{i1} = G(x_i\gamma)$. The second part of the model for values $y_i < 1$ is based on the conditional distribution:

$$f(y_i|y_i<1,x_i,n_i) = (1-P_1(n_i,x_i))\frac{P_2(k_i,x_i)}{1-P_2(n_i,x_i)}$$

implying that the probability distribution $f(y_i|x_i,n_i)$ is divided by $1-G(x_i\beta)^{n_i}$ to ensure that the conditional probabilities sum up to one. The conditional mean of the two-part model, thus, is given by[4]:

$$\begin{aligned} E(y_i|x_i,n_i) &= (1-P_1(n_i,x_i))\,E(y_i|y_i<1,x_i,n_i) + P_1(n_i,x_i) \\ &= \tfrac{1-G(x_i\gamma)^{n_i}}{1-G(x_i\beta)^{n_i}}(G(x_i\beta) - G(x_i\beta)^{n_i}) + G(x_i\gamma)^{n_i} \end{aligned} \tag{5}$$

---

[3] See Ramalho *et al.* (2011 [3], 2013 [19]) for a comprehensive discussion on alternative functional forms for one-part and two-part fractional response models.

[4] In the case of zero boundary values the conditional mean of this two-part fractional response model modifies to:

$$E(y_i|x_i,n_i) = P(y_i>0|x_i,n_i)E(y_i|y_i>0,x_i,n_i) = \frac{1-(1-G(x_i\gamma))^{n_i}}{1-(1-G(x_i\beta))^{n_i}}G(x_i\beta)$$

Equation (5) shows that in case of $\gamma = \beta$, the conditional mean of this two-part model reverts to the simple one-part formulation. The standard two-part literature typically uses a simplified version of the conditional mean, which ignores the fact that the second part of the model also assigns a non-zero probability to boundary values. For example, Ramalho and da Silva (2009, [11], p. 630) specify the conditional mean $E(y_i|y_i > 0, x_i, n_i)$ as $G(x_i\beta)$.

Defining $z_i = 1$ if $y_i = 1$ and zero otherwise, the likelihood of the two part-model contains the individual contributions:

$$\begin{aligned} \ln(f(\gamma, \beta; y_i, x_i)) &= (1 - z_i)[n_i(y_i \ln(G(x_i\beta)) + (1 - y_i) \ln(1 - G(x_i\beta))) - \ln(1 - G(x_i\beta)^{n_i})] \\ &\quad + (1 - z_i) \ln(1 - G(x_i\gamma)^{n_i}) + z_i n_i \ln(G(x_i\gamma)) + constant \end{aligned} \tag{6}$$

Under this specification, maximum likelihood estimation is straight forward, since it separates into the estimation of the model explaining $P(y_i = 1|x_i, n_i)$ using all observations and the estimation of the fractional response model for the observations with $y_i < 1$ only. In the following, we assume that the distributions, upon which the one-part and two-part models are based, are correctly specified and concentrate on maximum likelihood estimation.

The main advantage of the proposed two-part model is that it nests the one-part fractional response model since, as demonstrated in Equations (5) and (6), under $\theta_{i1}^{n_i} = \theta_{i2}^{n_i}$ (or equivalently, $\gamma = \beta$), the two-part-model reverts to the one-part fractional response model.[5] In the case of x_i being the same for the one-part and the two-part model and their parameters being equal under ($\gamma = \beta$), the two models coincide and have the same likelihood functions. This hypothesis can easily be tested by an LM or a Wald test of $H_0 : \gamma = \beta$. The LM test only requires the estimation of the restricted one-part model and might therefore be preferred. If one or both parts of the two-part model contain additional explanatory variables denoted by w_{1i} and w_{2i} with parameter vectors ϕ_1 and ϕ_2, respectively, the underlying H_0 to test is $\gamma = \beta$, $\phi_1 = 0$, $\phi_2 = 0$.[6]

In Appendix A1, we derive an LM test, which is based on the estimated parameters from the one-part fractional response model that are indexed by OP. Similar to Mullahy (1986) [7], the LM test uses the parametrization $\gamma = \beta + \delta$ and tests H_0: $\delta = 0$ *vs.* H_0: $\delta \neq 0$. It is easy to calculate and is given by:

$$LM = \widehat{s}'_{\delta,OP} \left(A^{-1}(\widehat{\beta}_{OP}) + \left(B(\widehat{\beta}_{OP}) - A(\widehat{\beta}_{OP}) \right)^{-1} \right) \widehat{s}_{\delta,OP} \tag{7}$$

Thereby, $\widehat{s}_{\delta,OP} = \sum_{i=1}^{N} \widehat{C}_{i\beta,OP}(z_i - G(x_i\widehat{\beta}_{OP})^{n_i})x_i'$ and $\widehat{C}_{i\beta,OP} = n_i \frac{1-G(x_i\widehat{\beta}_{OP})}{1-G(x_i\widehat{\beta}_{OP})^{n_i}}$. $A(\widehat{\beta}_{OP}) = \sum_{i=1}^{N} \widehat{C}_{i\beta,OP}^2(1 - G(x_i\widehat{\beta}_{OP})^{n_i})G(x_i\widehat{\beta}_{OP})^{n_i}x_i'x_i$ and $B(\widehat{\beta}_{OP}) = \sum_{i=1}^{N} n_i((1 - G(x_i\widehat{\beta}_{OP})) * G(x_i\widehat{\beta}_{OP}))x_i'x_i$. Note, x_i is defined as a $1 \times k$ vector. Under standard assumptions, this LM test is asymptotically distributed as $\chi^2(k)$.

---

[5] In a related setting, Lin and Schmidt (1984) [18] derive an LM test for testing a Tobit model against the more general Cragg's (1971) [20] two-part hurdle model under normality using a similar nesting hypothesis. Mullahy (1986) [7] proposes LM and Hausman test statistics in order to discriminate between one- and two-part (hurdle) count data models.

[6] In the quasi maximum likelihood framework, the literature commonly applies non-nested P tests to discriminate between the one-part and two-part fractional response models. Following Davidson and MacKinnon (1981) [16] and Ramalho *et al.* (2011) [3], the P test for the null hypothesis that the one-part model is the true one and the two-part model is the alternative is based on an artificial regression. However, in their Propositions 4.1.2 and 4.3.2, Gourieroux, Monfort and Trognon (1984) [21] prove that under the nested parametrization, this test is not applicable.

In the case that one is interested in estimating the more complex two-part model first, a Wald test allows one to investigate the accuracy of this model. Appendix A2 derives the Wald test statistic that uses the parameter estimates of this two-part model with index TP and is given by:

$$\widehat{W} = (\hat{\gamma}_{TP} - \hat{\beta}_{TP})' \left(A(\widehat{\gamma}_{TP})^{-1} + \left(B(\widehat{\beta}_{TP}) - A(\widehat{\beta}_{TP}) \right)^{-1} \right)^{-1} (\hat{\gamma}_{TP} - \hat{\beta}_{TP}) \tag{8}$$

which is likewise asymptotically distributed as $\chi^2(k)$. When applying this Wald test, it is crucial to use the weighted form of the likelihood or to assume $n_i = n$. This is necessary, because the likelihood is not defined for a group size of $n_i = 1$.

In the case that both alternatives are estimated, simple and standard likelihood ratio (LR) tests are also available for model selection. The advantage of the LM and Wald tests are that one only needs to either estimate the one-part or the two-part model, respectively. In this regard, there might be no need for estimating both models, in the case that the proposed LM or Wald tests do not reject the corresponding null hypothesis.

2.2. The Beta-Binomial Two-Part Model

In many cases, the assumption of independent zero-one decisions of the individuals is not plausible, as there may exist pronounced overdispersion in the data. To give an example, the presence of group-specific random effects generates equicorrelation within groups (see, e.g., Heckman and Willis 1977 [22] and McCulloch and Searle 2001 [23]) and violates the independence assumption of the binomial distribution made above. Following, e.g., Heckman and Willis (1977) [22], Prentice (1986) [24], McCulloch and Searle (2001) [23] and Santos Silva and Murteira (2009) [25] for these situations, a beta-binomial model forms a plausible and tractable alternative.[7] While for such situations, one still obtains consistent estimates using the Bernoulli-QMLE for the one-part model under H_0 (see Papke and Wooldridge 1996 [1]), the two-part model based on the binomial distribution (6) cannot be estimated consistently when the beta-binomial is the true data generating process. As a result, the above proposed Wald and LM tests are also invalid in this more general setting.

Following Prentice (1986) [24], we assume that the random variable K_i is distributed as beta-binomial taking values $k_i = n_i y_i$ with probabilities:

$$\begin{aligned} P_2(K_i = k_i, x_i) &= \binom{n_i}{k_i} \int_0^1 \pi_i^{k_i + a_{i2} - 1} (1 - \pi_i)^{n_i - k_i + b_{i2} - 1} \frac{\Gamma(a_{i2} + b_{i2})}{\Gamma(a_{i2})\Gamma(b_{i2})} d\pi_i \\ &= \binom{n_i}{k_i} \frac{\Gamma(k_i + a_{i2})\Gamma(n_i - k_i + b_{i2})}{\Gamma(n_i + a_{i2} + b_{i2})} \frac{\Gamma(a_{i2} + b_{i2})}{\Gamma(a_{i2})\Gamma(b_{i2})} = \binom{n_i}{k_i} \frac{\prod_{j=0}^{k_i - 1} (a_{i2} + j) \prod_{j=0}^{n_i - k_i - 1} (b_{i2} + j)}{\prod_{j=0}^{n_i - 1} (a_{i2} + b_{i2} + j)} \end{aligned}$$

[7] The Dirichlet-multinomial distribution allows one to generalize the beta-binomial model for multivariate fractional response data with more than two alternative choices (see, e.g., Johnson, Kemp and Kotz 2005 [26], Mullahy 2010 [27], Murteira and Ramalho 2014 [28]).

This distribution results from a data generating process that is based on $K_i|\pi_{i2} \sim Bin(n_i, \pi_{i2})$ and $\pi_{i2} \sim Beta(a_{i2}, b_{i2})$. In order to parametrize the model, we specify:

$$\theta_{i2} := E[\pi_i|x_i] = \frac{a_{i2}}{a_{i2}+b_{i2}},\ c_i = a_{i2} + b_{i2}$$

or

$$a_{i2} = c_i\theta_{i2},\ b_{i2} = c_i(1-\theta_{i2}),\ \theta_{i2} = \frac{e^{x_i\beta}}{1+e^{x_i\beta}}$$

This model introduces intra-group correlation and, therefore, allows for overdispersion in the data (see Prentice 1986 [24] and McCulloch and Searle 2001 [23]). A model with group-specific random effects that follow a $Beta(a_{i2}, b_{i2})$ distribution yields the same model structure (see McCulloch and Searle 2001 [23]).

Letting $K_i = \sum_{j=1}^{n_i} K_{ij}$, where K_{ij} are correlated Bernoulli-random variables that take the value one with probability π_i and $\pi_i \sim Beta(a_{i2}, b_{i2})$, the variance and covariance are given by:

$$\begin{aligned} Var[K_i] &= n_i\frac{a_ib_i}{c_i^2} + n_i\,(n_i-1)\,\frac{a_ib_i}{c_i^2(1+c_i)} \\ &= n_i\theta_{i2}(1-\theta_{i2})(1+(n_i-1)\rho) \end{aligned} \tag{9}$$

and:

$$Cov[K_{ij}, K_{il}] = \frac{a_ib_i}{c_i^2(1+c_i)} \text{ and } \rho = corr[K_{ij}, K_{il}] = \frac{1}{1+c_i} \tag{10}$$

respectively.[8] For the boundary value $K_i = n_i$, we specify $P_1(n_i, x_i)$ analogously and assume that:

$$P_1(n_i, x_i) = \frac{\Gamma\,(n_i + a_{i1})}{\Gamma\,(n_i + a_{i1} + b_{i1})}\frac{\Gamma\,(a_{i1} + b_{i1})}{\Gamma\,(a_{i1})} \tag{11}$$

where $a_{i1} = c_i\theta_{i1}$, $b_{i1} = c_i(1-\theta_{i1})$. For $P_1(n_i)$, the nested specification is given by: $\theta_{i1} = \frac{e^{x_i\gamma}}{1+e^{x_i\gamma}} = \frac{e^{x_i(\beta+\delta)}}{1+e^{x_i(\beta+\delta)}}$.

We impose the same parameter c_i for $P_1(n_i)$ and $P_2(k_i)$, as a different value for $P_1(n_i)$ remains unidentified for zero-one dependent variables. Principally, the parameter c_i can be made dependent on explanatory variables x_i.[9] In this case, a convenient parametrization would be $c_i = (e^{z_i\vartheta} - 1)/2$, so that $\rho = \frac{2}{1+e^{z_i\vartheta}}$, where z_i denotes the vector or explanatory variables for c_i and ϑ the corresponding parameter vector (see Prentice 1986 [24]). This parametrization guarantees that the coefficient of intra-group correlation is restricted to the [0,1] interval.

Given the maximum likelihood estimates (see Appendix B for details) of the unconstrained model, it is straight forward to use Wald or likelihood ratio (LR) tests for model discrimination between the two-part and one-part models. The resulting H_0 to test would again be $H_0: \gamma = \beta$ or $\delta = 0$ *vs.* $H_1: \gamma \neq \beta$ or $\delta \neq 0$. An analytical LM test is infeasible in this context, as a simple closed form of the

---

[8] Note, this model only allows for positive intra-group correlation, as one has to assume that $a_i > 0$ and $b_i > 0$. Prentice (1986) [24] proposes a transformation of the beta-binomial distribution that is also able to handle negative intra-group correlation.

[9] In the empirical application in Section 3, we introduce one specification, where c depends on the matching rate, firm size and the age of the pension plan. For further details, see Columns (7) and (8) in Table 1 and the corresponding discussion in the text.

expected Hessian of that model cannot be derived analytically under H_0. However, in the case that one prefers to apply LM-type tests, one can use the outer product of the first derivatives as the estimate of the variance of the score.

The working paper version of this paper contains two small-scale Monte Carlo exercises, which reveal that the herein proposed LM, Wald and LR tests are all properly sized and exhibit sufficient power. Accordingly, these tests seem to be suitable for discriminating between one-part and two-part fractional response models for dependent variables that are defined as ratios of integers. For further details on the Monte Carlo exercises, see Oberhofer and Pfaffermayr (2014) [29].

3. An Empirical Application: The 401(k) Pension Plan Participation Rates

This section offers an application of the nested two-part fractional response models for fractional responses defined as ratios of integers using 401(k) pension plan participation rate data that have also been used by Papke and Wooldridge (1996) [1]. The application was coded using version 13 of Stata. The data have been taken from the online archive of the *Journal of Applied Econometrics* [30]. The Stata code is available from the authors upon request.

In order to compare our estimation results with the non-weighted one-part model proposed by Papke and Wooldridge (1996) [1], we also replicate their results. Moreover, the beta-binomial model also allows one to highlight that the 401(k) plan participation rates are characterized by non-negligible intra-firm correlation. Finally, we document the usefulness of the proposed Wald, LM and LR tests for discriminating between one-part and two-part fractional response models.

In their empirical application, Papke and Wooldridge (1996) [1] model the participation in 401(k) pension plans using a sample of 4734 U.S. manufacturing firms. The dependent variable (PRATE) is measured as the fraction of active 401(k) pension plan accounts relative to the overall number of eligible employees, which amounts to one in 42.73 percent of all observations (*i.e.*, 2023 firms). The vector of covariates contains a firm's matching rate (MRATE), the firms overall number of employees (log (EMP)), the pension plan's age (AGE), as well as an indicator variable (SOLE) that takes on the value of one if the 401(k) pension plan is the only one offered by the firm. In their most general specification, Papke and Wooldridge (1996) [1] include squared terms of the former three covariates in order to control for additional non-linearities. Further details on different specifications and sub-sample results can be found in Papke and Wooldridge (1996) [1].

Table 1 reports the parameter estimates for various different fractional response models. To start with, Column (1) replicates the results from Column (4) of Table III in Papke and Wooldridge (1996) [1]. These results are based on the non-employment weighted QMLE estimator for the one-part fractional response model. In Column (2), we apply the same QMLE estimator, but additionally weight the observations by each firm's number of employees. Columns (3) and (4) report the results from the binomial two-part fractional response model. Thereby, Column (4) reports the restricted model results where the parameters are the same in both parts of the model. Note, Column (4) contains the same parameter estimates as Column (2), but the standard errors are much smaller. The reason is that the latter are MLE-estimates under the assumption of known group size and independent unit decisions within groups (*i.e.*, an absence of overdispersion in the data). Papke and Wooldridge (1996) [1] calculate

robust standard errors to account for potential overdispersion in the data. Finally, Columns (5) to (8) report estimation results from the beta-binomial fractional response model that explicitly allows for overdispersion in the data in a flexible MLE-setting.

Columns (5) to (8) in Table 1 highlight the existence of non-negligible intra-group correlation in the data on 401(k) pension plan participation. The estimated (average) ρ varies from 0.131 to 0.404 and (as indicated in Columns 5 and 8) is also affected by a firm's characteristics. An increase in a firm's matching rate decreases c_i and, thus, increases the intra-group correlation ρ, as can be seen from Equation (10). This finding is well in line with our expectations indicating that a larger matching rate increases the intra-group correlation, making 100 percent participation more likely. In a similar vein, a larger firm size also increases ρ. By contrast, in firms that offer older pension plans, the intra-group correlation is reduced. Overall, for the 401(k) pension plan rate data, the significant estimates for the intra-group correlation indicate that the beta-binomial model should be preferred over the more restrictive binomial model, which fails to account for this correlation. In addition, as compared to the two-part alternatives, the one-part models overestimate intra-group correlation considerably.

At the bottom of Table 1, we report the results from the alternative LM, LR and Wald tests. All of them indicate that the one-part models should be rejected in favor of the two-part fractional response alternatives. Given the discussion on the parameter estimates for the SOLE dummy variable, this result is not very surprising. This, together with the discussion on the intra-group correlation from above, suggests that the beta-binomial model might be most accurate for estimating the 401(k) plan participation rate data at hand.

In order to assess and compare the overall goodness-of-fit of the alternative models displayed in Table 1, we also report an *R*-squared measure based on Nagelkerke (1991) [31] defined as $R^2 = \left(1 - \left(\frac{L(0)}{L(\widehat{\theta})}\right)^{2/n}\right) / (1 - L(0)^{2/n})$, where the likelihoods enter in levels rather than in logs. This guarantees that the R^2 is bounded by zero and one, respectively. The corresponding numbers are reported at the bottom of Table 1. Since these numbers are not fully comparable across different model types, we restrict our discussion to the comparison of the restricted one-part models with their two-part alternatives. Starting with the more restrictive binomial model, the Nagelkerke *R*-squared measure, which corresponds to the one-part model, is given by only 0.056, while for the two-part alternative, we obtain a goodness-of-fit measure amounting to 0.216. In a similar vein, for both beta-binomial models reported in Columns (5) to (8), the two-part alternatives provide a much better fit for the data. Accordingly, the Nagelkerke *R*-squared measures amount to 0.569 and 0.576 for the two-part models, respectively, while for both (restricted) one-part models, the corresponding fit is below 0.240. To sum up, the Nagelkerke *R*-squared measures clearly indicate that the two-part models provide a much better fit for the 401(k) pension plan participation rates data. In this regard, the goodness-of-fit statistics confirm the results obtained from the proposed nested tests, which commonly reject the one-part models in favor of their two-part alternatives.

Table 1. Estimation results: 401(k) plan participation rates.

Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
First Part								
$MRATE$			1.313 ∗ ∗	1.372 ∗ ∗	1.349 ∗ ∗	1.779 ∗ ∗	1.203 ∗ ∗	1.220 ∗ ∗
			(0.083)	(0.002)	(0.085)	(0.085)	(0.097)	(0.068)
$MRATE^2$			−0.221 ∗ ∗	−0.290 ∗ ∗	−0.228 ∗ ∗	−0.312 ∗ ∗	−0.246 ∗ ∗	−0.288 ∗ ∗
			(0.022)	(0.001)	(0.022)	(0.022)	(0.024)	(0.018)
$\log(EMP)$			0.403 ∗ ∗	−0.602 ∗ ∗	−0.203∗	−0.597 ∗ ∗	−0.310 ∗ ∗	−0.674 ∗ ∗
			(0.111)	(0.004)	(0.115)	(0.111)	(0.131)	(0.075)
$\log(EMP)^2$			0.029 ∗ ∗	0.029 ∗ ∗	0.015 ∗ ∗	0.032 ∗ ∗	0.002	0.035 ∗ ∗
			(0.007)	(0.000)	(0.007)	(0.007)	(0.011)	(0.005)
AGE			−0.004	0.058 ∗ ∗	−0.004	0.006	0.011	0.044 ∗ ∗
			(0.009)	(0.000)	(0.009)	(0.009)	(0.010)	(0.006)
AGE^2			0.000∗	−0.001 ∗ ∗	0.000∗	0.000	0.000 ∗ ∗	−0.000 ∗ ∗
			(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$SOLE$			0.389 ∗ ∗	0.053 ∗ ∗	0.400 ∗ ∗	0.416 ∗ ∗	0.416 ∗ ∗	0.118 ∗ ∗
			(0.047)	(0.002)	(0.048)	(0.048)	(0.050)	(0.035)
$CONSTANT$			1.944 ∗ ∗	3.466 ∗ ∗	3.347 ∗ ∗	3.532 ∗ ∗	3.748 ∗ ∗	3.307 ∗ ∗
			(0.425)	(0.018)	(0.437)	(0.428)	(0.501)	(0.298)
Second Part/One Part								
$MRATE$	1.665 ∗ ∗	1.372 ∗ ∗	0.514 ∗ ∗	1.372 ∗ ∗	0.936 ∗ ∗	1.779 ∗ ∗	0.942 ∗ ∗	1.220 ∗ ∗
	(0.089)	(0.169)	(0.003)	(0.002)	(0.079)	(0.085)	(0.082)	(0.068)
$MRATE^2$	−0.332 ∗ ∗	−0.290 ∗ ∗	−0.178 ∗ ∗	−0.290 ∗ ∗	−0.226 ∗ ∗	−0.312 ∗ ∗	−0.238 ∗ ∗	−0.288 ∗ ∗
	(0.021)	(0.041)	(0.001)	(0.001)	(0.019)	(0.022)	(0.022)	(0.018)
$\log(EMP)$	−1.031 ∗ ∗	−0.602 ∗ ∗	−0.629 ∗ ∗	−0.602 ∗ ∗	−0.594 ∗ ∗	−0.597 ∗ ∗	−0.607 ∗ ∗	−0.674 ∗ ∗
	(0.112)	(0.256)	(0.004)	(0.004)	(0.088)	(0.111)	(0.088)	(0.075)
$\log(EMP)^2$	0.054 ∗ ∗	0.029 ∗ ∗	0.032 ∗ ∗	0.029 ∗ ∗	0.032 ∗ ∗	0.032 ∗ ∗	0.032 ∗ ∗	0.035 ∗ ∗
	(0.007)	(0.013)	(0.000)	(0.000)	(0.006)	(0.007)	(0.006)	(0.005)
AGE	0.055 ∗ ∗	0.058 ∗ ∗	0.061 ∗ ∗	0.058 ∗ ∗	0.048 ∗ ∗	0.006	0.050 ∗ ∗	0.044 ∗ ∗
	(0.008)	(0.008)	(0.000)	(0.000)	(0.006)	(0.009)	(0.006)	(0.006)

Table 1. *Cont.*

Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
AGE^2	$-0.001**$	$-0.001**$	$-0.001**$	$-0.001**$	$-0.001**$	0.000	$-0.001**$	$-0.000**$
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$SOLE$	0.064	0.053	$-0.170**$	$0.053**$	$-0.120**$	$0.416**$	$-0.107**$	$0.118**$
	(0.047)	(0.127)	(0.002)	(0.002)	(0.038)	(0.048)	(0.037)	(0.035)
$CONSTANT$	$5.105**$	$3.466**$	$3.376**$	$3.466**$	$3.040**$	$3.532**$	$3.119**$	$3.307**$
	(0.431)	(1.204)	(0.019)	(0.018)	(0.336)	(0.428)	(0.333)	(0.298)
Intra-group correlation								
$MRATE$							$-0.202**$	$-0.191**$
							(0.037)	(0.011)
$\log(EMP)$							$-0.236**$	$-0.088**$
							(0.028)	(0.009)
AGE							$0.014**$	$0.008**$
							(0.003)	(0.002)
$CONSTANT$					$2.648**$	$1.517**$	$3.315**$	$1.440**$
					(0.026)	(0.023)	(0.124)	(0.070)
Tests								
LM/LR Test ($\chi^2(8)$)			204.46 ***		2702.33 ***		489.28 ***	
Robust LM Test ($\chi^2(8)$)			977.72 ***					
Wald Test ($\chi^2(8)$)			82322.39 ***		5020.01 ***		619.41 ***	
ρ					0.132	0.360	0.131	0.404
Nagelkerke-R^2	0.109	0.056	0.216	0.056	0.569	0.237	0.576	0.223
Observations	4734	4734	4734	4734	4734	4734	4734	4734

Notes: Parameter estimates are reported. *, **, *** denote statistically significant at 10%-, 5%- and 1%-level, respectively. The results in Column (1) are identical to Column (4) of Table 3 reported in Papke and Wooldridge (1996) [1]. In the logit model, the dependent variable amounts to one if all employees participate in the 401(k) pension plan and zero otherwise. The QMLE of the two-part model is estimated only for $PRATE < 1$. LM tests are calculated for the binomial models in Columns (3) and (4), while the LR tests are reported for the beta-binomial models from Columns (5) to (8).

For an overall quantitative comparison of the impacts of all covariates across the different econometric models, one has to focus on marginal effects. In the following, we provide two alternatives on how to assess the quantitative differences across the competing models. First, we again follow Papke and Wooldridge (1996) [1] and plot the predicted participation rates for different matching rates between zero and one. Thereby, we set firm size and the pension plan age at their median values of 628 employees and eight years, respectively, and assume that no other pension plan is offered (*i.e.*, SOLE = 0). Finally, we vary the matching rate from zero (*i.e.*, no matching offered at all) to one (*i.e.*, 100 percent matching). Figure 1 displays the predicted participation rates and the respective 95 percent pointwise confidence intervals for the QMLE proposed by Papke and Wooldridge (1996) [1] and both alternative two-part models. The left panel compares the predictions from the QMLE (Column 1 of Table 1) with the ones from the two-part binomial alternative (Column 3). In the right panel, the alternative predictions are based on the parameters from the two-part beta-binomial model that parametrizes the intra-group correlation with MRATE, log(EMP) and AGE (Column 7). The 95 percent confidence intervals are constructed using the delta method.

Second, we calculate marginal effects based on the analytical expressions delivered in Appendix C. These marginal effects are calculated for each individual firm, and Table 2 reports the average marginal effects (AMEs) of MRATE for the deciles of the MRATE distribution for firms with MRATE ≤ 1. For small values of MRATE, the reported AMEs differ most substantially across the alternative models. To give an example, based on the two-part binomial model, the AME for the first decile of MRATE is given by 0.163, while the two-part beta-binomial model (from Column 7 in Table 1) delivers an AME of 0.242. In general, both alternative one-part models lead to rather similar AMEs, while the two-part alternatives result in rather different estimates. As discussed above, this might be driven by the strong assumption of independent decisions upon which the binomial model relies. Comparing both one-part models with the preferred beta-binomial two-part model, one can easily see that the one-part models result in lower marginal effects estimates. To sum up, the beta-binomial model, which explicitly accounts for both the mass points of ones, as well as intra-firm correlation (in the workers' participation decisions), results in a more significant and positive (estimated) impact of MRATE for the overall share of participants in 401(k) pension plans.

Figure 1 reveals some interesting results: First of all, in the absence of any matching (*i.e.*, MRATE = 0), the predicted conditional mean of PRATE is lowest (highest) for the QMLE (two-part binomial model). This difference is statistically significant, as indicated by the respective confidence intervals. Second, as can be seen from the slopes of the corresponding curves, the predicted marginal effect of MRATE, by contrast, is lowest for the two-part binomial model. Most interestingly, however, the QMLE and the (ML based) two-part beta-binomial model deliver relatively similar predictions. This is indicated by the overlap of the respective 95 percent confidence intervals over the whole range of matching rates considered. This finding highlights the usefulness of the QMLE estimator, but also points to the relevance of the (two-part) beta-binomial model as a valuable alternative. The QMLE is very easy to implement, but the two-part beta-binomial estimator allows one to explicitly specify intra-group correlation, thus providing additional (parametric) insights in the data generating process.

Figure 1. Participation rate *versus* matching rate: model predictions for firms with MRATE < 1: estimated marginal effects and 95% pointwise confidence intervals.

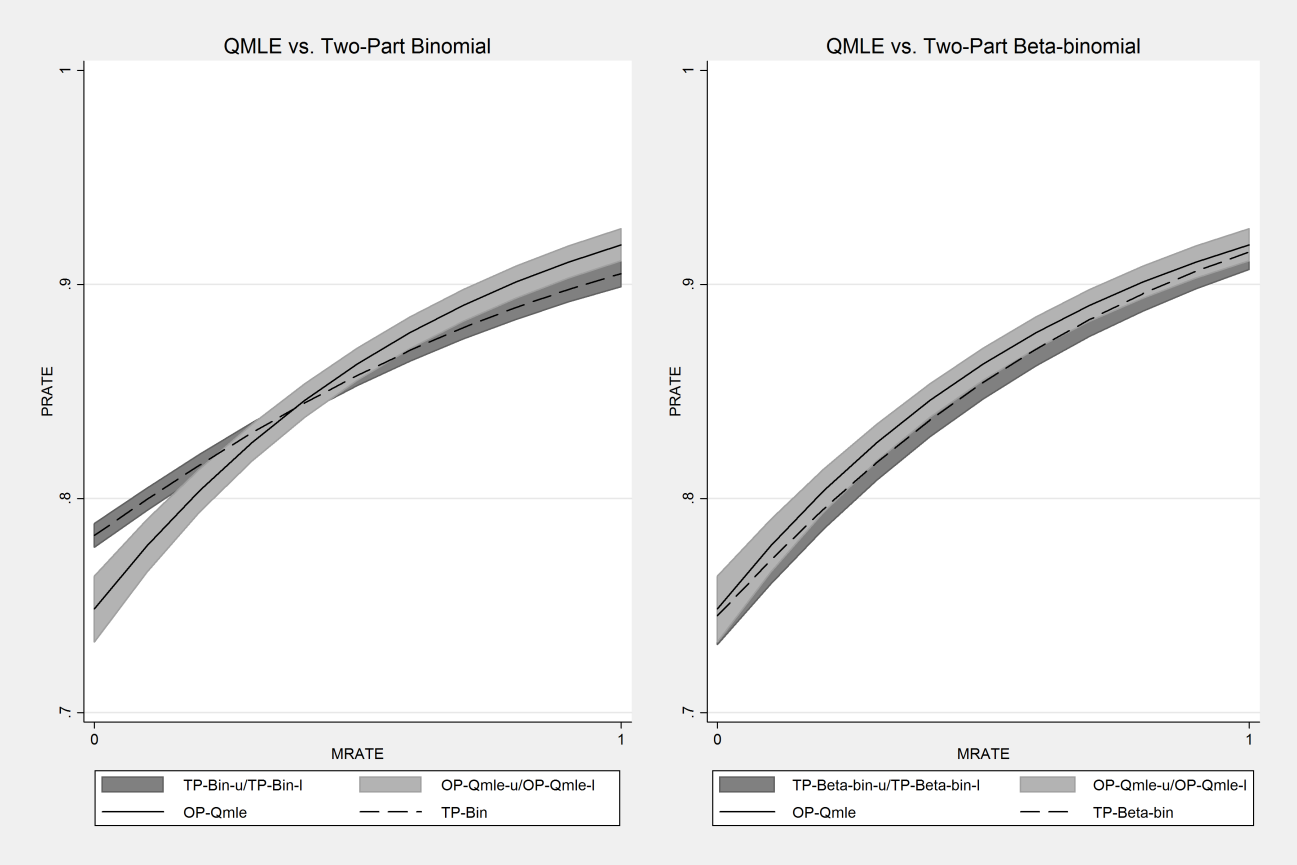

Table 2. Marginal effect estimates for MRATE.

Decile	Obs.	MRATE	TP-Bin	OP-QMLE	TP-Beta-bin	OP-Beta-bin
1	379	0.091	0.163	0.213	0.242	0.209
2	378	0.173	0.156	0.198	0.228	0.194
3	379	0.228	0.148	0.181	0.212	0.177
4	378	0.295	0.139	0.169	0.199	0.168
5	378	0.351	0.131	0.153	0.184	0.151
6	379	0.399	0.127	0.145	0.175	0.140
7	378	0.457	0.116	0.129	0.156	0.123
8	379	0.540	0.106	0.115	0.140	0.109
9	378	0.675	0.087	0.092	0.110	0.083
10	378	0.872	0.066	0.068	0.080	0.060
Total/Average	3.784	0.408	0.124	0.146	0.172	0.141

Notes: Average marginal effects (AMEs) and the average MRATE within each decile are reported. OP-QMLE and TP-Bin AMEs refer to Columns (3) and (4) in Table 1, respectively. AMEs reported for the beta-binomial models are based on the most general specification depicted in Columns (7) and (8) of Table 1.

4. Conclusions

In many applications of fractional response models, the number of units per group is observed, and consequently, the fractional response variable is based on a fraction of integers. In such a situation, one can use this additional information and specify two-part models that nest the one-part fractional response model proposed by Papke and Wooldridge (1996) [1] and account for intra-group correlation and overdispersion in the data induced by group-specific random effects.

These nested two-part models also have the advantage that they allow one to apply simple LM, LR and Wald tests to discriminate between one-part and two-part fractional response models. Based on the proposed two-part models, this paper also derives explicit formulas for the Wald and the LM tests. We apply our alternative estimators and tests to a sample of 401(k) pension plan participation rates and are able to show that the one-part models are rejected in favor of their two-part alternatives and that the data at hand are characterized by non-negligible intra-group correlation. Based on the derived nested tests and the goodness-of-fit of the alternative models, we are able to conclude that a two-part model based on the beta-binomial distribution most accurately explains the 401(k) pension plan participation rates data used in Papke and Wooldridge (1996) [1].

Acknowledgments

We would like to thank Joaquim Ramalho and three anonymous referees for valuable comments on a previous version of this paper.

Author Contributions

The authors contributed jointly to the paper.

Appendix

A. LM and Wald Tests for the Binomial Two-Part Model

A.1. Derivation of the LM Test

To derive the LM test, we reparametrize the model and set $\gamma = \beta + \delta$. Then, the likelihood function is given by:

$$\begin{aligned}&\textstyle\sum_{i=1}^{N} \ln(f(\beta; y_i, x_i)) \\ &= (1 - z_i)\big[\ln(1 - G(x_i\,(\beta + \delta))^{n_i}) + n_i(y_i \ln G(x_i\beta) + (1 - y_i)\ln(1 - G(x_i\beta))) \\ &- \ln(1 - G(x_i\beta)^{n_i}) + const\big] + z_i.\big[n_i \ln(G(x_i\,(\beta + \delta)))\big]\end{aligned}$$

To derive the score, define $z_i = 0$ if $y_i < 0$ and $z_i = 1$ if $y_i = 1$ and denote $G_i' = g_i$. To simplify the notation, we first assume that the model only contains a constant ($x_i = 1$) and introduce the vector of explanatory variables below. The score is given by:

$$\begin{aligned}
\frac{\partial \ln(l(\delta,\beta))}{\partial \delta} &= (1-z_i)\frac{-n_i G(\beta+\delta)^{n_i-1} g(\beta+\delta)}{1-G(\beta+\delta)^{n_i}} + z_i \frac{n_i g(\beta+\delta)}{G(\beta+\delta)} \\
\frac{\partial \ln(l(\delta,\beta))}{\partial \beta} &= (1-z_i)\frac{-n_i G(\beta+\delta)^{n_i-1} g(\beta+\delta)}{1-G(\beta+\delta)^{n_i}} \\
&\quad +(1-z_i) n_i \left(\frac{g(\beta) y_i}{G(\beta)} - \frac{g(\beta)(1-y_i)}{1-G(\beta)} \right) \\
&\quad +(1-z_i)\frac{n_i G(\beta)^{n_i-1} g(\beta)}{1-G(\beta)^{n_i}} \\
&\quad +z_i \frac{n_i g(\beta+\delta)}{G(\beta+\delta)}
\end{aligned}$$

Now, define $C_{i\delta} = n_i \frac{1-G(\delta)}{1-G(\delta)^{n_i}}$, $C_{i\beta+\delta} = n_i \frac{1-G(\beta+\delta)}{1-G(\beta+\delta)^{n_i}}$ and observe that $G_i = \frac{e^\beta}{1+e^\beta}$ and $g_i = \frac{e^\beta\left(1+e^\beta\right)-e^\beta e^\beta}{\left(1+e^\beta\right)^2} = \frac{e^\beta}{1+e^\beta}\frac{1}{1+e^\beta} = G_i(1-G_i)$. Inserting and simplifying yields:

$$\begin{aligned}
s_{i\delta} &= \frac{\partial \ln(l(\delta,\beta))}{\partial \delta} = C_{i\beta+\delta}\left(z_i - G(\beta+\delta)^{n_i}\right) \\
s_{i\beta} &= \frac{\partial \ln(l(\delta,\beta))}{\partial \beta} = n_i\left(y_i - G(\beta)\right) + (1-z_i)\left[G(\beta+\delta)^{n_i} C_{i\beta+\delta} - G(\beta)^{n_i} C_{i\beta}\right]
\end{aligned}$$

where we use $y_i = 1$ if $z_i = 1$. The Hessian thus can be derived as:

$$\begin{aligned}
\frac{\partial^2 \ln(l(\delta,\beta))}{\partial \delta^2} &= \frac{\partial C_{i\beta+\delta}}{\partial \delta}\left(z_i - G(\beta+\delta)^{n_i}\right) - C_{i\beta+\delta}^2\left(1-G(\beta+\delta)^{n_i}\right) G(\beta+\delta)^{n_i} \\
\frac{\partial \ln(l(\delta,\beta))}{\partial \delta \partial \beta} &= \frac{\partial C_{i\beta+\delta}}{\partial \beta}\left(z_i - G(\beta+\delta)^{n_i}\right) - C_{i\beta+\delta}^2\left(1-G(\beta+\delta)^{n_i}\right) G(\beta+\delta)^{n_i} \\
\frac{\partial^2 \ln(l(\delta,\beta))}{\partial \beta^2} &= -n_i\left(\left(1-G(\beta)\right) G(\beta)\right) + \frac{\partial (1-z_i)\left[G(\beta+\delta)^{n_i} C_{i\beta+\delta} - G(\beta)^{n_i} C_{i\beta}\right]}{\partial \beta}
\end{aligned}$$

so that under $H_0: \ \delta = 0$, one obtains:

$$\begin{aligned}
E[H_i]|_{\delta=0} &= -\begin{bmatrix} C_{i\beta}^2\left(1-G(\beta)^{n_i}\right) G(\beta)^{n_i} & C_{i\beta}^2\left(1-G(\beta)^{n_i}\right) G(\beta)^{n_i} \\ C_{i\beta}^2\left(1-G(\beta)^{n_i}\right) G(\beta)^{n_i} & n_i\left(\left(1-G(\beta)\right) G(\beta)\right) \end{bmatrix} \\
&:= -\begin{bmatrix} A_i & A_i \\ A_i & B_i \end{bmatrix}
\end{aligned}$$

Now, introducing a $(1 \times k)$ vector of explanatory variables x_i for unit i under $H_0: \ \delta = 0$ yields:

$$\begin{aligned}
s_{i\delta} &= C_{i\beta}\left(z_i - G(x_i\beta)^{n_i}\right) x_i' \\
s_{i\beta} &= n_i\left(y_i - G(x_i\beta)\right) x_i' = 0
\end{aligned}$$

We define:

$$\begin{aligned} s_\delta &= \sum_{i=1}^{N} s_{i\delta} \\ s_\beta &= \sum_{i=1}^{N} s_{i\beta} \\ A(\beta) &= \sum_{i=1}^{N} C_{i\beta}^2 G(x_i\beta)^{n_i} \left(1 - G(x_i\beta)^{n_i}\right) x_i' x_i \\ B(\beta) &= \sum_{i=1}^{N} \left[n_i \left(G(x_i\beta)(1 - G(x_i\beta))\right) x_i' x_i\right] \end{aligned}$$

$$I(\beta) = \begin{bmatrix} A & A \\ A & B \end{bmatrix} \text{ and } I(\beta)^{-1} = \begin{bmatrix} A^{-1}\left(I + A\left(B - A\right)^{-1}\right) AA^{-1} & -\left(B - A\right)^{-1} \\ -\left(B - A\right)^{-1} & \left(B - A\right)^{-1} \end{bmatrix}$$

Note $A^{-1}\left(A + A\left(B - A\right)^{-1} A\right) A^{-1} = A^{-1}\left(I + A\left(B - A\right)^{-1}\right) = A^{-1} + \left(B - A\right)^{-1}$. The estimated LM statistic uses the estimated parameters from the one-part model with index OP and is given by:

$$\begin{aligned} \widehat{LM} &= \begin{bmatrix} \widehat{s}_{\delta_{OP}}' & 0 \end{bmatrix} \begin{bmatrix} A(\widehat{\beta}_{OP})^{-1} + \left(B(\widehat{\beta}_{OP}) - A(\widehat{\beta}_{OP})\right)^{-1} & -\left(B(\widehat{\beta}_{OP}) - A(\widehat{\beta}_{OP})\right)^{-1} \\ -\left(B(\widehat{\beta}_{OP}) - A(\widehat{\beta}_{OP})\right)^{-1} & \left(B(\widehat{\beta}_{OP}) - A(\widehat{\beta}_{OP})\right)^{-1} \end{bmatrix} \begin{bmatrix} \widehat{s}_{\delta_{OP}} \\ 0 \end{bmatrix} \\ &= \widehat{s}_{\delta_{OP}}' \left(A(\widehat{\beta}_{OP})^{-1} + \left(B(\widehat{\beta}_{OP})^{-1} - A(\widehat{\beta}_{OP})\right)^{-1}\right) \widehat{s}_{\delta_{OP}} \end{aligned}$$

which is asymptotically distributed as $\chi^2(k)$.

A.2. Derivation of the Wald Test

For the Wald test, consider the model in original parametrization, so that the following restriction is tested:

$$\begin{bmatrix} I_k & -I_k \end{bmatrix} \begin{bmatrix} \gamma \\ \beta \end{bmatrix} = \gamma - \beta = 0$$

The expected Hessian of the unrestricted model can be derived similarly as above and, for observation i, reads as:

$$E[H_i] = -\begin{bmatrix} C_{i\gamma}^2 \left(1 - G(x_i\gamma)^{n_i}\right) G(x_i\gamma)^{n_i} & 0 \\ 0 & n_i \left(\left(1 - G(x_i\beta)\right) G(x_i\beta)\right) - C_{i\beta}^2 \left(1 - G(x_i\beta)^{n_i}\right) G(x_i\beta)^{n_i} \end{bmatrix}$$

The Wald test uses the estimated parameters of the two-part model that are indexed by TP. Denoting the estimated variance covariance matrix of $\hat{\gamma_{TP}}$ by $\hat{V}_{\gamma_{TP}}$ and that of $\hat{\beta_{TP}}$ by $\hat{V}_{\beta_{TP}}$, it can easily be shown that:

$$\hat{V}_{\gamma_{TP}} + \hat{V}_{\beta_{TP}} = A(\widehat{\gamma}_{TP})^{-1} + \left(B(\widehat{\beta}_{TP}) - A(\widehat{\beta}_{TP})\right)^{-1}$$

and:

$$\widehat{W} = (\hat{\gamma}_{TP} - \hat{\beta}_{TP})' \left(A(\widehat{\gamma}_{TP})^{-1} + \left(B(\widehat{\beta}_{TP}) - A(\widehat{\beta}_{TP})\right)^{-1}\right)^{-1} (\hat{\gamma}_{TP} - \hat{\beta}_{TP})$$

Asymptotically, this Wald test is also $\chi^2(k)$ distributed.

B. The Beta-Binomial Two-Part Fractional Response Model

In the two-part model, the probability for $k_i = y_i n_i \leq n_i$ successes in n_i trials can be rewritten as:

$$g(y_i|x_i, n_i) = \begin{cases} P_1(n_i, x_i) & \text{if} \quad y_i = 1 \\ (1 - P_1(n_i, x_i)) \frac{P_2(k_i, x_i)}{1 - P_2(n_i, x_i)} & \text{if} \quad y_i < 1 \end{cases}$$

Defining $\Delta \ln \Gamma(y, a) = \ln \Gamma(y + a) - \ln \Gamma(a)$, one can write:

$$\begin{aligned} \ln P_2(k_i) &= \ln \binom{n_i}{k_i} + \Delta \ln \Gamma(k_i, a_i) + \Delta \ln \Gamma(n_i - k_i, b_i) - \Delta \ln \Gamma(n_i, a_i + b_i) \\ &= \ln \binom{n_i}{k_i} + \Delta \ln \Gamma(k_i, c_i \theta_{i2}) + \Delta \ln \Gamma(n_i - k_i, c_i(1 - \theta_{i2})) - \Delta \ln \Gamma(n_i, c_i) \end{aligned}$$

$P_1(n, x_i)$ is defined analogously:

$$\begin{aligned} P_1(n_i) &= \frac{\Gamma(n_i + a_{i1})}{\Gamma(n_i + a_{i1} + b_{i1})} \frac{\Gamma(a_{i1} + b_{i1})}{\Gamma(a_{i1})} = \frac{\prod_{j=0}^{n_i - 1} (a_{i1} + j)}{\prod_{j=0}^{n_i - 1} (a_{i1} + b_{i1} + j)} = \frac{\prod_{j=0}^{n_i - 1} (c_i \theta_{i1} + j)}{\prod_{j=0}^{n_i - 1} (c_i + j)} \\ \ln P_1(n_i) &= \sum_{j=0}^{n_i - 1} \ln(c_i \theta_{i1} + j) - \sum_{j=0}^{n_i - 1} \ln(c_i + j) \end{aligned}$$

Similarly,

$$\ln P_2(n_i) = \sum_{j=0}^{n_i - 1} \ln(c_i \theta_{i2} + j) - \sum_{j=0}^{n_i - 1} \ln(c_i + j)$$

In order to drive the likelihood of the model, we define $z_i = 1[k_i = n_i] = 1[y_i = 1]$. Then, the contribution of each group i to the likelihood is:

$$\begin{aligned} &\ln L_i = (1 - z_i)\big[\ln(1 - P_1(\theta_{i1}, c_i, n_i)) + \ln P_2(\theta_{i2}, c_i, k_i) \\ &- \ln(1 - P_2(\theta_{i2}, c_i, n_i)) + const\big] + z_i\big[\ln(P_1(\theta_{i1}, c_i, n_i))\big] \\ &:= L_{i1} + L_{i2} \\ &\text{with} \\ &L_{i1} = (1 - z_i) \ln(1 - P_1(\theta_{i1}, c_i, n_i)) + z_i \ln(P_1(\theta_{i1}, c_i, n_i)) \\ &L_{i2} = -(1 - z_i) \ln(1 - P_2(\theta_{i2}, c_i, n_i)) + (1 - z_i) \ln P_2(\theta_{i2}, c_i, k_i)) \end{aligned}$$

Note, under $H_0: \quad P_1(\theta_{i1}, c_i, n_i) = P_2(\theta_{i2}, c_i, n_i)$, the contribution of each group i to the likelihood reduces to:

$$\ln L_i^1 = (1 - z_i) \ln(1 - P_2(\theta_{i2}, c_i)) + z_i \ln(P_2(\theta_{i2}, c_i))$$

We assume $\theta_{i2} = G(x_i \beta)$, while $\theta_{i1} = G(x_i(\beta + \delta)$. In order to derive the score, one needs the first derivatives of $\ln P_1(\theta_{i1}, c_i, n_i)$, $\ln P_2(\theta_{i2}, c_i, n_i)$ and $\ln P_2(\theta_{i2}, c_i, k_i)$. Using:

$$\begin{aligned} \tfrac{\partial}{\partial \ln p} \ln(1 - e^{\ln p}) &= -\tfrac{e^{\ln p}}{1 - e^{\ln p}} = -\tfrac{p}{1-p} \\ \tfrac{\partial}{\partial \ln p} \tfrac{z - e^{\ln p}}{1 - e^{\ln p}} &= \tfrac{-e^{\ln p}\left(1 - e^{\ln p}\right) + (z - e^{\ln p}) e^{\ln p}}{\left(1 - e^{\ln p}\right)^2} = e^{\ln p} \tfrac{-\left(1 - e^{\ln p}\right) + (z - e^{\ln p})}{\left(1 - e^{\ln p}\right)^2} \\ &= e^{\ln p} \tfrac{-1 + e^{\ln p} + z - e^{\ln p}}{\left(1 - e^{\ln p}\right)^2} = \tfrac{z-1}{\left(1 - e^{\ln p}\right)^2} e^{\ln p} = \tfrac{(z-1)p}{(1-p)^2} \end{aligned}$$

one can write:

$$\begin{aligned}
\frac{\partial \ln L_i}{\partial \delta} &= \left[(1-z_i)\frac{-P_1(\theta_{i1}(\beta,\delta),c_i,n_i)}{1-P_1(\theta_{i1}(\beta,\delta),c_i,n_i)} + z_i\right]\frac{\partial \ln P_1(\theta_{i1}(\beta,\delta),c_i,n_i)}{\partial \theta_{i1}}\frac{\partial \theta_{i1}(\beta,\delta)}{\partial \delta} \\
&= \left[\frac{-(1-z_i)P_1(\theta_{i1}(\delta),c_i,n_i)+z_i(1-P_1(\theta_{i1}(\delta),c_i,n_i))}{1-P_1(\theta_{i1}(\delta),c_i,n_i)}\right]\frac{\partial \ln P_1(\theta_{i1}(\beta,\delta),c_i,n_i)}{\partial \theta_{i1}}\frac{\partial \theta_{i1}(\beta,\delta)}{\partial \delta} \\
&= \left[\frac{z_i-P_1(\theta_{i1}(\beta,\delta),c_i,n_i))}{1-P_1(\theta_{i1}(\beta,\delta),c_i,n_i)}\right]\frac{\partial \ln P_1(\theta_{i1}(\beta,\delta),c_i,n_i)}{\partial \theta_{i1}}\frac{\partial \theta_{i1}(\beta,\delta)}{\partial \delta} \\
\frac{\partial \ln L_i}{\partial \beta} &= \left[\frac{z_i-P_1(\theta_{i1}(\beta,\delta),c_i,n_i))}{1-P_1(\theta_{i1}(\beta,\delta),c_i,n_i)}\right]\frac{\partial \ln P_1(\theta_{i1}(\beta,\delta),c_i,n_i)}{\partial \theta_{i1}}\frac{\partial \theta_{i1}(\beta,\delta)}{\partial \beta} \\
&\quad - \left[\frac{z_i-P_2(\theta_{i2}(\beta),c_i,n_i))}{1-P_2(\theta_{i2}(\beta),c_i,n_i)}\right]\frac{\partial \ln P_2(\theta_{i2}(\beta),c_i,n_i)}{\partial \theta_{i2}}\frac{\partial \theta_{i2}(\beta)}{\partial \beta} \\
&\quad + \left((1-z_i)\frac{\partial \ln P_2(\theta_{i2}(\beta),c_i,k_i)}{\partial \theta_{i2}} + z_i\frac{\partial \ln P_2(\theta_{i2}(\beta),c_i,n_i)}{\partial \theta_{i2}}\right)\frac{\partial \theta_{i2}(\beta)}{\partial \beta} \\
\frac{\partial \ln L_i}{\partial c_i} &= \left[\frac{z_i-P_1(\theta_{i1}(\beta,\delta),c_i,n_i))}{1-P_1(\theta_{i1}(\beta,\delta),c_i,n_i)}\right]\frac{\partial \ln P_1(\theta_{i1}(\beta,\delta),c_i,n_i)}{\partial c_i} \\
&\quad + \left[\frac{z_i-P_2(\theta_{i2}(\beta),c_i,n_i))}{1-P_2(\theta_{i2}(\beta),c_i,n_i)}\right]\frac{\partial \ln P_2(\theta_{i2}(\beta),c_i,n_i)}{\partial c_i} \\
&\quad + (1-z_i)\frac{\partial \ln P_2(\theta_{i2}(\beta),c_i,k_i)}{\partial c_i} + z_i\frac{\partial \ln P_2(\theta_{i2}(\beta),c_i,n_i)}{\partial c_i}
\end{aligned}$$

and:

$$\begin{aligned}
P_1(n_i) &= \Delta \ln \Gamma(n_i, c_i q_i) - \Delta \ln \Gamma(n_i, c_i) \\
P_2(n_i) &= \Delta \ln \Gamma(n_i, c_i p_i) - \Delta \ln \Gamma(n_i, c_i)
\end{aligned}$$

where $\Delta \ln \Gamma(y,a) = \ln \Gamma(y+a) - \ln \Gamma(a)$. $\frac{\partial \Delta \ln \Gamma(y,a)}{\partial a} = \psi(y+a) - \psi(a) = \Delta\psi(y,a)$, $\psi(x) = \frac{\partial \ln \Gamma(x)}{\partial x}$ denotes a di-gamma function.

$$\begin{aligned}
\frac{\partial \ln P_1(\theta_{i1}(\beta,\delta),c_i,n_i)}{\partial \theta_{i1}} &= \frac{\partial(\Delta \ln \Gamma(n_i,c_i q_i(\beta,\delta))-\Delta \ln \Gamma(n_i,c_i))}{\partial \theta_{i1}} = \Delta\psi(n_i, c_i\theta_{i1}(\beta,\delta))c_i \\
\frac{\partial \ln P_2(\theta_{i2}(\beta),c_i,n_i)}{\partial \theta_{i2}} &= \Delta\psi(n_i, c_i\theta_{i2}(\beta))c_i \\
\frac{\partial \ln P_2(\theta_{i2}(\beta),c_i,k_i)}{\partial \theta_{i2}} &= \Delta\psi(k_i, c_i\theta_{i2}(\beta))c_i \\
\frac{\partial \ln P_1(\theta_{i1}(\beta,\delta),c_i,n_i)}{\partial c_i} &= \Delta\psi(n_i, c_i\theta_{i1}(\beta,\delta))\theta_{i1}(\beta,\delta) - \Delta\psi(n_i, c_i) \\
\frac{\partial \ln P_2(\theta_{i2}(\beta),c_i,n_i)}{\partial c_i} &= \Delta\psi(n_i, c_i\theta_{i2}(\beta))\theta_{i2}(\beta) - \Delta\psi(n_i, c_i) \\
\frac{\partial \ln P_2(\theta_{i2}(\beta),c_i,k_i)}{\partial c_i} &= \Delta\psi(k_i, c_i\theta_{i2}(\beta))\theta_{i2}(\beta) - \Delta\psi(k_i, c_i)
\end{aligned}$$

Defining:

$$\begin{aligned}
u(\theta_{i1}(\beta,\delta), c_i, n_i) &= \frac{z_i-P_1(\theta_{i1}(\beta,\delta),c_i,n_i)}{1-P_1(\theta_{i1}(\beta,\delta),c_i,n_i)} \\
\frac{\partial}{\partial \ln P_1} u(\theta_{i1}(\beta,\delta), c_i, n_i) &= v(\theta_{i1}(\beta,\delta), c_i, n_i) = \frac{(z_i-1)P_1(\theta_{i1}(\beta,\delta),c_i,n_i)}{(1-P_1(\theta_{i1}(\beta,\delta),c_i,n_i))^2}
\end{aligned}$$

the derivation of the score can be based on:

$$\begin{aligned}
\frac{\partial \ln L_i^1}{\partial \delta_k} &= u(\theta_{i1}(\beta,\delta),c_i,n_i)\Delta\psi(n_i,c_iq_i)c_i\frac{\partial\theta_{i1}}{\partial\delta_k} = 0 \\
\frac{\partial \ln L_i^2}{\partial \beta_k} &= u(\theta_{i1}(\beta,\delta),c_i,n_i)\Delta\psi(n_i,c_iq_i)c_i\frac{\partial\theta_{i1}}{\partial\beta_k} = 0 \\
\frac{\partial \ln L_i^2}{\partial \beta_k} &= u(\theta_{i1}(\beta,\delta),c_i,n_i)\Delta\psi(n_i,c_iq_i)c_i\frac{\partial\theta_{i1}}{\partial\beta_k} \\
&\quad -u(\theta_{i1}(\beta,\delta),c_i,n_i)\Delta\psi(n_i,c_ip_i)c_i\frac{\partial\theta_{i2}}{\partial\beta_k} \\
&\quad +(1-z_i)\Delta\psi(k_i,c_ip_i)c_i\frac{\partial\theta_{i2}}{\partial\beta_k} + z_i\Delta\psi(n_i,c_ip_i)c_i\frac{\partial\theta_{i2}}{\partial\beta_k} \\
\frac{\partial \ln L_i}{\partial c_i} &= u(\theta_{i1}(\beta,\delta),c_i,n_i)\left(\Delta\psi(n_i,c_iq_i)\theta_{i1} - \Delta\psi(n_i,c_i)\right) \\
&\quad -u(\theta_{i2},c_i,n_i)\left(\Delta\psi(n_i,c_ip_i)\theta_{i2} - \Delta\psi(n_i,c_i)\right) \\
&\quad +\left(\Delta\psi(k_i,c_ip_i)\theta_{i2} - \Delta\psi(k_i,c_i)\right)
\end{aligned}$$

To implement the likelihood estimator, we provide the score, but rely on the numerical derivation of the Hessian. Then, Wald and LR tests are readily available from standard ML estimation routines.

C. Calculation of Marginal Effects in the Two-Part Models

For the following discussion, we denote $P_{i1}(n_i,x_i)$ by P_{i1}, $P_{i2}(n_i,x_i)$ by P_{i2} and $E[y_i|x_i,n_i]$ under the one-part model by F_i. According to (5), the expectation under the two-part model is given by:

$$\begin{aligned}
E[y_i|x_i,n_i] &= \frac{1-P_{i1}}{1-P_{i2}}(F_i - P_{i2}) + P_{i1} = \frac{(1-P_{i1})F_i - P_{i2} + P_{i1}P_{i2} + P_{i1} - P_{i1}P_{i2}}{1-P_{i2}} \\
&= \frac{(1-P_{i1})F_i - P_{i2} + P_{i1}}{1-P_{i2}} := \frac{u_i}{v_i}
\end{aligned}$$

The corresponding derivatives of u_i and v_i with respect to the k-th explanatory variable, x_{ik}, read as:

$$\begin{aligned}
\frac{\partial u_i}{\partial x_{ik}} &= -\frac{\partial P_{i1}}{\partial x_{ik}}F_i + (1-P_{i1})\frac{\partial F_i}{\partial x_{ik}} - \frac{\partial P_{i2}}{\partial x_{ik}} + \frac{\partial P_{i1}}{\partial x_{ik}} \\
&= \frac{\partial P_{i1}}{\partial x_{ik}}(1-F_i) + (1-P_{i1})\frac{\partial F_i}{\partial x_{ik}} - \frac{\partial P_{i2}}{\partial x_{ik}} \\
\frac{\partial v_i}{\partial x_{ik}} &= -\frac{\partial P_{i2}}{\partial x_{ik}}
\end{aligned}$$

Hence, one obtains:

$$\begin{aligned}
\frac{\partial E[y_i|x_i,n_i]}{\partial x_{ik}} &= \frac{1}{(1-P_{i2})}\frac{\partial u_i}{\partial x_{ik}} - \frac{u_i}{(1-P_{i2})}\frac{1}{(1-P_{i2})}\frac{\partial v_i}{\partial x_{ik}} \\
&= \frac{1}{(1-P_{i2})}\frac{\partial u_i}{\partial x_{ik}} - E[y_i|x_i,n_i]\frac{1}{(1-P_{i2})}\frac{\partial v_i}{\partial x_{ik}} \\
&= \frac{\frac{\partial P_{i1}}{\partial x_{ik}}(1-F_i) + (1-P_{i1})\frac{\partial F_i}{\partial x_{ik}} - \frac{\partial P_{i2}}{\partial x_{ik}}}{1-P_{i2}} + E[y_i|x_i,n_i]\frac{\frac{\partial P_{i2}}{\partial x_{ik}}}{(1-P_{i2})}
\end{aligned}$$

Under H_0, it holds that $\beta_k = \gamma_k$, $P_{i1} = P_{i2}$ and $E[y_i|x_i, n_i] = F_i$, and the marginal effects reduce to:

$$\begin{aligned}\frac{\partial E[y_i|x_i,n_i]}{\partial x_{ik}} &= \frac{\frac{\partial P_{i2}}{\partial x_{ik}}(1-F_i)+(1-P_{i2})\frac{\partial F_i}{\partial x_{ik}}-\frac{\partial P'_{i2}}{\partial x_{ik}}}{1-P_{i2}}+\frac{F_i\frac{\partial P_{i2}}{\partial x_{ik}}}{(1-P_{i2})}\\ &= \frac{\frac{\partial P_{i2}}{\partial x_{ik}}+(1-P_{i2})\frac{\partial F_i}{\partial x_{ik}}-\frac{\partial P_{i2}}{\partial x_{ik}}}{1-P_{i2}}\\ &= \frac{(1-P_{i2})\frac{\partial F_i}{\partial x_{ik}}}{1-P_{i2}}=\frac{\partial F_i}{\partial x_{ik}}\end{aligned}$$

Under the binomial two-part model, we insert:

$$\begin{aligned}P_{i1} &= G(x_i\gamma)^{n_i}\\ P_2 &= G(x_i\beta)^{n_i}\\ F_i &= G(x_i\beta)\\ \frac{\partial P_{i1}}{\partial x_{ik}} &= n_iG(x_i\gamma)^{n_i}(1-G(x_i\gamma))\gamma_k\\ \frac{\partial P_{i2}}{\partial x_{ik}} &= n_iG(x_i\beta)^{n_i}(1-G(x_i\beta))\beta_k\\ \frac{\partial F_i}{\partial x_{ik}} &= G(x_i\beta)(1-G(x_i\beta))\beta_k\end{aligned}$$

while for the beta-binomial two-part model, using $c_i = \frac{\left(e^{z_i\vartheta}-1\right)}{2}$, one obtains:

$$\begin{aligned}\ln P_{i1} &= \Delta\ln\Gamma(n_i, c_i\theta_{i1})-\Delta\ln\Gamma(n_i,c_i)\\ \ln P_{i2} &= \Delta\ln\Gamma(n_i, c_i\theta_{i2})-\Delta\ln\Gamma(n_i,c_i)\\ F_i &= \theta_{i2}=G(x_i\beta)\\ \frac{\partial\ln P_{i1}}{\partial x_{ik}} &= \Delta\psi(n_i,c_i\theta_{i1})c_i\frac{\partial\theta_{i1}}{\partial x_{ik}}+\left[\Delta\psi(n_i,c_i\theta_{i1})\theta_{i1}-\Delta\psi(n_i,c_i)\right]\frac{\partial c_i}{\partial x_{ik}}\\ \frac{\partial\ln P_{i2}}{\partial x_{ik}} &= \Delta\psi(n_i,c_i\theta_{i2})c_i\frac{\partial\theta_{i2}}{\partial x_{ik}}+\left[\Delta\psi(n_i,c_i\theta_{i1})\theta_{i2}-\Delta\psi(n_i,c_i)\right]\frac{\partial c_i}{\partial x_{ik}}\\ \frac{\partial F_i}{\partial x_{ik}} &= G(x_i\beta)(1-G(x_i\beta))\beta_k\\ \frac{\partial\theta_{i1}}{\partial x_{ik}} &= G(x_i\gamma)(1-G(x_i\gamma))\gamma_k\\ \frac{\partial c_i}{\partial x_{ik}} &= \tfrac{1}{2}e^{z_i\vartheta}\vartheta_k\end{aligned}$$

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Papke, L.E.; Wooldridge, J.M. Econometric Methods for Fractional Response Variables with an Application to 401(k) Plan Participation Rates. *J. Appl. Econom.* **1996**, *11*, 619–632.

2. Papke, L.E.; Wooldridge, J.M. Panel Data Methods for Fractional Response Variables with an Application to Test Pass Rates. *J. Econom.* **2008**, *145*, 121–133.
3. Murteira, J.M.R.; Ramalho, E.A.; Ramalho, J.J.S. Alternative Estimating and Testing Empirical Strategies for Fractional Regression Models. *J. Econ. Surv.* **2011**, *25*, 19–68.
4. Paolino, P. Maximum Likelihood Estimation of Models with Beta-Distributed Dependent Variables. *Polit. Anal.* **2001**, *9*, 325–346.
5. Kieschnick, R.; McCullough, B.D. Regression Analysis of Variates Observed on (0, 1): Percentages, Proportions and Fractions. *Stat. Model.* **2003**, *3*, 193–213.
6. Cribari-Neto, F.; Ferrari, S.L.P. Beta Regression for Modelling Rates and Proportions. *J. Appl. Stat.* **2004**, *31*, 799–815.
7. Mullahy, J. Specification and Testing of Some Modified Count Data Models. *J. Econom.* **1986**, *33*, 341–365.
8. Lambert, D. Zero-inflated Poisson Regression, with an Application to Defects in Manufacturing. *Technometrics* **1992**, *34*, 1–14.
9. Cameron, C.A.; Trivedi, P.K. *Microeconometrics: Methods and Applications*; Cambridge University Press: Cambridge, UK, 2005.
10. Wooldridge, J.M. *Econometric Analysis of Cross Section and Panel Data*; MIT: Cambridge, MA, USA, 2002.
11. Da Silva, J.V.; Ramalho, J.J.S. A Two-Part Fractional Regression Model for the Financial Leverage Decisions of Micro, Small, Medium and Large Firms. *Quant. Financ.* **2009**, *9*, 621–636.
12. Da Silva, J.V.; Ramalho, J.J.S. Functional Form Issues in the Regression Analysis of Corporate Capital Structure. *Empir. Econ.* **2013**, *44*, 799–831.
13. Oberhofer, H.; Pfaffermayr, M. Fractional Response Models—A Replication Exercise of Papke and Wooldridge (1996). *Contemp. Econ.* **2012**, *6*, 56–64.
14. Cook, D.O.; Kieschnick, R.; McCullough, B.D. Regression Analysis of Proportions in Finance with Self Selection. *J. Empir. Financ.* **2008**, *15*, 860–867.
15. Ferrari, S.L.P.; Ospina, R. A General Class of Zero-or-One Inflated Beta Regression Models. *Comput. Stat. Data Anal.* **2012**, *56*, 1609–1623.
16. Davidson, R.; MacKinnon, J.G. Several Tests for Model Specification in the Presence of Alternative Hypotheses. *Econometrica* **1981**, *49*, 781–793.
17. Papke, L.E.; Wooldridge, J.M. *Econometric Methods for Fractional Response Variables with an Application to 401(k) Plan Participation Rates*; National Bureau of Economic Research Technical Working Paper No. 147; National Bureau of Economic Research: Cambridge, MA, USA, 1993.
18. Lin, T.F.; Schmidt, P. A Test of the Tobit Specification Against an Alternative Suggested by Cragg. *Rev. Econ. Stat.* **1984**, *66*, 174–177.
19. Murteira, J.M.R.; Ramalho, E.A.; Ramalho, J.J.S. A Generalized Goodness-of-Functional Form Test for Binary and Fractional Regression Models. *Manch. Sch.* **2013**, *82*, 488–507.
20. Cragg, J.G. Some Statistical Models for Limited Dependent Variables with Application to the Demand for Durable Goods. *Econometrica* **1971**, *39*, 829–844.
21. Gourieroux, C.; Monfort, A.; Trognon A. Pseudo-maximum Likelihood Methods: Theory. *Econometrica* **1984**, *52*, 681–700.

22. Heckman, J.J.; Willis, R.J. A Beta-Logistic Model for the Analysis of Sequential Labor Force Participation by Married Women. *J. Polit. Econ.* **1977**, *85*, 27–58.
23. McCulloch, C.E.; Searle, A.F.M. *Generalized, Linear and Mixed Models*; John Wiley & Sons: Hoboken, NJ, USA, 2001.
24. Prentice, R.L. Binary Regression Using an Extended Beta-Binomial Distribution, with Discussion of Correlation Induced by Covariate Measurement Errors. *J. Am. Stat. Assoc.* **1986**, *81*, 321–327.
25. Murteira, J.M.R.; Santo Silva, J.M.C. Estimation of Default Probabilities Using Incomplete Contracts Data. *J. Empir. Financ.* **2009**, *16*, 457–465.
26. Johnson, N.L.; Kemp, A.W.; Kotz, S. *Univariate Discrete Distributions*, 3rd ed.; John Wiley & Sons: Hoboken, NJ, USA, 2005.
27. Mullahy, J. *Multivariate Fractional Regression Estimation of Econometric Share Models*; NBER Working Papers 16354; National Bureau of Economic Research: Cambridge, MA, USA, 2010.
28. Murteira, J.M.R.; Ramalho, J.J.S. Regression Analysis of Multivariate Fractional Data. *Econom. Rev.* **2014**, forthcoming.
29. Oberhofer, H.; Pfaffermayr, M. *Two-Part Models for Fractional Responses Defined as Ratios of Integers*; WIFO Working Papers 472; Austrian Institute of Economic Research: Vienna, Austria, 2014.
30. Journal of Applied Econometrics Data Archive. Available online: http://qed.econ.queensu.ca/jae/1996-v11.6/papke-wooldridge/ (accessed on 27 August 2014).
31. Nagelkerke, N.J.D. A Note on a General Definition of the Coefficient of Determination. *Biometrika* **1991**, *78*, 691–692.