A negative binomial model and moment conditions for count panel data*

Yoshitsugu Kitazawa†

June 15, 2009

Abstract
This paper proposes some moment conditions associated with an appropriate specification of negative binomial model for count panel data, which is proposed by Hausman et al. (1984). The newly proposed moment conditions enable researchers to conduct the consistent estimation of the model under much weaker assumptions than those configured by Hausman et al. (1984). In some Monte Carlo experiments, it is shown that the GMM estimators using the new moment conditions perform well in the DGP configurations conforming to the specification above.

Keywords: count panel data, predetermined explanatory variable, linear feedback model, overdispersion, negative binomial model, implicit operation, cross-linkage moment conditions, GMM, Monte Carlo experiments

JEL classification: C23

1. Introduction

It is often said that for count data the variance exceeds the mean (see Cameron and Trivedi, 2005, etc). For count panel data model with fixed effects when number of cross-sectional units is large and number of time periods is small, Hausman et al. (1984) propose a specification of the overdispersion based on the negative binomial model and an estimator for it. However, their estimator, which employs the conditional maximum likelihood approach, is not consistent when the explanatory variables are predetermined and/or the model is dynamic. In count panel data model, it is much admissible to regard the explanatory variables as being predetermined instead of being strictly exogenous and presumably much preferable to incorporate dynamics into the model. To take as an example a patent production function of a firm where number of patents as a flow variable is a function of R&D expenditures, it is conceivable that the current number of patents affects the future R&D expenditures as well as the current and past R&D expenditures affect the current number of patents. In addition, it is quite likely that the past numbers of patents affect the current number of patents.

For the case of allowing for the predetermined explanatory variables and the dynamics, the distribution-free GMM (generalized method of moments) estimators proposed by Hansen (1982) are exclusively utilized by using the moment conditions proposed by Chamberlain (1992), Wooldridge (1997), Windmeijer (2000), Blundell et al. (2002) and Kitazawa (2007) for the purpose of consistent estimations, except for the case where the pre-sample mean (PSM) estimator proposed

* Discussion Paper Series, Faculty of Economics, Kyushu Sangyo University Discussion Paper, June 2009, No. 34
Errata information: http://www.ip.kyusan-u.ac.jp/J/kitazawa/ERRATA/errata_negb.html
† Correspondence: Faculty of Economics, Kyushu Sangyo University, Matsukadai 2-3-1, Higashi-ku, Fukuoka, Japan.
E-mail: kitazawa@ip.kyusan-u.ac.jp

by Blundell et al. (1999,2002) is usable.[1] However, no moment condition is developed which gives the overdispersion a distinction, although the moment conditions associated with the equidispersion are developed by Kitazawa (2007, 2009).

In this paper, some moment conditions are proposed in association with a specification of the overdispersion for the count panel data model. The specification is based on the fixed effects negative binomial model proposed by Hausman et al. (1984) and allows for the dynamics and the predetermined explanatory variables in the model.[2] The moment conditions are constructed by using the implicit operation proposed by Kitazawa (2007) and are of the form of the cross-linkage moment conditions setting up the relationships between variances and covariances in the disturbances in the model.[3]

Some Monte Carlo experiments are carried out for both configurations of the equidispersion and the overdispersion. The experiments show that for the larger cross-sectional size, the GMM estimators incorporating the cross-linkage moment conditions associated with the overdispersion never perform better and remain biased for the configuration of the equidisperion, reflecting the inconsistency, while they perform better for the configuration of the overdispersion, reflecting the consistency and that the usage of the cross-linkage moment conditions associated with the overdispersion improve or do not at least vitiate the small sample performances for the configuration of the overdispersion.

The rest of the paper is organized as follows. In section 2, the cross-linkage moment conditions are proposed with respect to the overdispersion. In section 3, some Monte Carlo experiments are carried out. Section 4 concludes.

2. Model, moment conditions and GMM estimators

In this section, some sets of the moment conditions associated with the overdispersion in the framework of Hausman et al. (1984) for the linear feedback model (LFM) proposed by Blundell et al. (2002) in count panel data are proposed for the three cases: the case of predetermined explanatory variables, the case of strictly exogenous explanatory variables and the case of mean-stationary dependent variables. The method of deriving these sets is that based on the implicit operation proposed by Kitazawa (2007) and the moment conditions proposed in this paper are constructed in the framework of the cross-linkage moment conditions proposed by Kitazawa (2009).[4] Then, the GMM estimators are constructed by using the cross-linkage moment conditions.

2.1. Linear feedback model

A simple form of the linear feedback model (LFM) proposed by Blundell et al. (2002) is as follows:

$$y_{it}=\gamma y_{i,t-1}+\exp(\beta x_{it}+\eta_i)+v_{it} , \qquad \text{for} \quad t=2,\dots,T , \tag{2.1.1}$$

where the subscript i denotes the individual unit with $i=1,\dots,N$, t denotes the time period and it is assumed that T is fixed and $N\rightarrow\infty$. The count dependent variable y_{it} is able to have zero or positive integer values and the explanatory variable x_{it} is able to have the real number. The unobservable variables η_i and v_{it} are the individual specific effect and the

1 For the case of allowing for the strictly exogenous explanatory variables and the dynamics, the distribution-free GMM estimators are also available by using the moment conditions proposed by Crépon and Duguet (1997) for the purpose of consistent estimations (see Kitazawa, 2007).

2 The fixed effects negative binomial model is described in Winkelmann (2008) in a way easy to understand.

3 The cross-linkage moment conditions are proposed by Kitazawa (2009), associated with the equidispersion.

4 Ahn (1990) and Ahn and Schmidt (1995) propose the method of constructing the efficient set of the moment conditions based on the error components in the framework of ordinary dynamic panel data model. The implicit operation is developed with the aim of incorporating their method into the count panel data model.

disturbance respectively. The parameters of interest are γ (with $|\gamma|<1$) and β .

Equation (2.1.1) is rewritten as follows:

$$y_{it}=\gamma y_{i,t-1}+u_{it} \ , \qquad \text{for} \quad t=2,\ldots,T \ , \qquad (2.1.2)$$

$$u_{it}=\phi_i \mu_{it}+v_{it} \ , \qquad \text{for} \quad t=2,\ldots,T \ , \qquad (2.1.3)$$

where $\phi_i=\exp(\eta_i)$ and $\mu_{it}=\exp(\beta x_{it})$. Based on (2.1.2), it can be seen that u_{it} is observable in the sense that it is written in terms of data and parameter. That is, $u_{it}=y_{it}-\gamma y_{i,t-1}$, which is plugged into the moment conditions to be hereafter described.

2.2. Case of predetermined explanatory variables

In this case, the assumption on the disturbance v_{it} is

$$E[v_{it}\,|\,y_{i1},\eta_i,v_i^{t-1},x_i^t]=0 \ , \qquad \text{for} \quad t=2,\ldots,T \ , \qquad (2.2.1)$$

where $v_i^{t-1}=(v_{i1},\ldots,v_{i,t-1})$ and $x_i^t=(x_{i1},\ldots,x_{it})$. The assumption (2.2.1) is referred to as the "original assumption" for the case of predetermined explanatory variables. Kitazawa (2007) constructs the implicit standard assumptions from the original assumption (2.2.1) as follows:

$$E[y_{i1}v_{it}\,|\,y_{i1},\eta_i,v_i^{t-1},x_i^t]=0 \ , \qquad (2.2.2)$$

$$E[v_{is}v_{it}\,|\,y_{i1},\eta_i,v_i^{t-1},x_i^t]=0 \ , \ \text{for} \quad 2\le s\le t-1 \ , \qquad (2.2.3)$$

$$E[x_{is}v_{it}\,|\,y_{i1},\eta_i,v_i^{t-1},x_i^t]=0 \ , \ \text{for} \quad 1\le s\le t \ , \qquad (2.2.4)$$

$$E[\eta_i v_{it}\,|\,y_{i1},\eta_i,v_i^{t-1},x_i^t]=0 \ . \qquad (2.2.5)$$

Here, the following assumption with respect to the overdispersion specified in the framework of Hausman et al. (1984) is imposed in addition to the implicit standard assumptions (2.2.2) – (2.2.5):

$$E[(\,v_{it}^2-(1+\phi_i)y_{it}\,)\,|\,y_{i1},\eta_i,v_i^{t-1},x_i^t]=0 \ , \qquad \text{for} \quad t=2,\ldots,T \ , \qquad (2.2.6)$$

where the construction of (2.2.6) is written in Appendix A.

Kitazawa (2007, 2009) obtains some moment conditions for estimating γ and β consistently under the assumptions (2.2.2) – (2.2.5) and the assumption of the equidispersion instead of (2.2.6). In this paper, the moment conditions associated with the overdispersion are constructed under the assumptions (2.2.2) – (2.2.5) and the assumption (2.2.6).[5] The implicit operation proposed by Kitazawa (2007) is used for constructing the moment conditions.

According to Kitazawa (2007), the observable analogues for (2.2.2) – (2.2.4) are obtained by replacing the unobservable variables v_{it} by the observable variables u_{it} :

5 That is, the cross-linkage moment conditions associated with the overdispersion are obtained under the assumption (2.2.1) with (2.2.6).

$$E[y_{i1}u_{it} \mid y_{i1}, \eta_i, v_i^{t-1}, x_i^t] = y_{i1}\phi_i\mu_{it} \ , \quad (2.2.7)$$

$$E[u_{is}u_{it} \mid y_{i1}, \eta_i, v_i^{t-1}, x_i^t] = \phi_i^2\mu_{is}\mu_{it} + v_{is}\phi_i\mu_{it} \ , \quad \text{for} \quad 2 \le s \le t-1 \ , \quad (2.2.8)$$

$$E[x_{is}u_{it} \mid y_{i1}, \eta_i, v_i^{t-1}, x_i^t] = x_{is}\phi_i\mu_{it} \ , \quad \text{for} \quad 1 \le s \le t \ , \quad (2.2.9)$$

respectively. In addition, the observable analogues for (2.2.6) are obtained by replacing the unobservable variables v_{it} by the observable variables u_{it} and further replacing the unobservable variables $\phi_i y_{it}$ by the observable variables $(y_{it}u_{it} - u_{it}^2)/\mu_{it}$:

$$E[((u_{it}^2 - y_{it}) + (1/\mu_{it})(u_{it}^2 - y_{it}u_{it})) \mid y_{i1}, \eta_i, v_i^{t-1}, x_i^t] = \phi_i^2\mu_{it}(\mu_{it}+1) \ ,$$

$$\text{for} \quad t = 2, \dots, T \ , \quad (2.2.10)$$

where the derivation of (2.2.10) is described in Appendix B.

When the explanatory variables are predetermined, the condensed full set of moment conditions associated with the overdispersion is obtained in the framework of constructing that associated with the equidispersion. Kitazawa (2009) constructs a condensed full set of moment conditions associated with the equidispersion. It is derived from both of the relationships between $y_{i1}u_{i,t-1}$ and $y_{i1}u_{it}$ for $t=3,\dots,T$ and the relationships between $u_{is}u_{i,t-1}$ and $u_{is}u_{it}$ for $s=2,\dots,t-2$ and $t=4,\dots,T$, both of the relationships between $u_{i,t-1}^2$ and $u_{i,t-1}u_{it}$ for $t=3,\dots,T$ and the relationships between $u_{i,t-1}u_{it}$ and u_{it}^2 for $t=3,\dots,T$, and the relationships between $x_{is}u_{i,t-1}$ and $x_{is}u_{it}$ for $s=1,\dots,t-1$ and $t=3,\dots,T$, all of which are solved through the intermediary of the unconditional expectation operator after weighting them with appropriate transformations of explanatory variables x_{it} for $t=1,\dots,T$. The same relationships are used for deriving the full set of moment conditions for the case of the overdispersion as for the case of the equidispersion.

From the relationships between $y_{i1}u_{i,t-1}$ and $y_{i1}u_{it}$ for $t=3,\dots,T$, the relationships between $u_{is}u_{i,t-1}$ and $u_{is}u_{it}$ for $s=2,\dots,t-2$ and $t=4,\dots,T$ and the relationships between $x_{is}u_{i,t-1}$ and $x_{is}u_{it}$ for $s=1,\dots,t-1$ and $t=3,\dots,T$, Kitazawa (2007) obtains the following $(T-2)(T-1)/2$ and $(T-1)T/2-1$ quasi-differenced moment conditions based on the transformation proposed by Chamberlain (1992) and Wooldridge (1997):

$$E[y_{is}((\mu_{i,t-1}/\mu_{it})u_{it} - u_{i,t-1})] = 0 \ , \text{for} \quad s=1,\dots,t-2 \ ; \quad t=3,\dots,T \ , \quad (2.2.11)$$

$$E[x_{is}((\mu_{i,t-1}/\mu_{it})u_{it} - u_{i,t-1})] = 0 \ , \text{for} \quad s=1,\dots,t-1 \ ; \quad t=3,\dots,T \ , \quad (2.2.12)$$

where the quasi-differenced moment conditions (2.2.11) are those extended as the application to the LFM in Blundell et al. (2002). The moment conditions (2.2.11) and (2.2.12) hold even if the assumption (2.2.6) is not imposed.

The moment conditions based on the relationships between $u_{i,t-1}^2$ and $u_{i,t-1}u_{it}$ for $t=3,\dots,T$ and the relationships between $u_{i,t-1}u_{it}$ and u_{it}^2 for $t=3,\dots,T$ are referred to as the cross-linkage moment conditions, according to Kitazawa (2009). For the case of the equidispersion and predetermined explanatory variables, Kitazawa (2009) derives the cross-linkage moment conditions. From now on, two types of the cross-linkage moment conditions are solved for the case of the overdispersion and predetermined explanatory variables by using the relationships between $u_{i,t-1}^2$ and $u_{i,t-1}u_{it}$ for $t=3,\ldots,T$ and the relationships between $u_{i,t-1}u_{it}$ and u_{it}^2 for $t=3,\ldots,T$.

First, the relationship is solved between the transformation using $u_{i,t-1}^2$ (i.e. $(u_{i,t-1}^2-y_{i,t-1})+(1/\mu_{i,t-1})(u_{i,t-1}^2-y_{i,t-1}u_{i,t-1})$) and $u_{i,t-1}u_{it}$ (weighted with $(\mu_{i,t-1}+1)/\mu_{it}$), through the intermediary of the unconditional expectation operator. Multiplying both sides of (2.2.8) for $s=t-1$ by $(\mu_{i,t-1}+1)/\mu_{it}$ gives

$$E[u_{i,t-1}((\mu_{i,t-1}+1)/\mu_{it})u_{it}\,|\,y_{i1},\eta_i,v_i^{t-1},x_i^t]=\phi_i^2\mu_{i,t-1}(\mu_{i,t-1}+1)+v_{i,t-1}\phi_i(\mu_{i,t-1}+1)\ . \tag{2.2.13}$$

Applying the law of total expectation to (2.2.10) dated $t-1$ and (2.2.13), it follows that

$$E[((u_{i,t-1}^2-y_{i,t-1})+(1/\mu_{i,t-1})(u_{i,t-1}^2-y_{i,t-1}u_{i,t-1}))]=E[\phi_i^2\mu_{i,t-1}(\mu_{i,t-1}+1)]\ , \tag{2.2.14}$$

$$E[u_{i,t-1}((\mu_{i,t-1}+1)/\mu_{it})u_{it}]=E[\phi_i^2\mu_{i,t-1}(\mu_{i,t-1}+1)]\ . \tag{2.2.15}$$

Subtracting (2.2.14) from (2.2.15) gives

$$\begin{aligned}&E[u_{i,t-1}((\mu_{i,t-1}/\mu_{it})u_{it}-u_{i,t-1})]\ +\ E[y_{i,t-1}]\\ &\ +\ E[u_{i,t-1}((1/\mu_{it})u_{it}-(1/\mu_{i,t-1})u_{i,t-1})]\ +\ E[y_{i,t-1}(1/\mu_{i,t-1})u_{i,t-1}]=0\ ,\end{aligned} \tag{2.2.16}$$

At this stage, it should be noted that the following two relationships hold under the assumptions (2.2.1):

$$E[u_{i,t-1}((\mu_{i,t-1}/\mu_{it})u_{it}-u_{i,t-1})]=E[y_{i,t-1}((\mu_{i,t-1}/\mu_{it})u_{it}-u_{i,t-1})]\ , \tag{2.2.17}$$

$$E[u_{i,t-1}((1/\mu_{it})u_{it}-(1/\mu_{i,t-1})u_{i,t-1})]=E[y_{i,t-1}((1/\mu_{it})u_{it}-(1/\mu_{i,t-1})u_{i,t-1})]\ , \tag{2.2.18}$$

whose derivations are written in Appendix C. Accordingly, plugging (2.2.17) and (2.2.18) into (2.2.16) gives the following $T-2$ cross-linkage moment conditions:

$$E[y_{i,t-1}(((\mu_{i,t-1}+1)/\mu_{it})u_{it}-(u_{i,t-1}-1))]=0\ , \qquad \text{for}\quad t=3,\ldots,T\ , \tag{2.2.19}$$

in which the order reduction with respect to y is realized, compared to (2.2.16).

Next, the relationship through the intermediary of the unconditional expectation operator is solved between $u_{i,t-1}u_{it}$ (weighted with $1/\mu_{it}$) and the transformation of u_{it}^2 (i.e. $(u_{it}^2-y_{it})+(1/\mu_{it})(u_{it}^2-y_{it}u_{it})$ weighted with $\mu_{i,t-1}/(\mu_{it}(\mu_{it}+1))$). Multiplying (2.2.8) for $s=t-1$ by $1/\mu_{it}$ gives

$$E[u_{i,t-1}u_{it}(1/\mu_{it})\,|\,y_{i1},\eta_i,v_i^{t-1},x_i^t]=\phi_i^2\mu_{i,t-1}+v_{i,t-1}\phi_i \qquad (2.2.20)$$

and multiplying (2.2.10) by $\mu_{i,t-1}/(\mu_{it}(\mu_{it}+1))$ gives

$$\begin{aligned}&E[(\ (u_{it}^2-y_{it})(\ \mu_{i,t-1}/(\mu_{it}(\mu_{it}+1))\ )\ )\,|\,y_{i1},\eta_i,v_i^{t-1},x_i^t]\\&\quad+\ E[(\ (1/\mu_{it})(u_{it}^2-y_{it}u_{it})(\ \mu_{i,t-1}/(\mu_{it}(\mu_{it}+1))\ )\ )\,|\,y_{i1},\eta_i,v_i^{t-1},x_i^t]\ =\ \phi_i^2\mu_{i,t-1}\end{aligned}\ .$$

(2.2.21)

Applying the law of total expectation to (2.2.20) and (2.2.21), it follows that

$$E[u_{i,t-1}u_{it}(1/\mu_{it})]=E[\phi_i^2\mu_{i,t-1}]\ , \qquad (2.2.22)$$

$$\begin{aligned}&E[(\ (u_{it}^2-y_{it})(\ \mu_{i,t-1}/(\mu_{it}(\mu_{it}+1))\ )\ )]\\&\quad+\ E[(\ (1/\mu_{it})(u_{it}^2-y_{it}u_{it})(\ \mu_{i,t-1}/(\mu_{it}(\mu_{it}+1))\ )\ )]\ =\ E[\phi_i^2\mu_{i,t-1}]\end{aligned}\ . \qquad (2.2.23)$$

Subtracting (2.2.22) from (2.2.23), the following $T-2$ cross-linkage moment conditions are obtained:

$$E[(1/\mu_{it})(\ ((\mu_{i,t-1}/\mu_{it})u_{it}-u_{i,t-1})u_{it}-y_{it}(\mu_{i,t-1}/\mu_{it})u_{it}+y_{it}(\mu_{i,t-1}/(\mu_{it}+1))(u_{it}-1)\ )]=0\ ,$$

$$\text{for}\quad t=3,\ldots,T\ . \qquad (2.2.24)$$

The detail of derivation of (2.2.24) is written in Appendix D.

Eventually, a condensed full set of the moment conditions for the case where the assumption with respect to the overdispersion is imposed in addition to the implicit standard assumptions associated with predetermined explanatory variables is composed of (2.2.11), (2.2.19), (2.2.24) and (2.2.12). That is, under the assumption (2.2.1) with (2.2.6), the condensed full set is composed of the moment conditions (2.2.11), (2.2.19), (2.2.24) and (2.2.12). The moment conditions (2.2.11), (2.2.19) and (2.2.12) are linear with respect to γ , while (2.2.24) nonlinear.

2.3. Case of strictly exogenous explanatory variables

In this case, the assumption on the disturbance v_{it} is

$$E[v_{it}\,|\,y_{i1},\eta_i,v_i^{t-1},x_i^T]=0\ ,\quad \text{for}\quad t=2,\ldots,T\ , \qquad (2.3.1)$$

where $v_i^{t-1}=(v_{i1},\ldots,v_{i,t-1})$ and $x_i^T=(x_{i1},\ldots,x_{iT})$. The assumption (2.3.1) is referred to as the "original assumption" for the case of strictly exogenous explanatory variables. Kitazawa (2007) constructs the implicit standard assumptions from the original assumption (2.3.1) as follows:

$$E[y_{i1}v_{it} \mid y_{i1}, \eta_i, v_i^{t-1}, x_i^T]=0 \quad , \tag{2.3.2}$$

$$E[v_{is}v_{it} \mid y_{i1}, \eta_i, v_i^{t-1}, x_i^T]=0 \quad \text{, for} \quad 2 \le s \le t-1 \quad , \tag{2.3.3}$$

$$E[x_{is}v_{it} \mid y_{i1}, \eta_i, v_i^{t-1}, x_i^T]=0 \quad \text{, for} \quad 1 \le s \le T \quad , \tag{2.3.4}$$

$$E[\eta_i v_{it} \mid y_{i1}, \eta_i, v_i^{t-1}, x_i^T]=0 \quad . \tag{2.3.5}$$

Here, the following assumption with respect to the overdispersion specified in the framework of Hausman et al. (1984) is imposed in addition to the implicit standard assumptions (2.3.2) – (2.3.5):

$$E[(v_{it}^2-(1+\phi_i)y_{it}) \mid y_{i1}, \eta_i, v_i^{t-1}, x_i^T]=0 \quad , \qquad \text{for} \quad t=2,\ldots,T \quad , \tag{2.3.6}$$

where the construction of (2.3.6) is the same as the contents written in Appendix A, except that the terminology "the predetermined explanatory variables" is replaced by "the strictly exogenous variables", x_i^t is replaced by x_i^T and (2.2.1) and (2.2.6) are replaced by (2.3.1) and (2.3.6) respectively.

Kitazawa (2007, 2009) obtains some moment conditions for estimating γ and β consistently under the assumptions (2.3.2) – (2.3.5) and the assumption of the equidispersion instead of (2.3.6). In this paper, the moment conditions associated with the overdispersion are constructed under the assumptions (2.3.2) – (2.3.5) and the assumption (2.3.6).[6] The implicit operation proposed by Kitazawa (2007) is used for constructing the moment conditions.

According to Kitazawa (2007), the observable analogues for (2.3.2) – (2.3.4) are obtained by replacing the unobservable variables v_{it} by the observable variables u_{it} :

$$E[y_{i1}u_{it} \mid y_{i1}, \eta_i, v_i^{t-1}, x_i^T]=y_{i1}\phi_i\mu_{it} \quad , \tag{2.3.7}$$

$$E[u_{is}u_{it} \mid y_{i1}, \eta_i, v_i^{t-1}, x_i^T]=\phi_i^2\mu_{is}\mu_{it}+v_{is}\phi_i\mu_{it} \quad , \qquad \text{for} \quad 2 \le s \le t-1 \quad , \tag{2.3.8}$$

$$E[x_{is}u_{it} \mid y_{i1}, \eta_i, v_i^{t-1}, x_i^T]=x_{is}\phi_i\mu_{it} \quad , \qquad \text{for} \quad 1 \le s \le T \quad , \tag{2.3.9}$$

respectively. In addition, the observable analogues for (2.3.6) are obtained by replacing the unobservable variables v_{it} by the observable variables u_{it} and further replacing the unobservable variables $\phi_i y_{it}$ by the observable variables $(y_{it}u_{it}-u_{it}^2)/\mu_{it}$:

$$E[((u_{it}^2-y_{it})+(1/\mu_{it})(u_{it}^2-y_{it}u_{it})) \mid y_{i1}, \eta_i, v_i^{t-1}, x_i^T]=\phi_i^2\mu_{it}(\mu_{it}+1) \quad ,$$
$$\text{for} \quad t=2,\ldots,T \quad , \tag{2.3.10}$$

where the derivation of (2.3.10) is the same as that described in Appendix B, except that x_i^t is

6 That is, the cross-linkage moment conditions associated with the overdispersion are obtained under the assumption (2.3.1) with (2.3.6).

replaced by x_i^T and (2.2.1), (2.2.6) and (2.2.10) are replaced by (2.3.1), (2.3.6) and (2.3.10) respectively.

When the explanatory variables are strictly exogenous, the condensed full set of moment conditions associated with the overdispersion is obtained in the framework of that associated with the equidispersion. Kitazawa (2009) construct a condensed full set of moment conditions associated with the equidispersion. It is derived from both of the relationships between $y_{i1}u_{i,t-1}$ and $y_{i1}u_{it}$ for $t=3,\ldots,T$ and the relationships between $u_{is}u_{i,t-1}$ and $u_{is}u_{it}$ for $s=2,\ldots,t-2$ and $t=4,\ldots,T$, both of the relationships between $u_{i,t-1}^2$ and $u_{i,t-1}u_{it}$ for $t=3,\ldots,T$ and the relationships between $u_{i,t-1}u_{it}$ and u_{it}^2 for $t=3,\ldots,T$, and the relationships between $x_{is}u_{i,t-1}$ and $x_{is}u_{it}$ for $s=1,\ldots,T$ and $t=3,\ldots,T$, all of which are solved through the intermediary of the unconditional expectation operator after weighting them with appropriate transformations of explanatory variables x_{it} for $t=1,\ldots,T$. The same relationships are used for deriving the full set of moment conditions for the case of the overdispersion as for the case of the equidispersion.

From the relationships between $y_{i1}u_{i,t-1}$ and $y_{i1}u_{it}$ for $t=3,\ldots,T$, the relationships between $u_{is}u_{i,t-1}$ and $u_{is}u_{it}$ for $s=2,\ldots,t-2$ and $t=4,\ldots,T$ and the relationships between $x_{is}u_{i,t-1}$ and $x_{is}u_{it}$ for $s=1,\ldots,T$ and $t=3,\ldots,T$, Kitazawa (2007) obtains the following $(T-2)(T-1)/2$ and $(T-2)T$ quasi-differenced moment conditions reformed for the case of strictly exogenous explanatory variables:

$$E[y_{is}(u_{it}-(\mu_{it}/\mu_{i,t-1})u_{i,t-1})]=0 \text{ , for } s=1,\ldots,t-2 \text{ ; } t=3,\ldots,T \text{ ,} \tag{2.3.11}$$

$$E[x_{is}(u_{it}-(\mu_{it}/\mu_{i,t-1})u_{i,t-1})]=0 \text{ , for } s=1,\ldots,T \text{ ; } t=3,\ldots,T \text{ .} \tag{2.3.12}$$

The moment conditions (2.3.11) and (2.3.12) hold even if the assumption (2.3.6) is not imposed.

The moment conditions based on the relationships between $u_{i,t-1}^2$ and $u_{i,t-1}u_{it}$ for $t=3,\ldots,T$ and the relationships between $u_{i,t-1}u_{it}$ and u_{it}^2 for $t=3,\ldots,T$ are referred to as the cross-linkage moment conditions, according to Kitazawa (2009). For the case of the equidispersion and strictly exogenous explanatory variables, Kitazawa (2009) derives the cross-linkage moment conditions. From now on, two types of the cross-linkage moment conditions are solved for the case of the overdispersion and strictly exogenous explanatory variables by using the relationships between $u_{i,t-1}^2$ and $u_{i,t-1}u_{it}$ for $t=3,\ldots,T$ and the relationships between $u_{i,t-1}u_{it}$ and u_{it}^2 for $t=3,\ldots,T$.

First, the relationship is solved between the transformation using $u_{i,t-1}^2$ (i.e. $(u_{i,t-1}^2-y_{i,t-1})+(1/\mu_{i,t-1})(u_{i,t-1}^2-y_{i,t-1}u_{i,t-1})$ weighted with $\mu_{it}/(\mu_{i,t-1}+1)$) and $u_{i,t-1}u_{it}$, through the intermediary of the unconditional expectation operator. Multiplying both sides of (2.3.10) dated $t-1$ by $\mu_{it}/(\mu_{i,t-1}+1)$ gives

$$\begin{aligned}&E[((\mu_{it}/(\mu_{i,t-1}+1))(u_{i,t-1}^2-y_{i,t-1}))\,|\,y_{i1},\eta_i,v_i^{t-2},x_i^T]\\&+\;E[((\mu_{it}/(\mu_{i,t-1}+1))(1/\mu_{i,t-1})(u_{i,t-1}^2-y_{i,t-1}u_{i,t-1}))\,|\,y_{i1},\eta_i,v_i^{t-2},x_i^T]=\phi_i^2\mu_{i,t-1}\mu_{it}\end{aligned} .$$

(2.3.13)

Applying the law of total expectation to (2.3.13) and (2.3.8) for $s=t-1$, it follows that

$$E[((\mu_{it}/(\mu_{i,t-1}+1))(u_{i,t-1}^2-y_{i,t-1}))] + E[((\mu_{it}/(\mu_{i,t-1}+1))(1/\mu_{i,t-1})(u_{i,t-1}^2-y_{i,t-1}u_{i,t-1}))] = E[\phi_i^2\mu_{i,t-1}\mu_{it}] ,$$

(2.3.14)

$$E[u_{i,t-1}u_{it}]=E[\phi_i^2\mu_{i,t-1}\mu_{it}] . \quad (2.3.15)$$

Subtracting (2.3.14) from (2.3.15) gives

$$E[u_{i,t-1}(u_{it} - (\mu_{it}/(\mu_{i,t-1}+1))(1+(1/\mu_{i,t-1}))u_{i,t-1})] + E[y_{i,t-1}(\mu_{it}/(\mu_{i,t-1}+1))] + E[y_{i,t-1}(\mu_{it}/(\mu_{i,t-1}+1))(1/\mu_{i,t-1})u_{i,t-1}] = 0 .$$

(2.3.16)

At this stage, it should be noted that the following relationship holds:

$$E[u_{i,t-1}(u_{it} - (\mu_{it}/(\mu_{i,t-1}+1))(1+(1/\mu_{i,t-1}))u_{i,t-1})] = E[y_{i,t-1}(u_{it} - (\mu_{it}/(\mu_{i,t-1}+1))(1+(1/\mu_{i,t-1}))u_{i,t-1})] , \quad (2.3.17)$$

whose derivation is written in Appendix E. Accordingly, plugging (2.3.17) into (2.3.16) gives the following $T-2$ cross-linkage moment conditions:

$$E[y_{i,t-1}(u_{it}-(\mu_{it}/(\mu_{i,t-1}+1))(u_{i,t-1}-1))]=0 , \quad \text{for} \quad t=3,\ldots,T , \quad (2.3.18)$$

in which the order reduction with respect to y is realized, compared to (2.3.16).

Next, the relationship through the intermediary of the unconditional expectation operator is solved between $u_{i,t-1}u_{it}$ (weighted with $(\mu_{it}+1)/\mu_{i,t-1}$) and the transformation using u_{it}^2 (i.e. $(u_{it}^2-y_{it})+(1/\mu_{it})(u_{it}^2-y_{it}u_{it})$). Multiplying (2.3.8) for $s=t-1$ by $(\mu_{it}+1)/\mu_{i,t-1}$ gives

$$E[((\mu_{it}+1)/\mu_{i,t-1})u_{i,t-1}u_{it}\,|\,y_{i1},\eta_i,v_i^{t-1},x_i^T] = \phi_i^2\mu_{it}(\mu_{it}+1) + \phi_i v_{i,t-1}((\mu_{it}(\mu_{it}+1))/\mu_{i,t-1}) . \quad (2.3.19)$$

Applying the law of total expectation to (2.3.19) and (2.3.10),

$$E[((\mu_{it}+1)/\mu_{i,t-1})u_{i,t-1}u_{it}] = E[\phi_i^2\mu_{it}(\mu_{it}+1)] , \quad (2.3.20)$$

$$E[((u_{it}^2-y_{it})+(1/\mu_{it})(u_{it}^2-y_{it}u_{it}))] = E[\phi_i^2\mu_{it}(\mu_{it}+1)] . \quad (2.3.21)$$

Subtracting (2.3.20) from (2.3.21), the following $T-2$ cross-linkage moment conditions are obtained:

$$E[u_{it}(((\mu_{it}+1)/\mu_{it})u_{it}-((\mu_{it}+1)/\mu_{i,t-1})u_{i,t-1})+y_{it}(u_{it}-1)-y_{it}((\mu_{it}+1)/\mu_{it})u_{it}]=0 \quad ,$$

$$\text{for} \quad t=3,\ldots,T \quad . \qquad (2.3.22)$$

Eventually, a condensed full set of the moment conditions for the case where the assumption with respect to the overdispersion is imposed in addition to the implicit standard assumptions associated with strictly exogenous explanatory variables is composed of (2.3.11), (2.3.18), (2.3.22) and (2.3.12). That is, under the assumptions (2.3.1) with (2.3.6), the condensed full set is composed of the moment conditions (2.3.11), (2.3.18), (2.3.22) and (2.3.12). The moment conditions (2.3.11), (2.3.18) and (2.3.12) are linear with respect to γ, while (2.3.22) nonlinear.

2.4. Case of mean-stationary dependent variables

In this case, the stationarities of the dependent and explanatory variables are additionally assumed for the case of predetermined explanatory variables in the LFM (2.1.1) (see Kitazawa, 2007).

When

$$E[\exp(k\, x_{it})\,|\,\eta_i]=E[\varphi_i(k)\,|\,\eta_i] \quad , \qquad \text{for} \quad t=1,\ldots,T \qquad (2.4.1)$$

with k being any real number and

$$y_{i1}=(1/(1-\gamma))\phi_i\mu_{i1}+v_{i1} \qquad (2.4.2)$$

with

$$E[v_{i1}\,|\,\eta_i, x_{i1}]=0 \quad , \qquad (2.4.3)$$

the dependent variables in the LFM are mean-stationary:

$$E[y_{it}]=(1/(1-\gamma))E[\phi_i\varphi_i(\beta)] \quad , \qquad \text{for} \quad t=1,\ldots,T \quad . \qquad (2.4.4)$$

The assumption (2.4.1) implies that the explanatory variables x_{it} are stationary in the sense that their moment generating functions are equal over time. In this case, the observable analogue (2.2.7) is rewritten as

$$E[y_{i1}u_{it}\,|\,y_{i1},\eta_i,v_i^{t-1},x_i^t]=(1/(1-\gamma))\phi_i^2\mu_{i1}\mu_{it}+v_{i1}\phi_i\mu_{it} \quad . \qquad (2.4.5)$$

Using the observable analogues (2.4.5) (instead of (2.2.7)) and (2.2.8) with (2.4.1), the relationships between $y_{i1}u_{i3}$ and $u_{i2}u_{i3}$ and between $u_{i,t-2}u_{it}$ and $u_{i,t-1}u_{it}$ for $t=4,\ldots,T$ through the intermediary of the unconditional expectation operator after weighting them with appropriate transformations of explanatory variables x_{it} for $t=1,\ldots,T$ are realized by Kitazawa (2007) as the following $T-2$ stationarity moment conditions for the case without the assumption with respect to the overdispersion (2.2.6):

$$E[\Delta y_{i,t-1}(1/\mu_{it})u_{it}]=0 \quad , \qquad \text{for} \quad t=3,\ldots,T \quad , \qquad (2.4.6)$$

where Δ is the first-differencing operator.

In addition, the relationships between $x_{i,t-1}u_{it}$ and $x_{it}u_{it}$ for $t=2,\ldots,T$ are also realized by Kitazawa as the following $T-1$ stationarity moment conditions for the case without the assumption with respect to the overdispersion (2.2.6):

$$E[\Delta x_{it}(1/\mu_{it})u_{it}]=0 \quad , \qquad \text{for} \quad t=2,\ldots,T \quad . \tag{2.4.7}$$

From now on, the cross-linkage moment conditions for the case of mean-stationary dependent variables are constructed in the situation where the assumptions (2.4.1) and (2.4.2) with (2.4.3) are imposed in addition to the assumptions (2.2.2) – (2.2.6). They are solved by using the relationship between $y_{i1}u_{i2}$ and u_{i2}^2 and the relationships between $u_{i,t-1}u_{it}$ and u_{it}^2 for $t=3,\ldots,T$.

First, the relationship through the intermediary of the unconditional expectation operator is solved between $y_{i1}u_{i2}$ (weighted with $1/\mu_{i2}$) and the transformation using u_{i2}^2 (i.e. $(u_{i2}^2-y_{i2})+(1/\mu_{i2})(u_{i2}^2-y_{i2}u_{i2})$ weighted with $1/(\mu_{i2}+1)$). Multiplying (2.4.5) by $1/\mu_{it}$ gives

$$E[y_{i1}(1/\mu_{it})u_{it}\,|\,y_{i1},\eta_i,v_i^{t-1},x_i^t]=(1/(1-\gamma))\phi_i^2\mu_{i1}+v_{i1}\phi_i \quad . \tag{2.4.8}$$

In addition, multiplying (2.2.10) by $1/(\mu_{it}+1)$ gives

$$E[(\,(1/(\mu_{it}+1))(u_{it}^2-y_{it})+(1/(\mu_{it}+1))(1/\mu_{it})(u_{it}^2-y_{it}u_{it})\,)\,|\,y_{i1},\eta_i,v_i^{t-1},x_i^t]=\phi_i^2\mu_{it} \quad . \tag{2.4.9}$$

Applying the law of total expectation to (2.4.8) and (2.4.9) and allowing for (2.4.1), it follows that

$$E[y_{i1}(1/\mu_{it})u_{it}]=(1/(1-\gamma))E[\phi_i^2\varphi_i(\beta)] \quad , \tag{2.4.10}$$

$$E[(\,(1/(\mu_{it}+1))(u_{it}^2-y_{it})+(1/(\mu_{it}+1))(1/\mu_{it})(u_{it}^2-y_{it}u_{it})\,)]=E[\phi_i^2\varphi_i(\beta)] \quad . \tag{2.4.11}$$

Subtracting (2.4.10) for $t=2$ multiplied by $1-\gamma$ from (2.4.11) for $t=2$ gives

$$E[y_{i2}(1/(\mu_{i2}+1))(u_{i2}-1)\;-\;y_{i1}(1/\mu_{i2})u_{i2}]=0 \quad . \tag{2.4.12}$$

The detail of derivation of (2.4.12) is written in Appendix F.

Next, the relationship through the intermediary of the unconditional expectation operator is solved between $u_{i,t-1}u_{it}$ (weighted with $1/\mu_{it}$) and the transformation using u_{it}^2 (i.e. $(u_{it}^2-y_{it})+(1/\mu_{it})(u_{it}^2-y_{it}u_{it})$ weighted with $1/(\mu_{it}+1)$). Allowing for (2.4.1), equation (2.2.22) is written as

$$E[u_{i,t-1}(1/\mu_{it})u_{it}]=E[\phi_i^2\varphi_i(\beta)] \quad . \tag{2.4.13}$$

Subtracting (2.4.13) from (2.4.11) gives

$$E[y_{it}(1/(\mu_{it}+1))(u_{it}-1) \; + \; \Delta u_{it}(1/\mu_{it})u_{it} - y_{it}(1/\mu_{it})u_{it}]=0 \quad . \qquad (2.4.14)$$

The detail of derivation of (2.4.14) is written in Appendix G. In addition, creating the recursive equation

$$E[\Delta y_{it}(1/\mu_{it})u_{it}]=\gamma E[\Delta y_{i,t-1}(1/\mu_{it})u_{it}]+E[\Delta u_{it}(1/\mu_{it})u_{it}] \quad , \text{ for } \quad t=3,\ldots,T \quad . \qquad (2.4.15)$$

from the first-differences of (2.1.2) and applying the moment conditions (2.4.6), it can be seen that the following relationships hold:

$$E[\Delta u_{it}(1/\mu_{it})u_{it}]=E[\Delta y_{it}(1/\mu_{it})u_{it}] \quad , \qquad \text{for} \quad t=3,\ldots,T \quad . \qquad (2.4.16)$$

Accordingly, plugging (2.4.16) into (2.4.14) gives the following $T-2$ moment conditions:

$$E[y_{it}(1/(\mu_{it}+1))(u_{it}-1) \; - \; y_{i,t-1}(1/\mu_{it})u_{it}]=0 \quad , \qquad \text{for} \quad t=3,\ldots,T \quad , \qquad (2.4.17)$$

in which the order reduction with respect to γ is realized, compared to (2.4.14).

Writing (2.4.12) and (2.4.17) jointly, it follows that

$$E[y_{it}(1/(\mu_{it}+1))(u_{it}-1) \; - \; y_{i,t-1}(1/\mu_{it})u_{it}]=0 \quad , \qquad \text{for} \quad t=2,\ldots,T \quad , \qquad (2.4.18)$$

which are referred to as the stationarity moment conditions for the case of the overdispersion and whose number is $T-1$.

It should also be noted that for the case of the overdispersion formulated by (2.2.6) as well as for the case of the equidispersion, the intertemporal homoscedasticity moment conditions proposed by Ahn (1990) and Ahn and Schmidt (1995) in the context of the dynamic panel data model hold:

$$E[u_{it}^2-u_{i,t-1}^2]=0 \quad , \qquad \text{for} \quad t=3,\ldots,T \quad , \qquad (2.4.19)$$

when the dependent variables y_{it} are mean-stationary. That is, the moment conditions (2.4.19) are obtained by using the relationships (2.4.4) and the relationships

$$E[\phi_i y_{it}]=(1/(1-\gamma))E[\phi_i^2 \varphi_i(\beta)] \quad , \qquad \text{for} \quad t=1,\ldots,T \quad , \qquad (2.4.20)$$

both of which hold when the assumptions (2.4.1) and (2.4.2) with (2.4.3) are additionally imposed. The implication of (2.4.19) is that the disturbances v_{it} are homoscedastic over time (see Kitazawa, 2007).

Eventually, a condensed full set of the moment conditions for the case of stationary dependent variables when the assumption with respect to the overdispersion is imposed in addition to the implicit standard assumptions associated with predetermined explanatory variables is composed of (2.2.11), (2.2.19), (2.4.18), (2.2.12) and (2.4.7). That is, under the assumptions (2.2.1) with (2.2.6), (2.4.1) and (2.4.2) with (2.4.3), the condensed full set is composed of the moment conditions (2.2.11), (2.2.19), (2.4.18), (2.2.12) and (2.4.7), all of which are linear with respect to γ .

2.5. Discussion

There can be a case where a manipulation is needed, when using any of the moment conditions (2.2.24) and (2.4.18) for the estimation of γ and β. If all values in x_{it} are positive, the estimates of β using these moment conditions seem to be in danger of going to infinity. In this case, in order that x_{it} contains both positive and negative values evenly, x_{it} needs to be transformed in deviation from an appropriate value b. That is, $\tilde{x}_{it}$ needs to be used in the estimations instead of x_{it}, where $\tilde{x}_{it}=x_{it}-b$. The selection of b by Windmeijer (2000) is the overall mean of x_{it} (i.e. $b=(1/(N\ T))\sum_{i=1}^{N}\sum_{t=1}^{T} x_{it}$).

2.6. GMM estimators

Any set of the moment conditions for the LFM (2.1.1) can be collectively written in the following $m\times 1$ vector form:

$$E[g_i(\theta)]=0 \ , \tag{2.6.1}$$

where m is number of moment conditions, $\theta=[\gamma\ \ \beta]'$, $g_i(\theta)$ (which is the function of θ) is composed of the observable variables and θ for the individual i. Using the following empirical counterpart for (2.6.1):

$$\bar{g}(\theta)\ =\ (1/N)\sum_{i=1}^{N} g_i(\theta) \ , \tag{2.6.2}$$

the GMM estimator $\hat{\theta}$ is constructed by minimizing the following criterion function with respect to θ:

$$\bar{g}(\theta)'\ W_N(\hat{\theta}_1)\ \bar{g}(\theta) \ , \tag{2.6.3}$$

where the $m\times m$ optimal weighting matrix is given as follows by using a initial consistent estimator of θ (i.e. $\hat{\theta}_1$):

$$W_N(\hat{\theta}_1)=\left((1/N)\sum_{i=1}^{N} g_i(\hat{\theta}_1)g_i(\hat{\theta}_1)'\right)^{-1} \ . \tag{2.6.4}$$

The efficient asymptotic variance of $\hat{\theta}$ is estimated by using

$$\hat{V}(\hat{\theta})=(1/N)\left(D(\hat{\theta})'\ W_N(\hat{\theta}_1)\ D(\hat{\theta})\right)^{-1} \ , \tag{2.6.5}$$

where $D(\hat{\theta})=\partial\bar{g}(\theta)/\partial\theta'|_{\theta=\hat{\theta}}$.[7] The GMM estimations for the LFM are explained in detail in Windmeijer (2002, 2008).

Some GMM estimators are constructed, associated with the overdispersion. Firstly, for the case of predetermined explanatory variables, the GMM estimator using the moment conditions (2.2.11),

7 It is conceivable that the usage of the finite sample corrected variance proposed by Windmeijer (2005, 2008) would be preferable in small sample.

(2.2.12) and (2.2.19) is referred to as the GMM(qdcn) estimator, while that using (2.2.11), (2.2.12), (2.2.19) and (2.2.24) is referred to as the GMM(prcn) estimator. Secondly, for the case of strictly exogenous explanatory variables, the GMM estimator using the moment conditions (2.3.11), (2.3.12) and (2.3.18) is referred to as the GMM(qecn) estimator, while that using (2.3.11), (2.3.12), (2.3.18) and (2.3.22) is referred to as the GMM(excn) estimator. Thirdly, for the case of mean-stationary dependent variables, the GMM estimator using the moment conditions (2.2.11), (2.2.12), (2.2.19), (2.4.7) and (2.4.18) is referred to as the GMM(sacn) estimator.

Kitazawa (2009) proposes some GMM estimators associated with the equidispersion. In this paper, the GMM(qdc), GMM(prc), GMM(qec), GMM(exc) and GMM(sac) estimators proposed by Kitazawa (2009) are referred to as the GMM(qdcp), GMM(prcp), GMM(qecp), GMM(excp) and GMM(sacp) estimators, respectively. The GMM estimators corresponding to the GMM(qdcp), GMM(prcp), GMM(qecp), GMM(excp) and GMM(sacp) estimators are the GMM(qdcn), GMM(prcn), GMM(qecn), GMM(excn) and GMM(sacn) estimators in the case of the overdisperion.

In addition, there are some distribution-free GMM estimators: the GMM(qd), GMM(pr), GMM(qe), GMM(ex) and GMM(sa) (see Kitazawa, 2007 and 2009).

It should be noted that the transformation described in previous subsection is needed to implement the GMM(pr), GMM(prcp), GMM(prcn), GMM(sa), GMM(sacp) and GMM(sacn) estimators.

3. Monte Carlo

In this section, some small sample performances of a portion of the GMM estimators exhibited in previous section are investigated with Monte Carlo experiments. The GMM estimators to be looked into in this paper are the GMM(qd). GMM(qdcp), GMM(qdcn) and GMM(pr) estimators (which are tailored to the specification of predetermined explanatory variables), the GMM(qe). GMM(qecp), GMM(qecn) and GMM(ex) estimators (which are tailored to the specification of strictly exogenous explanatory variables) and the GMM(sa). GMM(sacp) and GMM(sacn) estimators (which are tailored to the specification of mean-stationarity dependent variables). The experiments are implemented by using an econometric software TSP version 4.5.[8]

3.1. Data generating process

Two types of data generating process (DGP) are configured: in one type, the dependent variables are generated from the Poisson distribution, while in another type, they are generated from the negative binomial distribution with the functional form based on the model in section 2.

The one type of DGP is as follows:

$$y_{it} \sim Poisson(\gamma y_{i,t-1} + \exp(\beta x_{it} + \eta_i)) \quad , \tag{3.1.1}$$

$$y_{i,-TG+1} \sim Poisson((1/(1-\gamma))\exp(\beta x_{i,-TG+1} + \eta_i)) \quad , \tag{3.1.2}$$

$$x_{it} = \rho x_{i,t-1} + \tau \eta_i + \varepsilon_{it} \quad , \tag{3.1.3}$$

$$x_{i,-TG+1} = (1/(1-\rho))\tau \eta_i + (1/(1-\rho^2)^{(1/2)})\varepsilon_{i,-TG+1} \quad , \tag{3.1.4}$$

$$\eta_i \sim N(0, \sigma_\eta^2) \quad ; \quad \varepsilon_{it} \sim N(0, \sigma_\varepsilon^2) \quad ,$$

8 See Hall and Cummins (2006) as for the details of the software.

where $t=-TG+1,\dots,-1,0,1,\dots,T$ with TG being the number of pre-sample periods to be generated. In the DGP, values are set to the parameters γ , β , ρ , τ , σ_η^2 and σ_ε^2 . The experiments are carried out with $TG=50$, the cross-sectional sizes $N=100$, 500 and 1000 , the numbers of periods used for the estimations $T=4$ and 8 and the number of replications $NR=1000$. This DGP setting is the same as that of Blundell et al. (2002), except for the initial condition of y_{it} . That is, the initial condition (3.1.2) denotes that the initial conditions of dependent variables are stationary. The DGP is configured with the explanatory variables x_{it} being strictly exogenous.

In the another type of DGP, equations (3.1.1) and (3.1.2) are replaced by the following ones respectively:

$$y_{it} \sim Negbin(\ \gamma y_{i,t-1}/\exp(\eta_i)+\exp(\beta x_{it})\ ,\ \exp(\eta_i)\ )\ , \tag{3.1.5}$$

$$y_{i,-TG+1} \sim Negbin(\ (1/(1-\gamma))\exp(\beta x_{i,-TG+1})\ ,\ \exp(\eta_i)\ )\ , \tag{3.1.6}$$

where the denotation $X \sim Negbin(\alpha,\theta)$ implies that the (non-negative integer-valued) count variable X is distributed as the negative binomial distribution whose probability function is

$$p(X)\ =\ (\Gamma(\alpha+X)/(\Gamma(\alpha)\Gamma(X+1)))(1/(1+\theta))^{\alpha}(\theta/(1+\theta))^{X}\ , \tag{3.1.7}$$

where $\Gamma(\cdot)$ is the gamma function written as $\Gamma(s)=\int_0^\infty z^{s-1}\exp(-z)\,dz$ for $s>0$ and α and θ are the parameters with $\alpha\geq 0$ and $\theta\geq 0$ respectively.

The DGP composed of (3.1.1), (3.1.2), (3.1.3) and (3.1.4) is referred to as the “DGP-Poisson”, while the DGP composed of (3.1.5), (3.1.6), (3.1.3) and (3.1.4) is referred to as the “DGP-Negbin”. In the DGP-Poisson, the GMM estimators incorporating the moment conditions associated with the equidispersion are consistent for large N and small T , while those incorporating the moment conditions associated with the negative binomial model are inconsistent. Antithetically, in the DGP-Negbin, the GMM estimators incorporating the moment conditions associated with the equidispersion are inconsistent, while those incorporating the moment conditions associated with the negative binomial model are consistent. Accordingly, it is expected that for large N , the considerable degrees of endemic bias and rmse are found in the DGP-Poisson for the GMM estimators incorporating the moment conditions associated with the negative binomial model, while the considerable degrees of endemic bias and rmse are found in the DGP-Negbin for the GMM estimators incorporating the moment conditions associated with the equidispersion.

3.2. Estimators for comparison

The following three estimators are used for comparison: the Level estimator, the within group (WG) mean scaling estimator and the pre-sample mean (PSM) estimator.[9] The Level and WG estimators are inconsistent in the DGP settings above, where N and T are able to be regarded as being large and small respectively. On the contrary, the PSM estimator is consistent if the long history is used in constructing the pre-sample means of dependent variables. The details on these estimators are described in Blundell et al. (1999, 2002) and Kitazawa (2007).

9 The WG estimator proposed by Blundell et al. (2002) is consistent for the case of allowing for the strictly exogenous explanatory variables and no dynamics. In addition, Lancaster (2002) and Blundell et al. (2002) point out that the Poisson maximum likelihood estimator is the same as the Poisson conditional maximum likelihood estimator and furthermore Blundell et al. (2002) show that they are identical to the WG estimator. The WG estimator requires no distributional assumption.

3.3. Monte Carlo results

For $T=4$, Monte Carlo results are shown in Table 1 for the Poisson model (i.e. the DGP-Poisson) and in Table 2 for the negative binomial model (i.e. the DGP-Negbin), while for $T=8$, they are shown in Table 3 for the Poisson model and in Table 4 for the negative binomial model.

In all tables, the endemic upward and downward biases are found for the Level and WG estimators respectively, while the PSM estimator behaves better as the longer pre-sample history is used.

The instruments used for the GMM estimators are curtailed so that the past dependent variables (y_{it}) dated $t-3$ and before are not used for the quasi-differenced equation dated t and further for the GMM(qd), GMM(qdcp), GMM(qdcn), GMM(pr), GMM(sa), GMM(sacp) and GMM(sacn) estimators the past explanatory variables (x_{it}) dated $t-3$ and before are not used for the quasi-differenced equation dated t.

The results on the GMM estimators say that if the cross-linkage moment conditions are valid for each specification of count dependent variables, the small sample properties could be considered to be improved by using the cross-linkage moment conditions, while if not so, the bias and rmse are of the considerable magnitude.

Firstly, it can be said that the GMM estimators using the cross-linkage moment conditions valid for each specification of count dependent variables outperform the conventional GMM(qd) estimator.

Secondly, looking at Tables 1 and 3 where the DGP-Poisson is used, the small sample properties of the GMM estimators incorporating the cross-linkage moment conditions associated with the equidispersion improve as the cross-sectional size (N) increases from 100, 500 to 1000, while those of the GMM estimators incorporating the cross-linkage moment conditions associated with the overdispersion remain to be poor. Conversely, looking at Tables 2 and 4 where the DGP-Negbin is used, the small sample properties of the GMM estimators incorporating the cross-linkage moment conditions associated with the overdispersion improve as the cross-sectional size (N) increases from 100, 500 to 1000, while those of the GMM estimators incorporating the cross-linkage moment conditions associated with the equidispersion remain to be poor. Further, it can be recognized that for the case of using the moment conditions invalid for each DGP, the augmentations of the Monte Carlo means of the Sargan statistic emerge as the cross-sectional size increases, which are said to be the reflections of the inconsistency.

Thirdly, comparing the results for the GMM(qd) estimator with those for the GMM(qdcp) estimator and comparing the results for the GMM(qe) estimator with those for the GMM(qecp) estimator in Tables 1 and 3 where the DGP is of the Poisson model, it can be said that some gains and no loss seem to be obtained in small sample by using the cross-linkage moment conditions associated with the equidispersion. Likewise, comparing the results for the GMM(qd) estimator with those for the GMM(qdcn) estimator and comparing the results for the GMM(qe) estimator with those for the GMM(qecn) estimator in Tables 2 and 4 where the DGP is of the negative binomial model, it can be said that some gains and no loss seem to be obtained in small sample by using the cross-linkage moment conditions associated with the overdispersion. It is shown in the limited Monte Carlo experiments that the additional usage of the cross-linkage moment conditions improve or do not at least vitiate the small sample performances as long as the cross-linkage moment conditions are valid.

Fourth, comparing the results for the GMM(sa) estimator with those for the GMM(sacp) estimator in Tables 1 and 3 where the DGP is of the Poisson model, it can be said that the GMM estimators utilizing the condensed full set incorporating the cross-linkage moment conditions associated with the equidispersion maximally does not underperform those without incorporating the cross-linkage moment conditions. Likewise, comparing the results for the GMM(sa) estimator with those for the GMM(sacn) estimator in Tables 2 and 4 where the DGP is of the negative binomial model, it can be said that the GMM estimators utilizing the condensed full set incorporating the cross-linkage moment conditions associated with the overdispersion maximally does not underperform those without incorporating the cross-linkage moment conditions.

4. Conclusion

In this paper, the cross-linkage moment conditions associated with an overdispersion for count panel data model were proposed for the case of predetermined explanatory variables, for the case of strictly exogenous explanatory variables and for the case of mean-stationary dependent variables. The specification of the model is an extension of the fixed effects negative binomial model proposed by Hausman et al. (1984) to the linear feedback model (LFM) proposed by Blundell et al. (2002) and the consistent (GMM) estimations with the overdispersion taken into consideration became possible for the negative binomial count panel data model. It was corroborated from some Monte Carlo experiments that for the negative binomial model, the cross-linkage moment conditions associated with the overdispersion are valid and the usage of the cross-linkage moment conditions associated with the overdispersion ameliorate or do not at least deteriorate the small sample performances.

Appendix A.

The conditional probability of the dependent variables in the negative binomial model specified by Hausman et al. (1984) is written as follows, incorporating the dynamics and allowing for the predetermined explanatory variables:

$$p(y_{it} \mid y_{i1}, \eta_i, v_i^{t-1}, x_i^t) = (\Gamma(\lambda_{it}+y_{it})/(\Gamma(\lambda_{it})\Gamma(y_{it}+1)))(1/(1+\phi_i))^{\lambda_{it}}(\phi_i/(1+\phi_i))^{y_{it}} \quad , \tag{A.1}$$

where $\Gamma(\cdot)$ is the gamma function written as $\Gamma(s)=\int_0^\infty z^{s-1}\exp(-z)\,dz$ for $s>0$ and $\lambda_{it}=E[y_{it}(1/\phi_i) \mid y_{i1}, \eta_i, v_i^{t-1}, x_i^t]$. In this case, the conditional variance is written as

$$Var[y_{it} \mid y_{i1}, \eta_i, v_i^{t-1}, x_i^t] = (1+\phi_i)E[y_{it} \mid y_{i1}, \eta_i, v_i^{t-1}, x_i^t] \quad . \tag{A.2}$$

Allowing for the system (2.1.1) with (2.2.1),

$$E[y_{it} \mid y_{i1}, \eta_i, v_i^{t-1}, x_i^t] = \gamma y_{i,t-1} + \exp(\eta_i + \beta x_{it}) \tag{A.3}$$

and accordingly

$$Var[y_{it} \mid y_{i1}, \eta_i, v_i^{t-1}, x_i^t] = Var[v_{it} \mid y_{i1}, \eta_i, v_i^{t-1}, x_i^t] \quad . \tag{A.4}$$

Plugging (A.4) into (A.2) gives (2.2.6).

Appendix B.

Firstly, the following relationship holds:

$$E[v_{it}^2 \mid y_{i1}, \eta_i, v_i^{t-1}, x_i^t] = E[u_{it}^2 \mid y_{i1}, \eta_i, v_i^{t-1}, x_i^t] - \phi_i^2 \mu_{it}^2 \ , \tag{B.1}$$

where the relationship $E[\phi_i \mu_{it} v_{it} \mid y_{i1}, \eta_i, v_i^{t-1}, x_i^t]=0$ originating from (2.2.1) is used.

Next, the following relationship is obtained:

$$E[\phi_i y_{it} \mid y_{i1}, \eta_i, v_i^{t-1}, x_i^t] = \gamma \phi_i y_{i,t-1} + \phi_i^2 \mu_{it} \ , \tag{B.2}$$

by utilizing the relationship $E[\phi_i v_{it} \mid y_{i1}, \eta_i, v_i^{t-1}, x_i^t]=0$ originating from (2.2.1). Further, the following relationship is obtained:

$$E[(y_{it} - u_{it}) u_{it} (1/\mu_{it}) \mid y_{i1}, \eta_i, v_i^{t-1}, x_i^t] = \gamma y_{i,t-1} \phi_i \ , \tag{B.3}$$

by utilizing the relationship $E[y_{i,t-1}(1/\mu_{it}) v_{it} \mid y_{i1}, \eta_i, v_i^{t-1}, x_i^t] = 0$ originating from (2.2.1). Employing (B.3) reduces (B.2) to

$$E[\phi_i y_{it} \mid y_{i1}, \eta_i, v_i^{t-1}, x_i^t] = E[(y_{it} - u_{it}) u_{it} (1/\mu_{it}) \mid y_{i1}, \eta_i, v_i^{t-1}, x_i^t] + \phi_i^2 \mu_{it} \ . \tag{B.4}$$

Plugging (B.1) and (B.4) into (2.2.6) gives (2.2.10).

Appendix C.

Firstly, creating the recursive equation

$$\begin{aligned} &E[y_{i,t-1}((\mu_{i,t-1}/\mu_{it}) u_{it} - u_{i,t-1})] \\ &= \gamma E[y_{i,t-2}((\mu_{i,t-1}/\mu_{it}) u_{it} - u_{i,t-1})] + E[u_{i,t-1}((\mu_{i,t-1}/\mu_{it}) u_{it} - u_{i,t-1})] \ , \end{aligned} \tag{C.1}$$

from (2.1.2) dated $t-1$ and applying the moment conditions (2.2.11) for $s=t-2$, it can be seen that equation (2.2.17) holds.

Next, creating the recursive equation

$$\begin{aligned} &E[y_{i,t-1}((1/\mu_{it}) u_{it} - (1/\mu_{i,t-1}) u_{i,t-1})] \\ &= \gamma E[y_{i,t-2}((1/\mu_{it}) u_{it} - (1/\mu_{i,t-1}) u_{i,t-1})] + E[u_{i,t-1}((1/\mu_{it}) u_{it} - (1/\mu_{i,t-1}) u_{i,t-1})] \ , \end{aligned} \tag{C.2}$$

from (2.1.2) dated $t-1$ and applying one of the moment conditions based on the Wooldridge transformation

$$E[y_{i,t-2}((1/\mu_{it}) u_{it} - (1/\mu_{i,t-1}) u_{i,t-1})]=0 \ , \qquad t=3,\ldots,T \ , \tag{C.3}$$

it can be seen that equation (2.2.18) holds.

It is corroborated that the moment conditions (C.3) are obtained by the implicit operation.

Multiplying both sides of the observable analogues (2.2.7) and (2.2.8) by $1/\mu_{it}$ gives

$$E[y_{i1}(1/\mu_{it})u_{it} \mid y_{i1},\eta_i,v_i^{t-1},x_i^t]=y_{i1}\phi_i \quad , \tag{C.4}$$

$$E[u_{is}(1/\mu_{it})u_{it} \mid y_{i1},\eta_i,v_i^{t-1},x_i^t]=\phi_i^2\mu_{is}+v_{is}\phi_i \quad , \qquad \text{for} \quad 2\le s\le t-1 \quad . \tag{C.5}$$

Applying the law of total expectation to (C.4) and (C.5) gives

$$E[y_{i1}(1/\mu_{it})u_{it}]=E[y_{i1}\phi_i] \quad , \tag{C.6}$$

$$E[u_{is}(1/\mu_{it})u_{it}]=E[\phi_i^2\mu_{is}] \quad , \qquad \text{for} \quad 2\le s\le t-1 \quad . \tag{C.7}$$

where (2.2.5) for $t=s$ is used to obtain (C.7). Taking the first-differences of (C.6) and (C.7) with respect to t gives

$$E[y_{i1}((1/\mu_{it})u_{it}-(1/\mu_{i,t-1})u_{i,t-1})]=0 \quad , \tag{C.8}$$

$$E[u_{is}((1/\mu_{it})u_{it}-(1/\mu_{i,t-1})u_{i,t-1})]=0 \quad . \qquad \text{for} \quad 2\le s\le t-2 \quad . \tag{C.9}$$

Creating the recursive equation

$$\begin{aligned} &E[y_{is}((1/\mu_{it})u_{it}-(1/\mu_{i,t-1})u_{i,t-1})] \\ &= \gamma E[y_{i,s-1}((1/\mu_{it})u_{it}-(1/\mu_{i,t-1})u_{i,t-1})]+E[u_{is}((1/\mu_{it})u_{it}-(1/\mu_{i,t-1})u_{i,t-1})] \ , \end{aligned}$$

$$\text{for} \quad 2\le s\le t-2 \quad , \tag{C.10}$$

from (2.1.2) for $t=s$ and applying the initial condition (C.8) and the innovation (C.9) to (C.10) successively generates the following moment conditions based on the Wooldridge transformation:

$$E[y_{is}((1/\mu_{it})u_{it}-(1/\mu_{i,t-1})u_{i,t-1})]=0 \quad , \text{for} \quad s=1,\ldots,t-2 \quad ; \quad t=3,\ldots,T \quad .$$

(C.11)

The moment conditions utilizing the Wooldridge transformation are explained in Wooldridge (1997) and Windmeijer (2000, 2008). The moment conditions (C.3) are the moment conditions (C.11) for $s=t-2$.

Appendix D.

Subtracting (2.2.22) from (2.2.23), it follows that

$$\begin{aligned} &E[(1/\mu_{it})((\mu_{i,t-1}/(\mu_{it}+1))u_{it}-u_{i,t-1})u_{it}] + E[(1/\mu_{it})((\mu_{i,t-1}/(\mu_{it}+1))(1/\mu_{it}))u_{it}^2] \\ &- E[y_{it}(1/\mu_{it})((\mu_{i,t-1}/(\mu_{it}+1))(1/\mu_{it}))u_{it}] - E[y_{it}(1/\mu_{it})(\mu_{i,t-1}/(\mu_{it}+1))] = 0 \quad . \end{aligned}$$

(D.1)

Using the relationship $(\mu_{i,t-1}/(\mu_{it}+1))(1+(1/\mu_{it}))=\mu_{i,t-1}/\mu_{it}$ and an arrangement, (D.1) is written as

$$\begin{aligned}&E[(1/\mu_{it})((\mu_{i,t-1}/\mu_{it})u_{it}-u_{i,t-1})u_{it}]\\ &-E[y_{it}(1/\mu_{it})((\mu_{i,t-1}/(\mu_{it}+1))(1/\mu_{it}))u_{it}]-E[y_{it}(1/\mu_{it})((\mu_{i,t-1}/(\mu_{it}+1)))u_{it}]\\ &+E[y_{it}(1/\mu_{it})((\mu_{i,t-1}/(\mu_{it}+1)))u_{it}]-E[y_{it}(1/\mu_{it})(\mu_{i,t-1}/(\mu_{it}+1))]=0\end{aligned} \quad (D.2)$$

Further, using the relationship $(\mu_{i,t-1}/(\mu_{it}+1))(1+(1/\mu_{it}))=\mu_{i,t-1}/\mu_{it}$, (D.2) is written as (2.2.24).

Appendix E.
Creating the recursive equations

$$\begin{aligned}&E[y_{i,t-1}(u_{it}-(\mu_{it}/\mu_{i,t-1})u_{i,t-1})]\\ &=\gamma E[y_{i,t-2}(u_{it}-(\mu_{it}/\mu_{i,t-1})u_{i,t-1})]+E[u_{i,t-1}(u_{it}-(\mu_{it}/\mu_{i,t-1})u_{i,t-1})],\end{aligned} \quad (E.1)$$

from (2.1.2) dated $t-1$ and applying the moment conditions (2.3.11) for $s=t-2$ and the relationship $\mu_{it}/\mu_{i,t-1}=(\mu_{it}/(\mu_{i,t-1}+1))(1+(1/\mu_{i,t-1}))$ to (E.1), it can be seen that equation (2.3.17) holds.

Appendix F.
It should be noted that (2.4.10) can be rewritten as

$$E[(1/(\mu_{it}+1))(1+(1/\mu_{it}))y_{i1}u_{it}]=(1/(1-\gamma))E[\phi_i^2\varphi(\beta)], \quad (F.1)$$

since $1/\mu_{it}=(1/(\mu_{it}+1))(1+(1/\mu_{it}))$.

Subtracting (F.1) for $t=2$ multiplied by $1-\gamma$ from (2.4.11) for $t=2$ gives

$$\begin{aligned}&E[(1/(\mu_{i2}+1))u_{i2}(u_{i2}-(1-\gamma)y_{i1})]-E[(1/(\mu_{i2}+1))y_{i2}]\\ &+E[(1/(\mu_{i2}+1))(1/\mu_{i2})u_{i2}(u_{i2}-(1-\gamma)y_{i1})]-E[(1/(\mu_{i2}+1))(1/\mu_{i2})y_{i2}u_{i2}]=0.\end{aligned} \quad (F.2)$$

Using the relationship $u_{it}-(1-\gamma)y_{i,t-1}=\Delta y_{it}$ stemming from (2.1.2) and an arrangement, (F.2) is rewritten as

$$\begin{aligned}&E[(1/(\mu_{i2}+1))u_{i2}\Delta y_{i2}]-E[(1/(\mu_{i2}+1))y_{i2}]+E[(1/(\mu_{i2}+1))(1/\mu_{i2})u_{i2}\Delta y_{i2}]\\ &+E[(1/(\mu_{i2}+1))y_{i2}u_{i2}]-E[(1/(\mu_{i2}+1))y_{i2}u_{i2}]-E[(1/(\mu_{i2}+1))(1/\mu_{i2})y_{i2}u_{i2}]=0.\end{aligned} \quad (F.3)$$

Applying the relationship $(1/(\mu_{it}+1))(1+(1/\mu_{it}))=1/\mu_{it}$ to (F.3), (F.3) reduces to (2.4.12).

Appendix G.
It should be noted that (2.4.13) can be rewritten as

$$E[(1/(\mu_{it}+1))(1+(1/\mu_{it}))u_{i,t-1}u_{it}]=E[\phi_i^2\varphi(\beta)] \ , \qquad (G.1)$$

since $1/\mu_{it}=(1/(\mu_{it}+1))(1+(1/\mu_{it}))$.

Subtracting (G.1) from (2.4.11) and employing an arrangement give

$$\begin{aligned} &E[(1/(\mu_{it}+1))u_{it}\Delta u_{it}]-E[(1/(\mu_{it}+1))y_{it}]+E[(1/(\mu_{it}+1))(1/\mu_{it})u_{it}\Delta u_{it}] \\ &+E[(1/(\mu_{it}+1))y_{it}u_{it}]-E[(1/(\mu_{it}+1))y_{it}u_{it}]-E[(1/(\mu_{it}+1))(1/\mu_{it})y_{it}u_{it}] = 0 \ . \end{aligned}$$

(G.2)

Applying the relationship $(1/(\mu_{it}+1))(1+(1/\mu_{it}))=1/\mu_{it}$ to (G.2), (G.2) reduces to (2.4.14).

References

Ahn, S.C. (1990). Three essays on share contracts, labor supply, and the estimation of models for dynamic panel data. *Unpublished Ph.D. Dissertation (Michigan State University, East Lansing, MI)*.

Ahn, S.C. & Schmidt, P. (1995). Efficient estimation of models for dynamic panel data. *Journal of Econometrics* **68**, 5-27.

Blundell, R., Griffith, R. & Van Reenen, J. (1999). Market share, market value and innovation in a panel of British manufacturing firms. *Review of Economic Studies* **66**, 529-554.

Blundell, R., Griffith, R. & Windmeijer, F. (2002). Individual effects and dynamics in count data models. *Journal of Econometrics* **108**, 113-131.

Cameron, A.C. & Trivedi, P.K. (2005). *Microeconometrics. Method and applications,* Cambridge.

Chamberlain, G. (1992). Comment: sequential moment restrictions in panel data. *Journal of Business and Economic Statistics* **10**, 20-26.

Crépon, B. & Duguet, E. (1997). Estimating the innovation function from patent numbers: GMM on count panel data. *Journal of Applied Econometrics* **12**, 243-263.

Hall, B.H. & Cummins, C. (2006). *TSP 5.0 user's guide*, TSP international.

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica* **50**, 1029-1054.

Hausman, J. A., Hall, B. H. & Griliches, Z. (1984). Econometric models for count data with an application to the patent-R&D relationship. *Econometrica* **52**, 909-938.

Kitazawa, Y. (2007). Some additional moment conditions for a dynamic count panel data model. *Kyushu Sangyo University, Faculty of Economics, Discussion Paper*, No. **29**.

Kitazawa, Y. (2009). Equidispersion and moment conditions for count panel data model. *Kyushu Sangyo University, Faculty of Economics, Discussion Paper*, No. **33**.

Lancaster, T. (2002). Orthogonal parameters and panel data. *Review of Economic Studies* **69**, 647-666.

Windmeijer, F. (2000). Moment conditions for fixed effects count data models with endogenous regressors. *Economics Letters* **68**, 21-24.

Windmeijer, F. (2002). ExpEnd, A Gauss programme for non-linear GMM estimation of ***Exp***onential models with ***End***ogenous regressors for cross section and panel data. *The Institute for Fiscal Studies, Department of Economics, UCL, Cemmap Working Paper*, **CWP 14/02**.

Windmeijer, F. (2005). A finite sample correction for the variance of linear efficient two-step GMM estimators. *Journal of Econometrics* **126**, 25-51.

Windmeijer, F. (2008). GMM for panel count data models. *Laszlo, M. and P. Sevestre (eds.), The econometrics of panel data. Fundamentals and recent developments in theory*, Springer.

Winkelmann, R. (2008). *Econometric analysis of count data, 5th edition,* Springer.

Wooldridge, J. M. (1997). Multiplicative panel data models without the strict exogeneity assumption. *Econometric Theory* **13**, 667-678.

Table 1 **Monte Carlo results for LFM, T=4** (Poisson model)

$\gamma=0.5$; $\beta=0.5$; $\rho=0.5$; $\tau=0.1$; $\sigma_{\eta}^{2}=0.5$; $\sigma_{\varepsilon}^{2}=0.5$

		N=100		N=500		N=1000	
		bias	rmse	bias	rmse	bias	rmse
		Sargan	df	Sargan	df	Sargan	df
Level	γ	0.256	0.264	0.273	0.275	0.278	0.279
	β	0.545	0.656	0.549	0.571	0.557	0.573
WG	γ	**−0.452**	**0.463**	**−0.446**	**0.449**	**−0.446**	**0.447**
	β	**−0.260**	**0.272**	**−0.261**	**0.263**	**−0.263**	**0.264**
GMM(qd)	γ	−0.274	0.398	−0.104	0.161	−0.061	0.112
	β	−0.259	0.371	−0.124	0.219	−0.078	0.172
		4.42	4	4.58	4	4.58	4
GMM(qdcp)	γ	**−0.054**	**0.155**	**−0.006**	**0.066**	**−0.001**	**0.045**
	β	**−0.134**	**0.288**	**−0.028**	**0.148**	**−0.013**	**0.104**
		8.72	**6**	**7.91**	**6**	**7.56**	**6**
GMM(qdcn)	γ	0.262	0.273	0.234	0.237	0.227	0.228
	β	0.082	0.379	0.420	0.474	0.542	0.568
		8.36	6	11.30	6	12.34	6
GMM(pr)	γ	**−0.090**	**0.214**	**−0.037**	**0.092**	**−0.016**	**0.067**
	β	**−0.159**	**0.282**	**−0.059**	**0.157**	**−0.026**	**0.134**
		5.62	**5**	**5.66**	**5**	**5.66**	**5**
GMM(qe)	γ	−0.222	0.306	−0.060	0.110	−0.031	0.075
	β	−0.148	0.220	−0.049	0.111	−0.025	0.086
		9.03	8	8.90	8	8.52	8
GMM(qecp)	γ	**−0.074**	**0.156**	**−0.004**	**0.060**	**0.000**	**0.042**
	β	**−0.083**	**0.185**	**−0.008**	**0.088**	**−0.001**	**0.067**
		13.88	**10**	**12.13**	**10**	**11.27**	**10**
GMM(qecn)	γ	0.233	0.247	0.255	0.257	0.260	0.261
	β	0.114	0.316	0.284	0.322	0.323	0.344
		14.18	10	20.35	10	26.97	10
GMM(ex)	γ	**−0.103**	**0.226**	**−0.035**	**0.092**	**−0.015**	**0.058**
	β	**−0.106**	**0.214**	**−0.032**	**0.106**	**−0.013**	**0.077**
		9.97	**9**	**9.79**	**9**	**9.35**	**9**
GMM(sa)	γ	−0.023	0.139	−0.019	0.079	−0.010	0.059
	β	−0.053	0.212	−0.023	0.137	−0.012	0.104
		9.70	9	9.46	9	9.46	9
GMM(sacp)	γ	**0.038**	**0.115**	**0.016**	**0.060**	**0.011**	**0.043**
	β	**−0.022**	**0.215**	**0.016**	**0.133**	**0.012**	**0.099**
		14.06	**12**	**12.87**	**12**	**12.91**	**12**
GMM(sacn)	γ	0.277	0.283	0.249	0.250	0.242	0.242
	β	0.036	0.263	0.350	0.391	0.471	0.490
		18.21	12	30.10	12	38.17	12
PSM	$\gamma(4)$	**0.132**	**0.156**	**0.157**	**0.162**	**0.163**	**0.167**
	$\beta(4)$	**0.191**	**0.296**	**0.205**	**0.225**	**0.211**	**0.229**
	$\gamma(8)$	0.104	0.132	0.125	0.131	0.130	0.135
	$\beta(8)$	0.141	0.228	0.148	0.165	0.152	0.165
	$\gamma(25)$	**0.046**	**0.091**	**0.061**	**0.072**	**0.066**	**0.073**
	$\beta(25)$	**0.058**	**0.139**	**0.062**	**0.083**	**0.065**	**0.078**
	$\gamma(50)$	0.020	0.081	0.033	0.050	0.038	0.048
	$\beta(50)$	0.031	0.119	0.032	0.059	0.035	0.052

See Notes of Tables.

Table 2 **Monte Carlo results for LFM, T=4** (Negative binomial model)

$\gamma=0.5$; $\beta=0.5$; $\rho=0.5$; $\tau=0.1$; $\sigma_\eta^2=0.5$; $\sigma_\varepsilon^2=0.5$

		N=100		N=500		N=1000	
		bias	rmse	bias	rmse	bias	rmse
		Sargan	df	Sargan	df	Sargan	df
Level	γ	0.133	0.165	0.154	0.162	0.158	0.163
	β	0.302	0.419	0.300	0.357	0.303	0.325
WG	γ	**−0.513**	**0.533**	**−0.500**	**0.506**	**−0.500**	**0.503**
	β	**−0.264**	**0.294**	**−0.275**	**0.280**	**−0.276**	**0.279**
GMM(qd)	γ	−0.350	0.521	−0.156	0.249	−0.098	0.165
	β	−0.327	0.496	−0.205	0.295	−0.132	0.242
		4.27	4	4.45	4	4.36	4
GMM(qdcp)	γ	**−0.448**	**0.485**	**−0.449**	**0.460**	**−0.444**	**0.450**
	β	**−0.389**	**0.503**	**−0.437**	**0.468**	**−0.443**	**0.459**
		7.13	**6**	**11.09**	**6**	**15.57**	**6**
GMM(qdcn)	γ	0.008	0.113	0.007	0.052	0.005	0.037
	β	−0.133	0.399	−0.054	0.193	−0.017	0.134
		6.68	6	6.96	6	6.59	6
GMM(pr)	γ	**−0.124**	**0.269**	**−0.068**	**0.154**	**−0.043**	**0.120**
	β	**−0.219**	**0.408**	**−0.129**	**0.244**	**−0.082**	**0.202**
		5.31	**5**	**5.48**	**5**	**5.26**	**5**
GMM(qe)	γ	−0.301	0.417	−0.106	0.167	−0.069	0.114
	β	−0.181	0.292	−0.082	0.153	−0.056	0.117
		9.18	8	8.81	8	8.77	8
GMM(qecp)	γ	**−0.459**	**0.495**	**−0.438**	**0.447**	**−0.433**	**0.438**
	β	**−0.253**	**0.314**	**−0.255**	**0.264**	**−0.256**	**0.261**
		12.67	**10**	**21.45**	**10**	**31.62**	**10**
GMM(qecn)	γ	−0.034	0.118	0.000	0.049	0.001	0.036
	β	−0.064	0.283	−0.007	0.119	−0.002	0.083
		11.67	10	11.06	10	10.91	10
GMM(ex)	γ	**−0.150**	**0.289**	**−0.063**	**0.141**	**−0.046**	**0.097**
	β	**−0.158**	**0.315**	**−0.064**	**0.158**	**−0.043**	**0.112**
		9.91	**9**	**9.73**	**9**	**9.74**	**9**
GMM(sa)	γ	−0.089	0.195	−0.050	0.111	−0.037	0.085
	β	−0.125	0.287	−0.078	0.173	−0.046	0.139
		9.59	9	9.83	9	9.19	9
GMM(sacp)	γ	**−0.469**	**0.525**	**−0.594**	**0.612**	**−0.623**	**0.636**
	β	**−0.029**	**0.331**	**−0.203**	**0.260**	**−0.264**	**0.292**
		17.06	**12**	**43.04**	**12**	**71.02**	**12**
GMM(sacn)	γ	0.013	0.082	0.008	0.039	0.005	0.028
	β	−0.108	0.345	−0.039	0.159	−0.012	0.106
		13.37	12	13.67	12	13.02	12
PSM	γ(4)	**0.057**	**0.121**	**0.085**	**0.103**	**0.092**	**0.104**
	β(4)	**0.144**	**0.276**	**0.153**	**0.337**	**0.156**	**0.206**
	γ(8)	0.041	0.113	0.068	0.090	0.076	0.089
	β(8)	0.109	0.236	0.113	0.177	0.120	0.156
	γ(25)	**0.005**	**0.106**	**0.032**	**0.067**	**0.040**	**0.061**
	β(25)	**0.053**	**0.197**	**0.053**	**0.111**	**0.059**	**0.090**
	γ(50)	−0.012	0.108	0.014	0.061	0.022	0.051
	β(50)	0.031	0.184	0.028	0.094	0.034	0.071

See Notes of Tables.

Table 3 **Monte Carlo results for LFM, T=8** (Poisson model)

$\gamma=0.5$; $\beta=0.5$; $\rho=0.5$; $\tau=0.1$; $\sigma_\eta^2=0.5$; $\sigma_\varepsilon^2=0.5$

		N=100		N=500		N=1000	
		bias Sargan	rmse df	bias Sargan	rmse df	bias Sargan	rmse df
Level	γ	0.262	0.267	0.275	0.277	0.278	0.279
	β	0.537	0.586	0.550	0.565	0.559	0.568
WG	γ	**−0.189**	**0.198**	**−0.184**	**0.186**	**−0.185**	**0.186**
	β	**−0.126**	**0.139**	**−0.127**	**0.130**	**−0.127**	**0.129**
GMM(qd)	γ	−0.229	0.261	−0.076	0.096	−0.046	0.062
	β	−0.232	0.265	−0.105	0.131	−0.066	0.091
		18.11	16	18.21	16	17.94	16
GMM(qdcp)	γ	**−0.147**	**0.185**	**−0.019**	**0.044**	**−0.007**	**0.027**
	β	**−0.217**	**0.257**	**−0.057**	**0.093**	**−0.024**	**0.057**
		29.39	**22**	**28.97**	**22**	**27.80**	**22**
GMM(qdcn)	γ	0.235	0.241	0.225	0.226	0.221	0.221
	β	0.058	0.224	0.372	0.391	0.504	0.513
		29.20	22	43.53	22	47.94	22
GMM(pr)	γ	**−0.006**	**0.128**	**−0.029**	**0.054**	**−0.023**	**0.040**
	β	**−0.117**	**0.190**	**−0.064**	**0.096**	**−0.043**	**0.069**
		22.06	**21**	**21.58**	**21**	**21.61**	**21**
GMM(qe)	γ	−0.321	0.337	−0.080	0.092	−0.041	0.050
	β	−0.233	0.243	−0.081	0.091	−0.042	0.053
		54.45	52	57.47	52	56.42	52
GMM(qecp)	γ	**−0.261**	**0.281**	**−0.035**	**0.049**	**−0.012**	**0.025**
	β	**−0.221**	**0.236**	**−0.053**	**0.069**	**−0.021**	**0.038**
		64.11	**58**	**68.44**	**58**	**65.95**	**58**
GMM(qecn)	γ	0.160	0.168	0.225	0.226	0.241	0.241
	β	0.014	0.130	0.193	0.207	0.233	0.241
		63.27	58	95.13	58	123.58	58
GMM(ex)	γ	**0.011**	**0.179**	**−0.021**	**0.055**	**−0.019**	**0.036**
	β	**−0.129**	**0.212**	**−0.039**	**0.065**	**−0.025**	**0.045**
		58.39	**57**	**57.30**	**57**	**56.85**	**57**
GMM(sa)	γ	−0.012	0.079	−0.012	0.043	−0.009	0.031
	β	−0.070	0.134	−0.027	0.073	−0.017	0.053
		30.40	29	30.36	29	29.90	29
GMM(sacp)	γ	**0.029**	**0.077**	**0.010**	**0.037**	**0.007**	**0.027**
	β	**−0.012**	**0.142**	**−0.003**	**0.074**	**0.004**	**0.058**
		39.58	**36**	**38.85**	**36**	**37.69**	**36**
GMM(sacn)	γ	0.263	0.266	0.248	0.248	0.237	0.237
	β	−0.026	0.191	0.244	0.274	0.398	0.410
		45.69	36	80.79	36	102.29	36
PSM	γ**(4)**	**0.145**	**0.155**	**0.162**	**0.165**	**0.165**	**0.167**
	β**(4)**	**0.197**	**0.231**	**0.210**	**0.222**	**0.216**	**0.221**
	γ(8)	0.115	0.127	0.131	0.135	0.134	0.136
	β(8)	0.145	0.178	0.155	0.164	0.160	0.165
	γ**(25)**	**0.054**	**0.075**	**0.068**	**0.073**	**0.070**	**0.073**
	β**(25)**	**0.063**	**0.100**	**0.068**	**0.078**	**0.071**	**0.076**
	γ(50)	0.027	0.059	0.039	0.047	0.040	0.044
	β(50)	0.033	0.078	0.036	0.049	0.039	0.045

See Notes of Tables.

Table 4 **Monte Carlo results for LFM, T=8** (Negative binomial model)

$\gamma = 0.5$; $\beta = 0.5$; $\rho = 0.5$; $\tau = 0.1$; $\sigma_\eta^2 = 0.5$; $\sigma_\varepsilon^2 = 0.5$

		N=100		N=500		N=1000	
		bias	rmse	bias	rmse	bias	rmse
		Sargan	df	Sargan	df	Sargan	df
Level	γ	0.143	0.157	0.159	0.163	0.159	0.161
	β	0.289	0.348	0.300	0.315	0.295	0.302
WG	γ	**−0.221**	**0.237**	**−0.215**	**0.220**	**−0.216**	**0.219**
	β	**−0.136**	**0.170**	**−0.140**	**0.148**	**−0.144**	**0.147**
GMM(qd)	γ	−0.280	0.333	−0.106	0.133	−0.070	0.092
	β	−0.298	0.360	−0.153	0.186	−0.106	0.135
		18.67	16	18.88	16	18.00	16
GMM(qdcp)	γ	**−0.494**	**0.515**	**−0.434**	**0.441**	**−0.423**	**0.427**
	β	**−0.333**	**0.422**	**−0.287**	**0.331**	**−0.298**	**0.322**
		28.57	**22**	**58.02**	**22**	**90.16**	**22**
GMM(qdcn)	γ	−0.016	0.077	−0.001	0.031	0.000	0.021
	β	−0.197	0.302	−0.071	0.123	−0.038	0.081
		24.79	22	25.79	22	25.67	22
GMM(pr)	γ	**−0.006**	**0.171**	**−0.034**	**0.084**	**−0.035**	**0.062**
	β	**−0.180**	**0.310**	**−0.107**	**0.157**	**−0.083**	**0.119**
		22.66	**21**	**23.27**	**21**	**23.23**	**21**
GMM(qe)	γ	−0.379	0.409	−0.119	0.136	−0.070	0.084
	β	−0.275	0.296	−0.118	0.134	−0.075	0.089
		54.72	52	57.63	52	57.57	52
GMM(qecp)	γ	**−0.539**	**0.557**	**−0.452**	**0.456**	**−0.443**	**0.446**
	β	**−0.316**	**0.337**	**−0.259**	**0.264**	**−0.254**	**0.256**
		62.03	**58**	**94.85**	**58**	**135.30**	**58**
GMM(qecn)	γ	−0.090	0.123	−0.014	0.035	−0.006	0.022
	β	−0.181	0.248	−0.048	0.090	−0.026	0.056
		60.45	58	63.09	58	63.02	58
GMM(ex)	γ	**0.018**	**0.231**	**−0.016**	**0.101**	**−0.025**	**0.064**
	β	**−0.261**	**0.425**	**−0.112**	**0.178**	**−0.071**	**0.100**
		59.24	**57**	**62.71**	**57**	**62.74**	**57**
GMM(sa)	γ	−0.068	0.133	−0.041	0.071	−0.032	0.052
	β	−0.174	0.242	−0.087	0.127	−0.059	0.092
		31.11	29	31.68	29	31.23	29
GMM(sacp)	γ	**−0.444**	**0.476**	**−0.593**	**0.603**	**−0.638**	**0.644**
	β	**0.059**	**0.273**	**−0.069**	**0.164**	**−0.149**	**0.182**
		44.30	**36**	**114.44**	**36**	**188.60**	**36**
GMM(sacn)	γ	0.005	0.058	0.004	0.027	0.002	0.018
	β	−0.153	0.311	−0.073	0.142	−0.038	0.086
		39.12	36	41.62	36	41.75	36
PSM	γ**(4)**	**0.078**	**0.106**	**0.098**	**0.106**	**0.099**	**0.103**
	β**(4)**	**0.153**	**0.265**	**0.156**	**0.173**	**0.154**	**0.164**
	γ(8)	0.061	0.094	0.080	0.090	0.082	0.087
	β(8)	0.115	0.189	0.121	0.138	0.120	0.130
	γ**(25)**	**0.025**	**0.076**	**0.042**	**0.059**	**0.044**	**0.053**
	β**(25)**	**0.054**	**0.134**	**0.059**	**0.083**	**0.059**	**0.072**
	γ(50)	0.006	0.073	0.023	0.047	0.025	0.038
	β(50)	0.027	0.118	0.033	0.064	0.032	0.051

See Notes of Tables.

Notes of Tables

The number of replications is 1000.

The instrument sets for GMM estimators include no time dummies.

The initial consistent estimates used for the GMM estimation are obtained in the framework of the way described in Kitazawa (2007).

The symbols "Sargan" and "df" denote the Monte Carlo mean of values of the Sargan statistic for each GMM estimator and its degree of freedom, respectively.

As for the PSM estimators, the figures in the parentheses next to γ and β imply numbers of the latest pre-sample periods used for the estimations.

The replications where no convergence is achieved in the estimations and/or where the estimates whose absolute values exceed 10 (the latter of which fairly infrequently arise in using the Level and PSM estimators) are eliminated when calculating the values of the Monte Carlo statistics. Their rates are below 5 % in total for each experiment.

The values of the Monte Carlo bias and rmse exhibited in the tables are those obtained using the true values of γ and β as the starting values in the optimization for each replication. The values of these statistics obtained using the true values are not much different from those obtained using two different types of the starting values. The differences are below about 0.01 in terms of the absolute value in nearly all cases and below about 0.02 in almost all cases, while the differences are greater than 0.02 (but below 0.08) when the values are stained.

The individuals where the pre-sample means are zero are eliminated in each replication when estimating the parameters of interest using the PSM estimator.

The Monte Carlo means of proportions of zeros for the count dependent variables are about 22 % in Tables 1 and 3 where where the DGP is of the Poisson model, while about 32 % in Table 2 and 4 where the DGP is of the negative binomial model.