Consistent estimation for the full-fledged fixed effects zero-inflated Poisson model*

Yoshitsugu Kitazawa†

April 5, 2014

Abstract

This paper advocates the transformations used for the consistent estimation of the full-fledged fixed effects zero-inflated Poisson model whose zero outcomes can arise from both of logit and Poisson parts and which equips both parts with the fixed effects. The valid moment conditions are constructed on the basis of the transformations. The finite sample behaviors of GMM and EL estimators employing the moment conditions are investigated by use of Monte Carlo experiments.

Keywords: fixed effects zero-inflated Poisson model; predetermined explanatory variables in Poisson part; moment conditions; GMM; EL; Monte Carlo experiments

JEL classification: C23, C25, C51

1 Introduction

The zero-inflated Poisson model (hereafter ZIP model) proposed by Lambert (1992) is one of the models dealing with count data with zero values being superabundant. Empirical studies using the ZIP model are often found in the literature on the econometric analysis: Gurmu and Trivedi (1996) on the relationship between the recreational boating trips and boat owners' attributes, Crépon & Duguet (1997) and Hu & Jefferson (2009) on the patents and R&D relationship, Tomlin (2000) on the FDI and foreign exchange relationship, List (2001) on the scheduled interviews and characteristics relationship with respect to job-seekers in the academic market, Campolieti (2002) on the relationship between recurrence of workers' injuries and employer accommodations, Durham et al. (2004) on the selection and characteristics relationship with respect to wine in a restaurant, Edmeades & Smale (2006) on the relationship on count of banana plants and traits with respect to household farms, and Frondel & Vance (2011) investigating the determinants of public transit ridership in Germany, etc.

However, to the best of the author's knowledge, two types of studies deal with the incipient ZIP model allowing for the fixed effects in the context of count panel data.[1] Majo (2010) and

*Discussion Paper Series, Faculty of Economics, Kyushu Sangyo University Discussion Paper, April 2014, No. 66.

†Correspondence: Faculty of Economics, Kyushu Sangyo University, Matsukadai 2-3-1, Higashi-ku, Fukuoka, Japan. E-mail: kitazawa@ip.kyusan-u.ac.jp

[1] A small number of studies deal with various kinds of the random effects ZIP model (e.g. Crépon & Duguet, 1997; Hall, 2000; Min & Agresti, 2005; Lam et al., 2006; Hasan & Sneddon, 2009; Feng & Zhu, 2011). However, the random effects model generally lacks flexibility, compared to the fixed effects model. One of the assumptions often claimed in the panel data analysis is that the individual heterogeneity can be arbitrarily correlated with the explanatory variables.

Majo & Van Soest (2011) consider the fixed effects ZIP model including the fixed effects in both of logit and Poisson parts, propose the estimation method for the model, and apply it to the micro level data concerning health care service utilization in Europe. Their model assumes the truncated-at-zero Poisson model in Poisson part, implying that the origin of the zero count outcomes is confined to the logit part. On the contrary, the quasi-conditional maximum likelihood estimator presented by Gilles (2012) and Gilles & Kim (2013) can consistently estimate the fixed effects ZIP model where the zero count outcomes originate from Poisson part as well as from the logit part. However, the fixed effects ZIP model which they assume incorporates no fixed effect in the logit part. Accordingly, it might be said that the fixed effects ZIP models assumed in both types of studies are still less plenary.

Different from the studies by Majo (2010) and Majo & Van Soest (2011) and by Gilles (2012) and Gilles & Kim (2013), the fixed effects ZIP model discussed in this paper has the Poisson part from which the zero count outcome is not improbable and the logit part with the fixed effects being built-in. The valid moment conditions for this ZIP model are constructed based on two transformations for different specifications of the explanatory variables in Poisson part and then the parameters of interest are consistently estimated by use of the GMM (Generalized Method of Moments) proposed by Hansen (1982) and EL (Empirical Likelihood) method proposed by Owen (1988, 1990, 1991, 2001) and advanced by Qin & Lawless (1994) using these moment conditions. Monte Carlo experiments (which are limited) show that the large cross-sectional size would be needed for enhancing the accuracy and precision of the estimators.

The rest of the paper is organized as follows. Section 2 provides the fixed effects ZIP model with the fixed effects being included in both of logit and Poisson parts and constructs the moment conditions for consistently estimating the parameters of interest in both parts. Section 3 outlines the GMM and EL estimators using the moment conditions. Section 4 lays out some Monte Carlo results for the estimators. Section 5 concludes the discussion.

2 Model and moment conditions

In this section, the fixed effects ZIP model is considered, which has two potential sources of outbreaks of zero count variables: logit probability and Poisson density and which furnishes both of logit and Poisson parts with the fixed effects. The model in this paper is distinct from that assumed in Majo (2010) and Majo & Van Soest (2011) and that assumed in Gilles (2012) and Gilles & Kim (2013), as is described in previous section. The fixed effects ZIP model is described in the implicit form and the mean and variance of its disturbance are specified. Then, presupposing that the disturbance and its square are uncorrelated with any transformations of the disturbances in past and the fixed effects, the moment conditions for consistently estimating the parameters of interest are constructed under both of the slightly strong assumptions and the mitigated ones. Under the slightly strong assumptions, the explanatory variables in both of the logit probability and the Poisson mean are slightly exogenous, while under the mitigated assumptions, the explanatory variables in the logit probability are slightly exogenous and those in the Poisson mean are predetermined. The overtone of the slight exogeneity introduced in this paper is that the count dependent variable at a given period wield no influence over the explanatory variable at the period just behind the occurrence of its count variable, whereas it could make some sorts of influences on the subsequent explanatory variables.

2.1 Fixed effects ZIP model

The fixed effects ZIP model has the following two potential sources of outbreaks of zero count dependent variables: $y_{it} = 0$ with probability $1 - p_{it}$, while $y_{it} \sim \text{Pois}(q_{it})$ with probability

p_{it}, where subscripts i and t denote the individual and the time period with $i = 1, \ldots, N$ and $t = 1, \ldots, T$, respectively. It is assumed that $N \to \infty$, whereas T is fixed.

In this paper, the logit probability of generating the binary process is specified as $p_{it} = \exp(\psi_i + \delta w_{it})/(1 + \exp(\psi_i + \delta w_{it}))$, while the mean of generating the Poisson process is specified as $q_{it} = \exp(\eta_i + \beta x_{it})$, where ψ_i and η_i are the fixed effects, w_{it} and x_{it} are the (continuous) explanatory variables, and δ and β are the parameters of interest. It is assumed that the variables in the model are independent and identically distributed among individuals.

The fixed effects ZIP model can be written in the following implicit form:

$$y_{it} = p_{it} q_{it} + v_{it}, \tag{1}$$

where the disturbance v_{it} is tailored to the specifications based on the slightly strong assumptions and the mitigated ones as is described in the following subsections.[2]

2.2 Slightly strong assumptions and moment conditions

In this case, the assumptions on the disturbances are as follows:

$$\begin{aligned} \mathrm{E}[v_{it} \mid \psi_i, w_i^{t+1}, \eta_i, x_i^{t+1}, v_i^{t-1}] &= 0, & (2) \\ \mathrm{E}[v_{it}^2 \mid \psi_i, w_i^{t+1}, \eta_i, x_i^{t+1}, v_i^{t-1}] &= p_{it} q_{it}(1 + (1 - p_{it}) q_{it}), & (3) \end{aligned}$$

where $w_i^{t+1} = (w_{i1}, \ldots, w_{i,t+1})$, $x_i^{t+1} = (x_{i1}, \ldots, x_{i,t+1})$, and $v_i^{t-1} = (v_{i0}, \ldots, v_{i,t-1})$ with v_{i0} being empty.

Under the assumptions (2) and (3) for the implicit form (1), the following moment conditions are constructed for consistently estimating δ and β when $N \to \infty$ and T is fixed:

$$\mathrm{E}[\Phi_{it}(\delta, \beta) \mid \psi_i, w_i^t, \eta_i, x_i^t, v_i^{t-2}] = 0, \quad \text{for } t = 2, \ldots, T, \tag{4}$$

with

$$\begin{aligned} \Phi_{it}(\delta, \beta) = {} & (\tanh(\delta\, \Delta w_{it}/2) - 1) \exp(-\beta\, \Delta x_{it})(y_{it}^2 - y_{it}) \\ & + (\tanh(\delta\, \Delta w_{it}/2) + 1) \exp(\beta\, \Delta x_{it})(y_{i,t-1}^2 - y_{i,t-1}) - 2 \tanh(\delta\, \Delta w_{it}/2) y_{it} y_{i,t-1}, \end{aligned} \tag{5}$$

where Δ is the first-differencing operator such as $\Delta w_{it} = w_{it} - w_{i,t-1}$ and $\Delta x_{it} = x_{it} - x_{i,t-1}$. The derivation of (4) with (5) is shown in Appendix A.

The transformation (5) is referred to as the "PHI transformation" in this paper. As is seen from (4), the PHI transformation can construct the unconditional moment conditions for consistently estimating the parameters of interest (i.e. δ and β) by using the functions of the information set $(\psi_i, w_i^t, \eta_i, x_i^t, v_i^{t-2})$, if the explanatory variables in both the logit probability and the Poisson mean are slightly exogenous.

2.3 Mitigated assumptions and moment conditions

In this case, the assumptions on the disturbances are as follows:

$$\mathrm{E}[v_{it} \mid \psi_i, w_i^{t+1}, \eta_i, x_i^t, v_i^{t-1}] = 0, \tag{6}$$

[2] As is seen from (1), the mean of the dependent variable y_{it} is the product of the logit probability and the exponential function. Accordingly, the conventional estimators for the ordinary fixed effects count data model, such as the within group mean scaling estimator proposed by Blundell et al. (2002) (which is equivalent to the conditional maximum likelihood estimator proposed by Hausman et al., 1984) and the quasi-differenced GMM estimators proposed by Chamberlain (1992) and Wooldridge (1997), are not applicable to the fixed effects ZIP model proposed in this paper.

$$\mathrm{E}[v_{it}^2 \mid \psi_i, w_i^{t+1}, \eta_i, x_i^t, v_i^{t-1}] \;=\; p_{it}q_{it}(1+(1-p_{it})q_{it}). \tag{7}$$

Under the assumptions (6) and (7) for the implicit form (1), the following moment conditions are constructed for consistently estimating δ and β when $N \to \infty$ and T is fixed:

$$\mathrm{E}[\Psi_{it}(\delta,\beta) \mid \psi_i, w_i^t, \eta_i, x_i^{t-1}, v_i^{t-2}] = 0, \quad \text{for } t = 2, \ldots, T, \tag{8}$$

with

$$\begin{aligned}\Psi_{it}(\delta,\beta) &= (\tanh(\delta\ \Delta w_{it}/2) - 1)\exp(-2\beta\ \Delta x_{it})(y_{it}^2 - y_{it}) \\ &\quad +(\tanh(\delta\ \Delta w_{it}/2)+1)(y_{i,t-1}^2 - y_{i,t-1}) - 2\tanh(\delta\ \Delta w_{it}/2)\exp(-\beta\ \Delta x_{it})y_{it}y_{i,t-1}.\end{aligned} \tag{9}$$

The derivation of (8) with (9) is shown in Appendix B.

The transformation (9) is referred to as the "PSI transformation" in this paper. As is seen from (8), the PSI transformation can construct the unconditional moment conditions for consistently estimating the parameters of interest (i.e. δ and β) by using the functions of the information set $(\psi_i, w_i^t, \eta_i, x_i^{t-1}, v_i^{t-2})$, if the explanatory variables in the logit probability are slightly exogenous and those in the Poisson mean are predetermined.

3 Estimation methods

This section reviews the two estimators using the unconditional moment conditions based on the PHI or PSI transformations in previous section. The GMM estimator is obtained by minimizing the quadratic form composed of the sample version vector of moment conditions and a weighting matrix. The EL estimator, as an alternative to the GMM estimator, is obtained by maximizing the log likelihood constructed by using the implied probability under the constraint of the sample version vector weighted by the implied probability. Many studies revealing that the EL estimator behaves better than the GMM estimator in small sample are reported using the theoretical analysis and Monte Carlo experiment (e.g. Newey & Smith, 2004; Anatolyev, 2005; Ramalho, 2005).

3.1 GMM estimator

Any set of the unconditional moment conditions constructed on the basis of (4) with (5) and (8) with (9) can be collectively written in the following $m \times 1$ vector form:

$$\mathrm{E}[g_i(\theta)] = 0, \tag{10}$$

where m is number of the moment conditions, $\theta = [\delta\ \beta]'$, $g_i(\theta)$ (which is the function of θ) is composed of the observables variables for the individual i and the parameter vector θ. Using the following sample moments without $g_i(\theta)$ being weighted:

$$\bar{g}(\theta) = (1/N)\sum_{i=1}^{N} g_i(\theta) = 0, \tag{11}$$

which is the ersatz of (10), and the $m \times m$ inverse of asymptotically optimal weighting matrix:

$$\bar{\Omega}(\hat{\theta}_1) = (1/N)\sum_{i=1}^{N} g_i(\hat{\theta}_1)g_i(\hat{\theta}_1)', \tag{12}$$

where $\hat{\theta}_1$ is any initial consistent estimator for θ, the GMM estimator is constructed as follows:

$$\hat{\theta}_{\mathrm{GMM}} = \arg\min_{\theta} \bar{g}(\theta)'(\bar{\Omega}(\hat{\theta}_1))^{-1}\bar{g}(\theta)\,. \tag{13}$$

Taking notice of the assumption that the variables are independent and identically distributed among individuals, it follows that

$$N^{1/2}(\hat{\theta}_{\text{GMM}} - \theta_0) \xrightarrow{\text{d}} \text{N}(0, (D(\theta_0)'(\Omega(\theta_0))^{-1}D(\theta_0))^{-1}), \tag{14}$$

where $D(\theta_0) = (\partial \text{E}[g_i(\theta)]/\partial\theta') \mid_{\theta=\theta_0}$ and $\Omega(\theta_0) = \text{E}[g_i(\theta_0)g_i(\theta_0)']$.

3.2 EL estimator

Hsueh & Lee (2009, 2012) are the pioneering works of applying the EL estimation to count panel data model. According to their papers, the probability π_i is defined for individual i, which is the probability of realization of the variables composing $g_i(\theta)$ and satisfies the following relationship by definition:

$$\sum_{i=1}^{N} \pi_i = 1. \tag{15}$$

In addition, the "surprisal" for individual i is $-\ln(1/N)$ when the following constraint is not imposed:

$$\sum_{i=1}^{N} \pi_i g_i(\theta) = 0, \tag{16}$$

which is interpreted as being the empirical counterpart of (10) weighted with the probability π_i, while it is defined as $-\ln\pi_i$ subject to (16) when the constraint (16) is imposed.[3] The EL estimator for θ (i.e. $\hat{\theta}_{\text{EL}}$) is obtained by maximizing the arithmetic mean of the differences of the former from the latter with respect to θ and $\pi_1, \ldots, \pi_N$. That is, the problem to be solved is as follows:

$$\min_{\theta,\pi_1,\ldots,\pi_N} -(1/N)\sum_{i=1}^{N}((-\ln(1/N)) - (-\ln\pi_i)), \tag{17}$$

subject to (15) and (16).

The usage of some algebras in this minimization problem leads to the solution of the following dual problem:

$$\hat{\theta}_{\text{EL}} = \arg\min_{\theta}(\max_{\lambda}\ (1/N)\sum_{i=1}^{N}\ln(1 - \lambda' g_i(\theta))), \tag{18}$$

where λ is the $m \times 1$ vector of Lagrange multipliers used in the minimization problem for (17) subject to (15) and (16). It should be noted that in transforming the problem from (17) subject to (15) and (16) to (18), number of the parameters to be estimated decreases from $2 + N$ to $2 + m$, as long as $N > m$. Qin & Lawless (1994) show that the EL estimator $\hat{\theta}_{\text{EL}}$ has the same limit distribution as the GMM estimator $\hat{\theta}_{\text{GMM}}$, which is represented by (14).

4 Monte Carlo

In this section, the finite sample behaviors of the GMM and EL estimators based on the PHI and PSI transformations are investigated with some Monte Carlo experiments. The experiments are carried out by using the programming language "R" (version 3.0.2) developed by R Core Team (2013).

[3] The terminology "surprisal" is coined by Tribus (1961).

4.1 Data generating process

The data generating process (DGP) is as follows:

$$
\begin{aligned}
y_{it} &= y_{it}^p y_{it}^q, \\
y_{it}^p &\sim \mathrm{Bin}(1, p_{it}), \\
p_{it} &= \exp(\psi_i + \delta w_{it})/(1 + \exp(\psi_i + \delta w_{it})), \\
w_{it} &= \alpha w_{i,t-1} + \iota \psi_i + \zeta_{it}, \\
w_{i1} &= (1/(1-\alpha))\iota\psi_i + (1/(1-\alpha^2)^{(1/2)})\zeta_{i1}, \\
& \quad \psi_i \sim \mathrm{N}(0, \sigma_\psi^2);\ \zeta_{it} \sim \mathrm{N}(0, \sigma_\zeta^2), \\
y_{it}^q &\sim \mathrm{Pois}(q_{it}), \\
q_{it} &= \exp(\eta_i + \beta x_{it}), \\
x_{it} &= \rho x_{i,t-1} + \tau \eta_i + \varepsilon_{it}, \\
x_{i1} &= (1/(1-\rho))\tau\eta_i + (1/(1-\rho^2)^{(1/2)})\varepsilon_{i1}, \\
& \quad \eta_i \sim \mathrm{N}(0, \sigma_\eta^2);\ \varepsilon_{it} \sim \mathrm{N}(0, \sigma_\varepsilon^2).
\end{aligned}
$$

In the DGP, values are set to the parameters δ, α, ι, σ_ψ^2, σ_ζ^2, β, ρ, τ, σ_η^2 and σ_ε^2. The cross-sectional size $N = 1000$, 5000 and 10000 and the number of time periods $T = 4$ and 8 are used in the experiments. The number of replications is 10000.

4.2 Estimators assayed

In the experiments, the GMM and EL estimators utilize the unconditional moment conditions based on two types of the conditional moment conditions: (4) with (5) (based on the PHI transformation) and (8) with (9) (based on the PSI transformation).

The unconditional moment conditions constructed based on the PHI transformation are as follows:

$$
\begin{aligned}
\mathrm{E}[\Phi_{it}(\delta, \beta)\ \Delta w_{it}] &= 0, \quad \text{for } t = 2, \ldots, T, & (19) \\
\mathrm{E}[\Phi_{it}(\delta, \beta)\ \Delta x_{it}] &= 0, \quad \text{for } t = 2, \ldots, T, & (20)
\end{aligned}
$$

while those based on the PSI transformation are as follows:

$$
\begin{aligned}
\mathrm{E}[\Psi_{it}(\delta, \beta)\ \Delta w_{it}] &= 0, \quad \text{for } t = 2, \ldots, T, & (21) \\
\mathrm{E}[\Psi_{it}(\delta, \beta)\ x_{is}] &= 0, \quad \text{for } s = 1, \ldots, t-1;\ t = 2, \ldots, T, & (22)
\end{aligned}
$$

where the moment conditions (22) are of the forms of the sequential moment conditions, in which the lagged levels of the explanatory variables x_{it} are the instruments for the PSI transformations, and the number of the sequential moment consitions grows as the number of time periods T increases.[4]

In this paper, the GMM and EL estimators using the moment conditions (19) and (20) are referred to as the “GMM(PHI)” and “EL(PHI)” estimators respectively, while those using the moment conditions (21) and (22) are referred to as the “GMM(PSI)” and “EL(PSI)” estimators respectively. As a control, the pooled maximum likelihood estimator (hereafter, the “ML(POOL)” estimator) is used, which ignores the individual heterogeneity and accordingly has the indigenous bias. The GMM and EL estimations are implemented by using the “R”

[4]The sequential moment conditions are proposed by Holtz-Eakin et al. (1988) and Arellano & Bond (1991) for the ordinary dynamic panel data model and Chamberlain (1992) and Wooldridge (1997) for the count panel data model.

package "gmm" developed by Chaussé (2010), while the ML estimation is implemented by using the "pscl" developed by Zeileis et al. (2008).[5]

4.3 Results

Monte Carlo results for the estimators assayed when $T = 4$ and 8 are shown in Table 1 and 2, respectively. It can be seen that the bias and rmse (root mean square error) for the GMM and EL estimators dwindle in size as the cross-sectional size N increases, reflecting the consistency, while the considerable upward bias of the inconsistent ML(POOL) estimator remains unchanged. Figure 1 and 2 are the boxplots of the GMM and EL estimators for δ and β when $T = 4$, respectively, while Figure 3 and 4 are those when $T = 8$. It can be seen that the interquantile range (hereafter IQR) and whisker length become narrower and less standoff outliers are found as the cross-sectional size N is larger.

When using the PSI transformations based on the mitigated assumptions, the EL estimator overwhelmingly outperforms the GMM estimator whose small sample performance is poor in the extreme, as is seen from the comparison of the performance of the EL(PSI) estimator with that of the GMM(PSI) estimator. The smaller sizes of bias and rmse, narrower IQR and whisker range, and less standoff outliers are recognizable for the EL estimator.[6] One potential explanation for them is that the GMM(PSI) estimator might suffer from the weak instruments problem pointed out by Bound et al. (1995) and Staiger & Stock (1997). That is, it could be that the lagged levels of the explanatory variables x_{it} in the moment conditions (22) are the weak instruments for the PSI transformations (9).[7] It might be reckoned that the weak instruments problem would be mollified by using the EL estimator instead of the GMM estimator in this case. Another one is that the GMM(PSI) estimator (which is the two-step estimator) might be afflicted with the higher-order bias characteristic of the GMM estimator shown by Newey & Smith (2004), leading to its poor small sample performance, judging from the fact that it uses many growing instruments for the PSI transformations as the number of time periods T increases. In addition, Newey & Smith (2004) theoretically show that the higher-order bias of the EL estimator is considerably smaller than that of the GMM estimator, which suggests that the former small sample property is superior to the latter one, especially in using the increasing number of the moment conditions, while the Monte Carlo evidences by Ramalho (2005) bear out it. In fact, it is sure that the discrepancy of the small sample performances is large between the GMM(PSI) estimator and the EL(PSI) estimator for the case with $T = 8$, compared to the case with $T = 4$. The number of the moment conditions based on the PSI transformations is 9 when $T = 4$, while it is 35 when $T = 8$.

The observations for which $(y_{it}, y_{i,t-1}) = (0, 0)$, $(0, 1)$, or $(1, 0)$ make no contribution to the identification using the GMM and EL estimators, as is seen from the PHI and PSI transformations (i.e. (5) and (9)). In the DGP, rate of the above combinations of the dependent variables attains to about 70 percent for each replication, which is discarded in the estimations. It can be said that these are the events peculiar to the ZIP model, behind which the mass generation of zero values of count dependent variables lies. Accordingly, a considerable degree of sample sizes would be needed for enhancing the accuracy and precision of the GMM and EL estimators, which is reflected in the Monte Carlo results shown in Table 1 and 2 and Figure 1 to 4.

[5] The functions "gmm()" and "el()" are used for the GMM and EL estimations in the "gmm" package (version 1.4–5), while the function "zeroinfl()" is used for the ML estimation in the "pscl" package (version 1.04.4).

[6] In the experiment using a different seed, the larger rmse of δ is found when $N = 10000$ than when $N = 5000$ for the GMM(PSI) estimator, although the boxplot exhibits the tendency of the incresing accuracy and precision for the larger N. This is because a far standoff outliers is estimated in a replication.

[7] The weak instruments problem often materializes in analyses using the dynamic panel data model and count panel data model (e.g. Blundell & Bond, 1998; Blundell et al., 2002).

5 Conclusion

In this paper, the two types of moment conditions were proposed for consistently estimating the parameters of interest in the fixed effects ZIP model in which zero count outcomes could germinate from the Poisson part as well as from the logit part: the moment conditions for the case of slightly exogenous explanatory variables in logit and Poisson parts and those for the case of slightly exogenous explanatory variables in logit part and predetermined ones in Poisson part. Monte Carlo experiments indicated that the large number of individuals would behooves for obtaining the accurate and precise GMM and EL estimates. It is conceivable that this would be caused by the virtual decrease of sample sizes contributing to the estimation, which is due to mass generation of zero count outcomes.

Appendix A

First, the following relationship is obtained from (1), (2) and (3):

$$\mathrm{E}[(y_{it}^2 - y_{it}) \mid \psi_i, w_i^{t+1}, \eta_i, x_i^{t+1}, v_i^{t-1}] = p_{it} q_{it}^2. \tag{A.1}$$

Multiplying (A.1) dated t by $q_{i,t-1}/q_{it}$ gives

$$\mathrm{E}[\exp(-\beta\ \Delta x_{it})(y_{it}^2 - y_{it}) \mid \psi_i, w_i^{t+1}, \eta_i, x_i^{t+1}, v_i^{t-1}] = p_{it} q_{it} q_{i,t-1}, \tag{A.2}$$

while multiplying (A.1) dated $t-1$ by $q_{it}/q_{i,t-1}$ gives

$$\mathrm{E}[\exp(\beta\ \Delta x_{it})(y_{i,t-1}^2 - y_{i,t-1}) \mid \psi_i, w_i^{t}, \eta_i, x_i^{t}, v_i^{t-2}] = p_{i,t-1} q_{it} q_{i,t-1}. \tag{A.3}$$

Next, the following relationship is obtained from (1) and (2):

$$\mathrm{E}[y_{it} y_{i,t-1} \mid \psi_i, w_i^{t+1}, \eta_i, x_i^{t+1}, v_i^{t-1}] = p_{it} p_{i,t-1} q_{it} q_{i,t-1} + p_{it} q_{it} v_{i,t-1}. \tag{A.4}$$

Third, the following relationship holds for the logit probability:

$$\tanh(\delta\ \Delta w_{it}/2) = (p_{it} - p_{i,t-1})/(p_{it} + p_{i,t-1} - 2p_{it} p_{i,t-1}), \tag{A.5}$$

using which Kitazawa (2012) constructs the first-order condition for the conditional maximum likelihood estimator for the static fixed effects logit model developed by Rasch (1960, 1961) and Chamberlain (1980). Multiplying the numerator and denominator of the right-hand side of (A.5) by $q_{it} q_{i,t-1}$ gives

$$\tanh(\delta\ \Delta w_{it}/2)\ (p_{it} q_{it} q_{i,t-1} + p_{i,t-1} q_{it} q_{i,t-1} - 2 p_{it} p_{i,t-1} q_{it} q_{i,t-1}) = (p_{it} q_{it} q_{i,t-1} - p_{i,t-1} q_{it} q_{i,t-1}). \tag{A.6}$$

Finally, plugging (A.2), (A.3) and (A.4) into (A.6) and then taking the expectation conditional on the information set $(\psi_i, w_i^t, \eta_i, x_i^t, v_i^{t-2})$ for its both sides, the conditional moment conditions (4) with (5) are obtained.

Appendix B

As is the case with Appendix A, the following four relationships are recognizable:

$$\mathrm{E}[(y_{i,t-1}^2 - y_{i,t-1}) \mid \psi_i, w_i^{t}, \eta_i, x_i^{t-1}, v_i^{t-2}] = p_{i,t-1} q_{i,t-1}^2, \tag{B.1}$$

$$\mathrm{E}[\exp(-2\beta\ \Delta x_{it})(y_{it}^2 - y_{it}) \mid \psi_i, w_i^{t+1}, \eta_i, x_i^{t}, v_i^{t-1}] = p_{it} q_{i,t-1}^2, \tag{B.2}$$

$$\mathrm{E}[\exp(-\beta\ \Delta x_{it}) y_{it} y_{i,t-1} \mid \psi_i, w_i^{t+1}, \eta_i, x_i^{t}, v_i^{t-1}] = p_{it} p_{i,t-1} q_{i,t-1}^2 + p_{it} q_{i,t-1} v_{i,t-1}, \tag{B.3}$$

$$\tanh(\delta\ \Delta w_{it}/2)\ (p_{it} q_{i,t-1}^2 + p_{i,t-1} q_{i,t-1}^2 - 2 p_{it} p_{i,t-1} q_{i,t-1}^2) = (p_{it} q_{i,t-1}^2 - p_{i,t-1} q_{i,t-1}^2). \tag{B.4}$$

Plugging (B.1), (B.2) and (B.3) into (B.4) and then taking the expectation conditional on the information set $(\psi_i, w_i^t, \eta_i, x_i^{t-1}, v_i^{t-2})$ for its both sides, the conditional moment conditions (8) with (9) are obtained.

References

Anatolyev, S. (2005), 'GMM, GEL, serial correlation, and asymptotic bias', *Econometrica*, **73**, 983–1002.
doi:10.1111/j.1468-0262.2005.00601.x

Arellano, M. & Bond, S. (1991), 'Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations', *Review of Economic Studies*, **58**, 277–297.
doi:10.2307/2297968

Blundell, R. & Bond, S. (1998), 'Initial Conditions and Moment Restrictions in Dynamic Panel Data Models', *Journal of Econometrics*, **87**, 115–143.
doi:10.1016/S0304-4076(98)00009-8

Blundell, R., Griffith, R. & Windmeijer, F. (2002), 'Individual effects and dynamics in count data models', *Journal of Econometrics*, **108**, 113–131.
doi:10.1016/S0304-4076(01)00108-7

Bound, J., Jaeger, D.A. & Baker, R.M. (1995), 'Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak', *Journal of the American Statistical Association*, **90**, 443–450.
doi:10.1080/01621459.1995.10476536

Campolieti, M. (2002), 'The recurrence of occupational injuries: estimates from a zero inflated count model', *Applied Economics Letters*, **9**, 595–600.
doi:10.1080/13504850110111207

Chamberlain, G. (1980). 'Analysis of covariance with qualitative data', *Review of Economic Studies*, **47**, 225–238.
doi:10.2307/2297110

Chamberlain, G. (1992), 'Comment: sequential moment restrictions in panel data', *Journal of Business and Economic Statistics*, **10**, 20–26.
doi:10.1080/07350015.1992.10509881

Chaussé, P. (2010), 'Computing generalized method of moments and generalized empirical likelihood with R', *Journal of Statistical Software*, **34**, 1–35.
http://www.jstatsoft.org/v34/i11/

Crépon, B. & Duguet, E. (1997), 'Research and development, competition and innovation pseudo-maximum likelihood and simulated maximum likelihood methods applied to count data models with heterogeneity', *Journal of Econometrics*, **79**, 355–378.
doi:10.1016/S0304-4076(97)00027-4

Durham, C.A., Pardoe, I. & Vega-H, E. (2004), 'A methodology for evaluating how product characteristics impact choice in retail settings with many zero observations: an application to restaurant wine purchase', *Journal of Agricultural and Resource Economics*, **29**, 112–131.

Edmeades, S. & Smale, M. (2006), 'A trait-based model of the potential demand for a genetically engineered food crop in a developing economy', *Agricultural Economics*, **35**, 351–361.
doi:10.1111/j.1574-0862.2006.00167.x

Feng, J. & Zhu, Z. (2011), 'Semiparametric analysis of longitudinal zero-inflated count data', *Journal of Multivariate Analysis*, **102**, 61–72.
doi:10.1016/j.jmva.2010.08.001

Frondel, M. & Vance, C. (2011), 'Rarely enjoyed? A count data analysis of ridership in Germany's public transport', *Transport Policy*, **18**, 425–433.
doi:10.1016/j.tranpol.2010.09.009

Gilles, R. (2012), *An Economic Analysis of Secure Messaging between Patients and Providers*, Ph.D. Dissertation, University of Washington, Seattle.
https://digital.lib.washington.edu/dspace/handle/1773/20865

Gilles, R. & Kim, S. (2013), 'Distribution-free estimation of zero-inflated models with unobserved heterogeneity', *University of Washington, Department of Economics, Working Papers*, **UWEC-2013-03**.
http://faculty.washington.edu/seikkim/seikkim_qcmle.pdf

Gurmu, S. & Trivedi, P.K. (1996), 'Excess zeros in count models for recreational trips', *Journal of Business and Economic Statistics*, **14**, 469-477.
doi:10.1080/07350015.1996.10524676

Hall, D. (2000), 'Zero-inflated Poisson and binomial regression with random effects: a case study', *Biometrics*, **56**, 1030–1039.
doi:10.1111/j.0006-341X.2000.01030.x

Hansen, L.P. (1982), 'Large sample properties of generalized method of moments estimators', *Econometrica*, **50**, 1029–1054.
doi:10.2307/1912775

Hasan, M.T. & Sneddon, G. (2009), 'Zero-inflated Poisson regression for longitudinal data', *Communications in Statistics - Simulation and Computation*, **38**, 638–653.
doi:10.1080/03610910802601332

Hausman, J.A., Hall B.H. & Griliches, Z. (1984), 'Econometric models for count data with an appication to the patent-R&D relationship', *Econometrica*, **52**, 909–938.
doi:10.2307/1911191

Holtz-Eakin, D., Newey, W. & Rosen, H.S. (1988) 'Estimating Vector Autoregressions with Panel Data', *Econometrica*, **56**, 1371–1395.
doi:10.2307/1913103

Hsueh, S.P. & Lee, W.M. (2009), 'Empirical likelihood estimation for panel count data models with fixed effects', *mimeo*.
http://econ.ccu.edu.tw/2009/conference/3A1.pdf

Hsueh, S.P. & Lee, W.M. (2012), 'A revisit to the relationship between patents and R&D using empirical likelihood estimation', *Economics Bulletin*, **32**, 1208–1214.
http://www.accessecon.com/Pubs/EB/2012/Volume32/EB-12-V32-I2-P115.pdf

Hu, A.G. & Jefferson, G.H. (2009), 'A great wall of patents: what is behind China's recent patent explosion?', *Journal of Development Economics*. **90**, 57–68.
doi:10.1016/j.jdeveco.2008.11.004

Kitazawa, Y. (2012), 'Hyperbolic transformation and average elasticity in the framework of the fixed effects logit model', *Theoretical Economics Letters*, **2**, 191–198.
doi:10.4236/tel.2012.22034

Lam, K.F., Xue, H. & Cheung, Y.B. (2006), 'Semiparametric analysis of zero-inflated count data', *Biometrics*, **62**, 996–1003.
doi:10.1111/j.1541-0420.2006.00575.x

Lambert, D. (1992), 'Zero-inflated Poisson regression, with an application to defects in manufacturing', *Technometrics*, **34**, 1–14.
doi:10.1080/00401706.1992.10485228

List, J. (2001), 'Determinants of securing academic interviews after tenure denial: evidence from a zero-inflated Poisson model', *Applied Economics*, **33**, 1423–1431.
doi:10.1080/00036840010009856

Majo, M.C. (2010), *A Microeconometric Analysis of Health Care Utilization in Europe*, Ph.D. Dissertation, Tilburg University, Tilburg.
http://oudesite.uvt.nl/research/institutes-and-research-groups/center/graduate-school/thesis/majo.html

Majo, M.C. & Van Soest, A. (2011), 'The fixed-effects zero-inflated Poisson model with an application to health care utilization', *Tilburg University, Center for Economic Research, Discussion Paper*, **No. 2011-083**.
http://arno.uvt.nl/show.cgi?fid=115295

Min, Y. & Agresti, A. (2005), 'Random effect models for repeated measures of zero-inflated count data', *Statistical Modeling*, **5**, 1–19.
doi:10.1191/1471082X05st084oa

Newey, W.K. & Smith, R.J. (2004), 'Higher order properties of GMM and generalized empirical likelihood estimators', *Econometrica*, **72**, 219–255.
doi:10.1111/j.1468-0262.2004.00482.x

Owen, A.B. (1988), 'Empirical likelihood ratio confidence intervals for a single functional', *Biometrika*, **75**, 237–249.

Owen, A.B. (1990). 'Empirical likelihood confidence regions', *Annals of Statistics*, **18**, 90–120.

Owen, A.B. (1991), 'Empirical likelihood for linear models', *Annals of Statistics*, **19**, 1725–1747.

Owen, A.B. (2001), *Empirical likelihood*, Chapman and Hall, London.

Qin, J. & Lawless, J. (1994), 'Empirical likelihood and generalized estimating equations', *Annals of Statistics*, **22**, 300–325.

Ramalho, J.J.S. (2005), 'Small sample bias of alternative estimation methods for moment condition models: Monte Carlo evidence for covariance structures', *Studies in Nonlinear Dynamics & Econometrics*, **9**, 1–18.
doi:10.2202/1558-3708.1202

Rasch, G. (1960), *Probabilistic models for some intelligence and attainment tests*, The Danish Institute for Educational Research, Copenhagen.

Rasch, G. (1961), 'On general laws and the meaning of measurement in psychology', *Preceeding of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, **4**, 321–333.

R Core Team (2013), *R: A language and environment for statistical computing*, R Foundation for Statistical Computing, Vienna.

Staiger, D. & Stock, J.H. (1997), 'Instrumental Variables Re- gression with Weak Instruments', *Econometrica*, textbf65, 557–586.
doi:10.2307/2171753

Tomlin, K.S.M. (2000), 'The effects of model specification on foreign direct investment models: an application of count data models', *Southern Economic Journal*, **67**, 460–468.
doi:10.2307/1061481

Tribus, M. (1961), *Thermodynamics and thermostatics: an introduction to energy, information and states of matter, with engineering applications*, D. Van Nostrand Company Inc., New York.

Wooldridge, J.M. (1997), 'Multiplicative panel data models without the strict exogeneity assumption', *Econometric Theory*, **13**, 667–678.
doi:10.1017/S0266466600006125

Zeileis, A., Kleiber, C. & Jackman, S. (2008), 'Regression models for count data in R', *Journal of Statistical Software*, **27**, 1–25.
http://www.jstatsoft.org/v27/i08/

Errata information

The errata information (if any) is available from the following site:
http://www.ip.kyusan-u.ac.jp/J/kitazawa/ERRATA/errata_fezip.html

Table 1: Monte Carlo results for the fixed effects ZIP model, $T = 4$

		N = 1000		N = 5000		N = 10000	
		bias	rmse	bias	rmse	bias	rmse
GMM(PHI)	δ	0.128	0.776	0.044	0.276	0.029	0.198
	β	0.003	0.110	0.002	0.054	0.001	0.039
GMM(PSI)	δ	0.935	43.754	0.064	0.399	0.039	0.242
	β	-0.092	0.277	-0.029	0.133	-0.014	0.091
EL(PHI)	δ	0.093	0.632	0.025	0.273	0.016	0.198
	β	0.004	0.113	0.002	0.056	0.001	0.041
EL(PSI)	δ	0.098	0.642	0.030	0.276	0.019	0.200
	β	-0.050	0.334	-0.006	0.134	0.000	0.090
ML(POOL)	δ	0.343	0.350	0.342	0.343	0.341	0.342
	β	0.476	0.479	0.477	0.478	0.477	0.478

Notes: 1) The initial consistent estimates used for the GMM estimations are the GMM estimates obtained by using the identity matrix as the weighting matrix. 2) Inappropriate replications (i.e. the non-convergence replications) are eliminated in calculating the statistics. Their number is zero or extremely small. 3) The values of the Monte Carlo statistics are obtained using the true values of the parameters of interest as the starting values in the optimization for each replication. The values of the statistics obtained using the true values are almost the same as those obtained using two different types of the starting values. 4) The starting values of Lagrange multipliers are zero in the optimization. 5) The Monte Carlo mean of proportions of zeros for the count dependent variables is about 70 percent.

Figure 1: Monte Carlo boxplots of the GMM and EL estimates for δ, $T = 4$

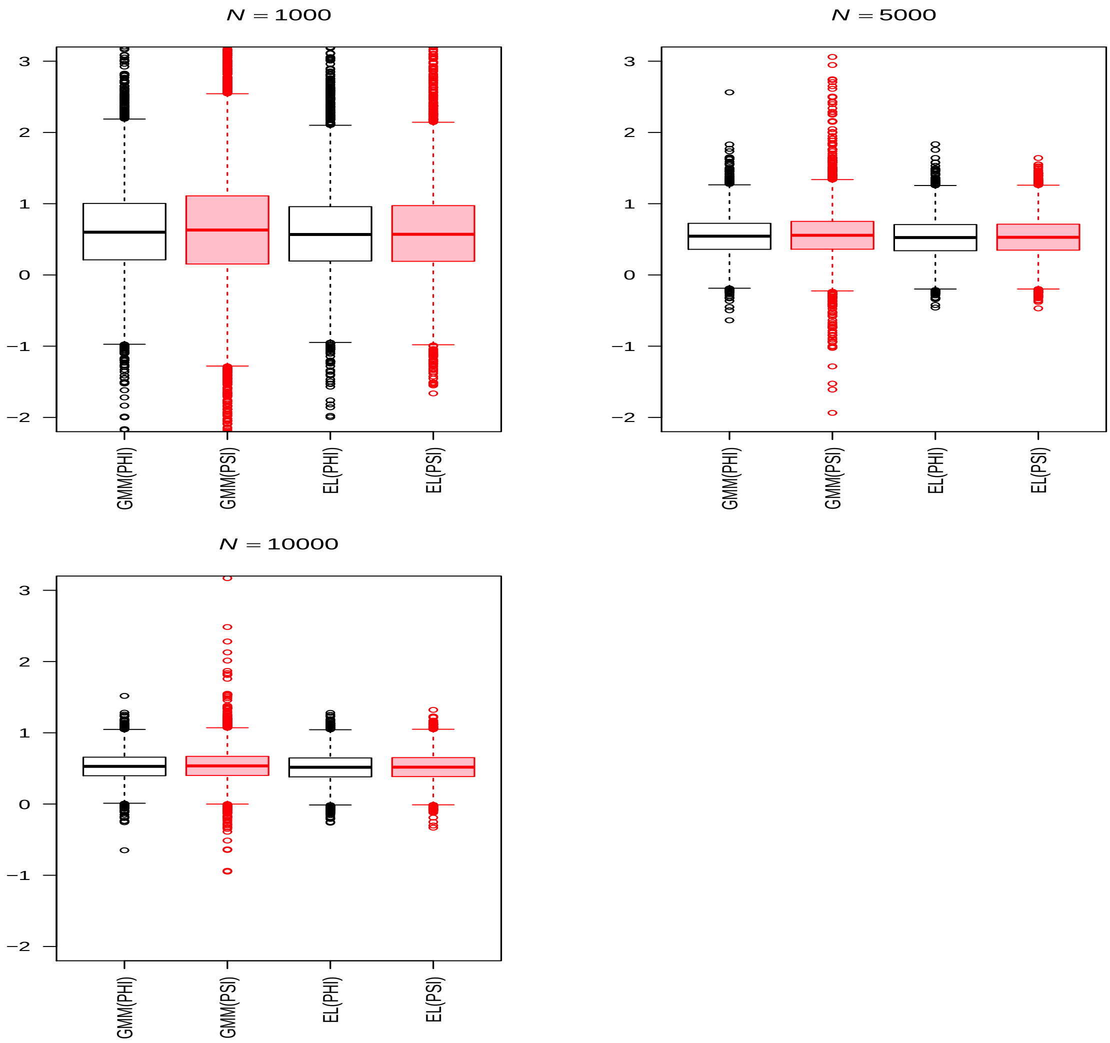

Figure 2: Monte Carlo boxplots of the GMM and EL estimates for β, $T = 4$

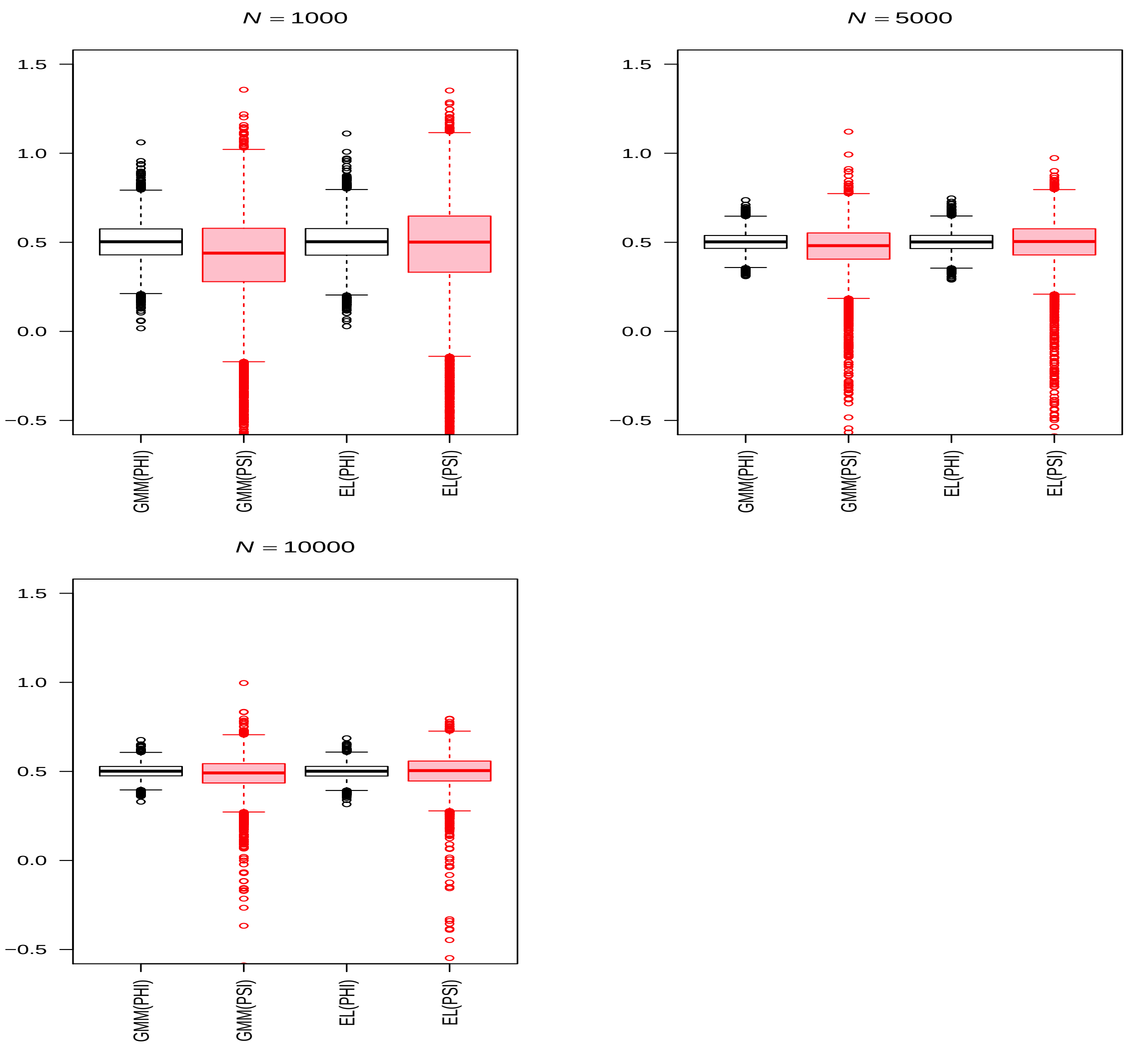

Table 2: Monte Carlo results for the fixed effects ZIP model, $T = 8$

		$N = 1000$		$N = 5000$		$N = 10000$	
		bias	rmse	bias	rmse	bias	rmse
GMM(PHI)	δ	0.085	2.633	0.046	0.212	0.032	0.131
	β	0.001	0.073	0.001	0.035	0.000	0.026
GMM(PSI)	δ	1.982	106.559	0.043	3.052	0.061	1.761
	β	-0.114	0.222	-0.032	0.080	-0.017	0.051
EL(PHI)	δ	0.082	0.395	0.027	0.177	0.018	0.130
	β	0.003	0.072	0.001	0.036	0.000	0.027
EL(PSI)	δ	0.119	0.440	0.040	0.178	0.025	0.129
	β	0.015	0.116	0.005	0.055	0.003	0.040
ML(POOL)	δ	0.342	0.346	0.341	0.342	0.341	0.342
	β	0.476	0.479	0.477	0.478	0.477	0.477

Notes: 1) The initial consistent estimates used for the GMM estimations are the GMM estimates obtained by using the identity matrix as the weighting matrix. 2) Inappropriate replications (i.e. the non-convergence replications) are eliminated in calculating the statistics. Their number is zero or extremely small. 3) The values of the Monte Carlo statistics are obtained using the true values of the parameters of interest as the starting values in the optimization for each replication. The values of the statistics obtained using the true values are almost the same as those obtained using two different types of the starting values. 4) The starting values of Lagrange multipliers are zero in the optimization. 5) The Monte Carlo mean of proportions of zeros for the count dependent variables is about 70 percent.

Figure 3: Monte Carlo boxplots of the GMM and EL estimates for δ, $T = 8$

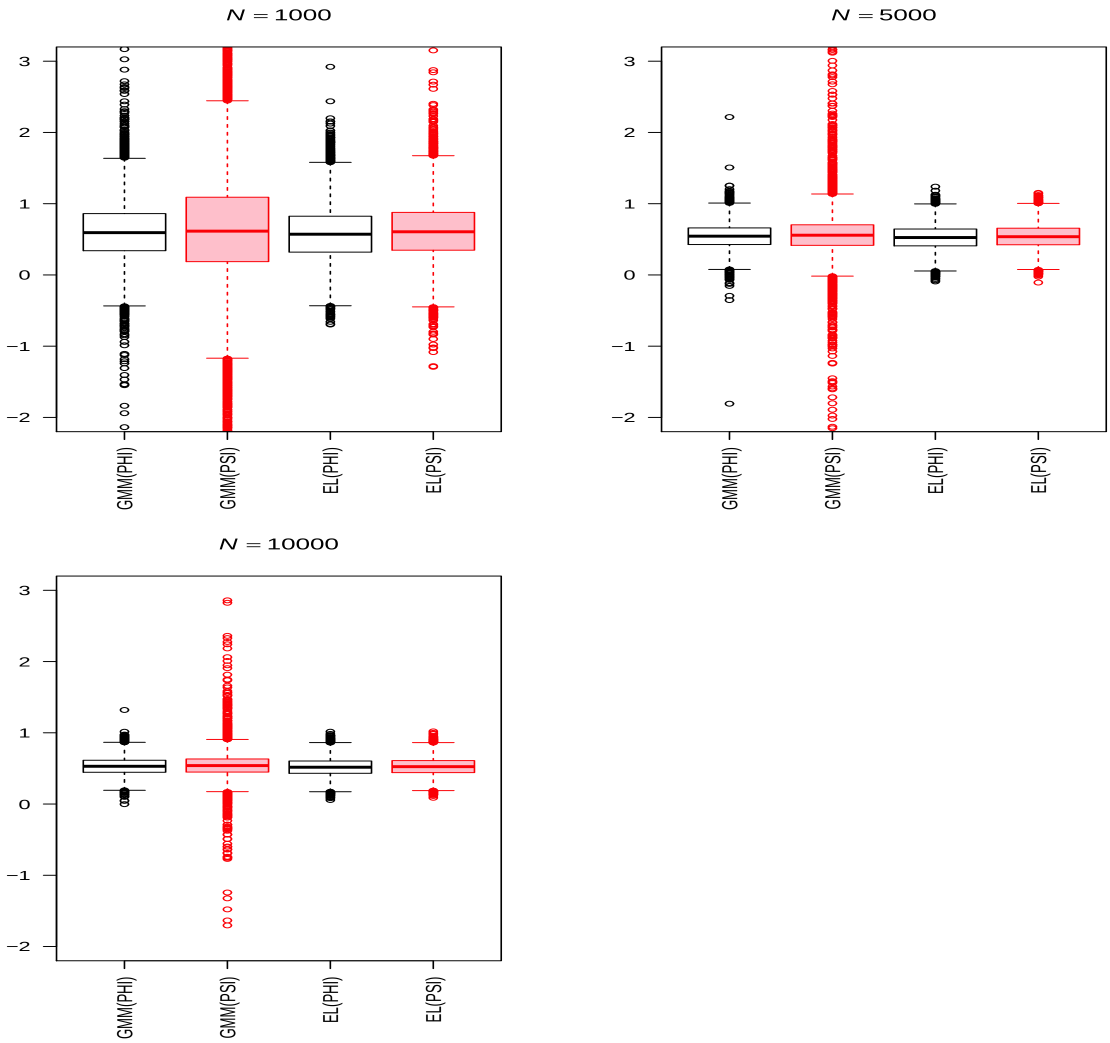

Figure 4: Monte Carlo boxplots of the GMM and EL estimates for β, $T = 8$