

# Some Additional Moment Conditions for a Dynamic Count Panel Data Model with Predetermined Explanatory Variables

Yoshitsugu Kitazawa

Faculty of Economics, Kyushu Sangyo University, Fukuoka, Japan

Email: [kitazawa@ip.kyusan-u.ac.jp](mailto:kitazawa@ip.kyusan-u.ac.jp)

Received July 10, 2013; revised August 10, 2013; accepted August 17, 2013

Copyright © 2013 Yoshitsugu Kitazawa. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

## ABSTRACT

This paper proposes some additional moment conditions for the linear feedback model with explanatory variables being predetermined, which is proposed by [1] for the purpose of dealing with count panel data. The newly proposed moment conditions include those associated with the equidispersion, the Negbin I-type model and the stationarity. The GMM estimators are constructed incorporating the additional moment conditions. Some Monte Carlo experiments indicate that the GMM estimators incorporating the additional moment conditions perform well, compared to that using only the conventional moment conditions proposed by [2,3].

**Keywords:** Count Panel Data; Linear Feedback Model; Moment Conditions; GMM; Monte Carlo Experiments

## 1. Introduction

Since the pioneering works conducted by [4,5], which aim at estimating the knowledge production function represented as the production of patents and innovations, various models and estimators have been proposed for the purpose of dealing with count panel data. The count panel data models are often discussed under the assumption that the time dimension is small but the cross-sectional size is large, which implies that the asymptotics of the estimators relies on the cross-sectional size. In this case, some problems need to be solved for the consistent estimation of the parameters of interest, when assuming the multiplicative fixed effects. Although [4] proposes the conditional maximum likelihood estimators (CMLEs) which rule out the fixed effects by using the reproductive property of Poisson and negative binomial distributions, these estimators are consistent only for the case with explanatory variables being strictly exogenous and dynamics being excluded.

In count panel data models, it is usual to regard the explanatory variables as being predetermined instead of being strictly exogenous. An example is the patent production function of a firm where the number of patents as a flow variable is a function of R&D expenditures. In this case, 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. For the case with explanatory variables being predetermined, [2,3] propose the quasi-differencing transformation, which eliminates the fixed effects to construct the valid moment conditions prepared for the Generalized Method of Moments (GMM) estimator proposed by [6]<sup>1</sup>, while [10] and [1] propose the Pre-Sample Mean (PSM) estimator, which uses the averages of the pre-sample histories of the count dependent variables as the proxies of the fixed effects<sup>2</sup>.

In addition, it is conceivable that incorporating dynamics into the model is also preferable for count panel data. The Linear Feedback Model (LFM) is proposed by [1], where the lagged dependent variables are included as additive regressors and therefore the problems associated with the explosive count dependent variables or the treatment of the zero-valued count dependent variables can be circumvented<sup>3</sup>. The LFM grows out of the integer-valued autoregressive model for the time series Poisson count model developed by [12-15]. The GMM and PSM esti-

<sup>1</sup>Some trailblazing estimations of the knowledge production function using the quasi-differencing transformation are conducted by [7-9], etc.

<sup>2</sup>The origin of the PSM estimator can be traced to [11].

<sup>3</sup>The alternative models incorporating dynamics are proposed by [7,10].

mators mentioned above are also applicable to the LFM<sup>4</sup>.

However, the GMM estimator based on the quasi-differencing transformation is afflicted with the undesirable small sample biases. Presumably this is due to the weak instruments problem when the cross-sectional size is small and/or the variables are persistent. In addition, the PSM estimator necessitates not only some strict assumptions for its consistency but also long pre-sample histories of the count dependent variables (whose availability would be ordinarily said to be low) for the improvement of its small sample performance<sup>5</sup>. These are indicated by Monte Carlo experiments previously conducted by [1].

In this paper, some valid additional moment conditions other than the conventional moment conditions based on the quasi-differencing transformation are proposed for the LFM with explanatory variables being predetermined, with the intension of improving the small sample performance of the GMM estimator. The additional moment conditions (and the conventional moment conditions) are derived on the basis of the variance-covariance structures originating from the conditional expectations for the disturbances in the LFM. The derivation method is analogous to that proposed by [23,24] in the framework of the ordinary dynamic panel data model, with the exception that the conditional expectation instead of the unconditional expectation is used in the variance-covariance restrictions.

The covariance restrictions among the disturbances give the conventional moment conditions and one type of the additional moment conditions related to predetermined regressors for the LFM. The former correspond to the first-differenced moment conditions proposed by [25, 26] in the framework of the ordinary dynamic panel data model, while the latter correspond to the additional non-linear moment conditions proposed by [23,24].

For the LFM with explanatory variables being predetermined, the relationships between variance and covariance for the disturbances can also give other types of the additional moment conditions. This paper proposes the moment conditions associated with the equidispersion reminiscent of Poisson distributed count dependent variables and those associated with the Negbin I-type model which is the negative binomial model introduced by [4] and characterizes one type of the overdispersion<sup>6</sup>. Although as shown by [1], the Poisson CMLE proposed by [4] requires no distributional assumption and accordingly

<sup>4</sup>Some empirical works with respect to the knowledge production function applying these estimators to the LFM are conducted by [16-22], etc.

<sup>5</sup>Even if the long pre-sample histories of the dependent variables are available, the PSM estimator is not consistent unless the fixed effects composing the explanatory variables are proportional to the fixed effects in the regression and the (finite) moment generating functions of the disturbance terms composing the explanatory variables are equal over time and for all individuals.

<sup>6</sup>The terminology "Negbin I-type model" is used in [27].

is also consistent for the Negbin I-type model, each of both GMM estimators using the moment conditions associated with the equidispersion and using those associated with the Negbin I-type model can discern each other's model.

If the stationary count dependent variables are assumed, the stationarity moment conditions are obtained as the additional moment conditions for the LFM with explanatory variables being predetermined: those based on the covariance restrictions, those associated with the equidispersion, and those associated with the Negbin I-type model. The first correspond to the stationarity moment conditions proposed by [28] and discussed by [24,29] in the framework of the ordinary dynamic panel data model.

Some Monte Carlo experiments are conducted for both configurations of the equidispersion and of the Negbin I-type model. It is shown that the joint usages of the conventional moment conditions with the additional moment conditions ameliorate the small sample performance of the GMM estimators, compared to their single usage. In addition, it is ascertained that both moment conditions associated with the equidispersion and associated with the Negbin I-type model can distinguish each other's model for large cross-sectional size.

The rest of the paper is organized as follows. In Section 2, the conventional and additional moment conditions are constructed on the basis of the variance-covariance restrictions on the disturbances for the LFM. In Section 3, some Monte Carlo experiments are carried out. Section 4 concludes.

## 2. Model, Moment Conditions, and GMM Estimators

In this section, some moment conditions are derived for the LFM with explanatory variables being predetermined: the moment conditions based on the covariance restrictions on the disturbances, the moment conditions associated with the equidispersion and the Negbin I-type model, and the stationarity moment conditions. The derivation method can be interpreted as an extension of the method proposed by [23,24] in the framework of the ordinary dynamic panel data model to the count panel data model.

### 2.1. Linear Feedback Model

With  $\phi_i = \exp(\eta_i)$  and  $\mu_{it} = \exp(\beta x_{it})$ , the simple LFM is written as follows:

$$y_{it} = \gamma y_{i,t-1} + u_{it}, \text{ for } t = 2, \dots, T, \quad (2.1.1)$$

$$u_{it} = \phi_i \mu_{it} + v_{it}, \text{ for } t = 2, \dots, T, \quad (2.1.2)$$

where the subscript  $i$  denotes the individual unit with  $i = 1, \dots, N$ , the subscript  $t$  denotes the time period,  $y_{it}$  is the (observable) count dependent variable,  $x_{it}$  is the (observable) continuous predetermined explanatory

variable,  $\eta_i$  is the (unobservable) individual specific fixed effect,  $v_{it}$  is the (unobservable) disturbance, and the parameters of interest are  $\gamma$  and  $\beta$ . The discussion is conducted for the case where  $N \rightarrow \infty$  but  $T$  is fixed.

Allowing for the uncorrelated structures between the initial dependent variable and the disturbances and between the fixed effect and the disturbances, the serially uncorrelated disturbances, and the predetermined explanatory variables, the assumptions for the disturbances are written as follows:

$$E[v_{it} | y_{i1}, \eta_i, v_i^{t-1}, x_i^t] = 0, \tag{2.1.3}$$

for  $t = 2, \dots, T$ ,

where  $v_i^{t-1} = (v_{i1}, \dots, v_{i,t-1})$  and  $x_i^t = (x_{i1}, \dots, x_{it})$ . The assumptions (2.1.3) are referred to as the original assumptions in this paper.

### 2.2. Covariance Restrictions and Moment Conditions

Different from the covariance restrictions considered by [23,24] in the context of the ordinary dynamic panel data model, those for the LFM, which originate from the original assumptions (2.1.3), are conditional on the information set  $(y_{i1}, \eta_i, v_i^{t-1}, x_i^t)$  as follows:

$$E[y_{i1}v_{it} | y_{i1}, \eta_i, v_i^{t-1}, x_i^t] = 0, \tag{2.2.1}$$

$$E[v_{is}v_{it} | y_{i1}, \eta_i, v_i^{t-1}, x_i^t] = 0, \text{ for } 2 \leq s \leq t-1, \tag{2.2.2}$$

$$E[x_{is}v_{it} | y_{i1}, \eta_i, v_i^{t-1}, x_i^t] = 0, \text{ for } 1 \leq s \leq t, \tag{2.2.3}$$

$$E[\eta_i v_{it} | y_{i1}, \eta_i, v_i^{t-1}, x_i^t] = 0, \tag{2.2.4}$$

where (2.2.4) is displayed for convenience, playing no role in constructing the valid moment conditions.

By replacing the unobservable variable  $v_{it}$  by the observable variable  $u_{it} = y_{it} - \gamma y_{i,t-1}$  (which is written by using the dependent variables and the parameter of interest) in (2.2.1) - (2.2.3), the following Equations are obtained:

$$E[y_{i1}u_{it} | y_{i1}, \eta_i, v_i^{t-1}, x_i^t] = y_{i1}\phi_t\mu_{it}, \tag{2.2.5}$$

$$E[u_{is}u_{it} | y_{i1}, \eta_i, v_i^{t-1}, x_i^t] = \phi_t^2\mu_{is}\mu_{it} + v_{is}\phi_t\mu_{it}, \tag{2.2.6}$$

for  $2 \leq s \leq t-1$ ,

$$E[x_{is}u_{it} | y_{i1}, \eta_i, v_i^{t-1}, x_i^t] = x_{is}\phi_t\mu_{it}, \tag{2.2.7}$$

for  $1 \leq s \leq t$ .

Utilizing the relationships holding among (2.2.5) - (2.2.7), the following  $(T-2)(T-1)/2$ ,  $T-3$ , and  $(T-1)T/2-1$  moment conditions are obtained:

$$E\left[y_{is} \left( \left( \frac{\mu_{i,t-1}}{\mu_{it}} \right) u_{it} - u_{i,t-1} \right) \right] = 0, \tag{2.2.8}$$

for  $s = 1, \dots, t-2$ ;  $t = 3, \dots, T$ ,

$$E\left[ \left( \left( \frac{\mu_{i,t-2}}{\mu_{i,t-1}} \right) u_{i,t-1} - u_{i,t-2} \right) \left( 1/\mu_{it} \right) u_{it} \right] = 0, \tag{2.2.9}$$

for  $t = 4, \dots, T$ ,

$$E\left[x_{is} \left( \left( \frac{\mu_{i,t-1}}{\mu_{it}} \right) u_{it} - u_{i,t-1} \right) \right] = 0, \tag{2.2.10}$$

for  $s = 1, \dots, t-1$ ;  $t = 3, \dots, T$ .

The moment conditions (2.2.8) are based on the relationships holding between  $y_{i1}u_{i,t-1}$  and  $y_{i1}u_{it}$  and holding between  $u_{is}u_{i,t-1}$  and  $u_{is}u_{it}$  for  $s \leq t-2$ , while the moment conditions (2.2.9) are based on the relationships holding between  $u_{i,t-2}u_{it}$  and  $u_{i,t-1}u_{it}$ . The moment conditions (2.2.10) are based on the relationships between  $x_{is}u_{i,t-1}$  and  $x_{is}u_{it}$  for  $s \leq t-1$ . The derivation of the moment conditions (2.2.8) - (2.2.10) is described in **Appendix A**.

The moment conditions (2.2.8) and (2.2.10) are the conventional moment conditions, which are linear with respect to  $\gamma$  and based on the quasi-differencing transformation proposed by [2,3], while the moment conditions (2.2.9) are the additional moment conditions nonlinear with respect to  $\gamma$  on the basis of the covariance restrictions on the disturbances<sup>7</sup>.

In this paper, the moment conditions (2.2.8) and (2.2.10) are referred to as the quasi-differenced moment conditions by convention, while the moment conditions (2.2.9) are referred to as the additional nonlinear moment conditions.

It can be said that the quasi-differenced moment conditions and the additional nonlinear moment conditions correspond to the first-differenced moment conditions (otherwise known as the standard moment conditions) proposed by [25,26] and the additional nonlinear moment conditions proposed by [23,24] in the framework of the ordinary dynamic panel data model, respectively.

The moment conditions (2.2.8) and (2.2.9) are the condensed full set of the relationships holding among (2.2.5) and (2.2.6) in the sense that the other relationships are indirectly traced based on these relationships, while the moment conditions (2.2.10) are the condensed full set of the relationships found among (2.2.7).

### 2.3. Equidispersion and Moment Conditions

If the assumptions of the equidispersion are imposed on the LFM, the variance-covariance restrictions are produced by the addition of the following restrictions to the

<sup>7</sup>For the model without dynamics, variants of the moment conditions analogous to (2.2.9) are proposed by [30] for the case of endogenous regressors and by [7] for the case of strictly exogenous regressors, respectively. The former is valid for the case of predetermined regressors, while the latter is not.

covariance restrictions (2.2.1) - (2.2.4):

$$E\left[\left(v_{it}^2 - y_{it}\right) \middle| y_{i1}, \eta_i, v_i^{t-1}, x_i^t\right] = 0. \tag{2.3.1}$$

By replacing the unobservable variable  $v_{it}$  by the observable variable  $u_{it}$  in (2.3.1), the following Equations are obtained:

$$E\left[\left(u_{it}^2 - y_{it}\right) \middle| y_{i1}, \eta_i, v_i^{t-1}, x_i^t\right] = \phi_i^2 \mu_{it}^2. \tag{2.3.2}$$

Utilizing the relationships holding among (2.2.6) and (2.3.2), the following two sets of  $T-2$  moment conditions are obtained:

$$E\left[y_{i,t-1} \left(\left(\mu_{i,t-1} / \mu_{it}\right) u_{it} - \left(u_{i,t-1} - 1\right)\right)\right] = 0, \tag{2.3.3}$$

for  $t = 3, \dots, T$ ,

$$E\left[\left(1 / \mu_{it}\right) \left(\left(\mu_{i,t-1} / \mu_{it}\right) u_{it} - u_{i,t-1}\right) u_{it} - \left(\mu_{i,t-1} / \mu_{it}\right) y_{it}\right] = 0, \text{ for } t = 3, \dots, T. \tag{2.3.4}$$

The moment conditions (2.3.3) are based on the relationships holding between  $u_{i,t-1}^2$  and  $u_{i,t-1}u_{it}$ , while the moment conditions (2.3.4) are based on the relationships holding between  $u_{i,t-1}u_{it}$  and  $u_{it}^2$ .

The moment conditions (2.3.3) are the additional moment conditions linear with respect to  $\gamma$  for the case of the equidispersion, while the moment conditions (2.3.4) are the additional moment conditions nonlinear with respect to  $\gamma$ .

In this paper, the moment conditions (2.3.3) and (2.3.4) are referred to as the additional linear equidispersion moment conditions and the additional nonlinear equidispersion moment conditions, respectively.

The moment conditions (2.2.8), (2.3.3) and (2.3.4) are the condensed full set of the relationships holding among (2.2.5), (2.2.6) and (2.3.2) (*i.e.* the condensed full set when the equidispersion is assumed). The derivation of the moment conditions (2.3.3) and (2.3.4) is described in **Appendix B**.

### 2.4. Negbin I-Type Model and Moment Conditions

If the assumptions of the Negbin I-type model are imposed on the LFM, the variance-covariance restrictions are produced by the addition of the following restrictions to the covariance restrictions (2.2.1) - (2.2.4):

$$E\left[\left(v_{it}^2 - (1 + \phi_i) y_{it}\right) \middle| y_{i1}, \eta_i, v_i^{t-1}, x_i^t\right] = 0. \tag{2.4.1}$$

By replacing the unobservable variable  $v_{it}$  by the observable variable  $u_{it}$  and the unobservable variable

$\phi_i y_{it}$  by the observable variable  $(y_{it} - u_{it}) u_{it} \left(\frac{1}{\mu_{it}}\right)$  in

(2.4.1), the following Equations are obtained:

$$E\left[\left(\left(u_{it}^2 - y_{it}\right) + \left(1 / \mu_{it}\right) \left(u_{it}^2 - y_{it} u_{it}\right)\right) \middle| y_{i1}, \eta_i, v_i^{t-1}, x_i^t\right] = \phi_i^2 \mu_{it} (\mu_{it} + 1). \tag{2.4.2}$$

Utilizing the relationships holding among (2.2.6) and (2.4.2), the following two sets of  $T-2$  moment conditions are obtained:

$$E\left[y_{i,t-1} \left(\left(\mu_{i,t-1} + 1\right) / \mu_{it}\right) u_{it} - \left(u_{i,t-1} - 1\right)\right] = 0, \tag{2.4.3}$$

for  $t = 3, \dots, T$ ,

$$E\left[\left(1 / \mu_{it}\right) \left(\left(\mu_{i,t-1} / \mu_{it}\right) u_{it} - u_{i,t-1}\right) u_{it} - y_{it} \cdot \left(\mu_{i,t-1} / \mu_{it}\right) u_{it} + y_{it} \left(\mu_{i,t-1} / \left(\mu_{it} + 1\right)\right) \left(u_{it} - 1\right)\right] = 0, \tag{2.4.4}$$

for  $t = 3, \dots, T$ .

The moment conditions (2.4.3) are based on the relationships holding between  $u_{i,t-1}^2$  and  $u_{i,t-1}u_{it}$ , while the moment conditions (2.4.4) are based on the relationships holding between  $u_{i,t-1}u_{it}$  and  $u_{it}^2$ .

The moment conditions (2.4.3) are the additional moment conditions linear with respect to  $\gamma$  for the case of the Negbin I-type model, while the moment conditions (2.4.4) are the additional moment conditions nonlinear with respect to  $\gamma$ .

In this paper, the moment conditions (2.4.3) and (2.4.4) are referred to as the additional linear Negbin I-type moment conditions and the additional nonlinear Negbin I-type moment conditions, respectively.

The moment conditions (2.2.8), (2.4.3) and (2.4.4) are the condensed full set of the relationships holding among (2.2.5), (2.2.6) and (2.4.2) (*i.e.* the condensed full set when the Negbin I-type model is assumed). The derivation of the moment conditions (2.4.3) and (2.4.4) is described in **Appendix C**.

### 2.5. Stationarity and Moment Conditions

The moment conditions concerning the stationarity are proposed and discussed in the ordinary dynamic panel data model (e.g. [24,28,29,31]). Likewise, they can be proposed in the LFM for count panel data.

If the explanatory variables  $x_{it}$  are stationary in terms of the moment generating functions as follows:

$$E\left[\exp(kx_{it}) \middle| \eta_i\right] = E\left[\varphi_i(k) \middle| \eta_i\right], \tag{2.5.1}$$

for  $t = 1, \dots, T$

(with  $k$  being any real number) and the initial condition of the count dependent variable is written as follows:

<sup>8</sup>Note that

$E\left[\phi_i y_{it} \middle| y_{i1}, \eta_i, v_i^{t-1}, x_i^t\right] = E\left[\left(y_{it} - u_{it}\right) u_{it} \left(1 / \mu_{it}\right) \middle| y_{i1}, \eta_i, v_i^{t-1}, x_i^t\right] + \phi_i^2 \mu_{it}$ , taking notice of (2.1.1) and (2.1.2) with (2.1.3).

$$y_{it} = (1/(1-\gamma))\phi_i\mu_{it} + v_{it}, \quad (2.5.2)$$

with

$$E[v_{it} | \eta_i, x_{it}] = 0, \quad (2.5.3)$$

the following  $T-2$  and  $T-1$  moment conditions are obtained by utilizing the relationships holding among (2.2.5) - (2.2.7):

$$E[\Delta y_{i,t-1} (1/\mu_{it}) u_{it}] = 0, \quad \text{for } t = 3, \dots, T, \quad (2.5.4)$$

$$E[\Delta x_{it} (1/\mu_{it}) u_{it}] = 0, \quad \text{for } t = 2, \dots, T. \quad (2.5.5)$$

The moment conditions (2.5.4) are based on the relationships holding between  $y_{it}u_{i3}$  and  $u_{i2}u_{i3}$  and holding between  $u_{i,t-2}u_{it}$  and  $u_{i,t-1}u_{it}$ , while the moment conditions (2.5.5) are based on the relationships holding between  $x_{i,t-1}u_{it}$  and  $x_{it}u_{it}$ <sup>9</sup>. The moment conditions (2.5.4) for  $t = 4, \dots, T$  are the replacement of the moment conditions (2.2.9).

If the assumptions of the equidispersion are imposed in addition to those concerning the stationarity, the following  $T-1$  moment conditions are obtained by utilizing the relationships holding among (2.2.5), (2.2.6), and (2.3.2):

$$E[(1/\mu_{it})(\Delta y_{it}u_{it} - y_{it})] = 0, \quad \text{for } t = 2, \dots, T. \quad (2.5.6)$$

Similarly, if the assumptions of the Negbin I-type model are imposed in addition to those concerning the stationarity, the following  $T-1$  moment conditions are obtained by utilizing the relationships holding among (2.2.5), (2.2.6), and (2.4.2):

$$E[y_{it} (1/(\mu_{it} + 1))(u_{it} - 1) - y_{i,t-1} (1/\mu_{it}) u_{it}] = 0, \quad (2.5.7)$$

for  $t = 2, \dots, T$ .

The moment conditions (2.5.6) and (2.5.7) for  $t = 3, \dots, T$  are the replacement of the moment conditions (2.3.4) and (2.4.4), respectively.

The moment conditions (2.5.4) and (2.5.5) are the moment conditions concerning the stationarity on the basis of the covariance restrictions among the disturbances and between the regressors and the disturbances, respectively, while the moment conditions (2.5.6) and (2.5.7) are the moment conditions concerning the stationarity for the cases of the equidispersion and the Negbin I-type model, respectively. These moment conditions are linear with respect to  $\gamma$ .

In this paper, the moment conditions (2.5.4) and (2.5.5) are referred to as the stationarity moment conditions,

<sup>9</sup>Although author proposes the moment conditions (2.5.4) and (2.5.5) in the discussion paper of Kyushu Sangyo University in December, 2007, [32] also proposes the same moment conditions in the framework of a specification of the LFM slightly different from that in this paper. It can be said that [32] also makes some other contribution to the estimation of the LFM.

while the moment conditions (2.5.6) and (2.5.7) are referred to as the stationarity & equidispersion moment conditions and the stationarity & Negbin I-type moment conditions, respectively.

It can be said that the stationarity moment conditions correspond to the stationarity moment conditions proposed by [28] and discussed by [24,29] in the framework of the ordinary dynamic panel data model.

When the stationarity is assumed, the moment conditions (2.2.8) and (2.5.4) are the condensed full set of the relationships holding among (2.2.5) and (2.2.6), while the moment conditions (2.2.10) and (2.5.5) are the condensed full set of the relationships holding among (2.2.7). The moment conditions (2.2.8), (2.3.3) and (2.5.6) are the condensed full set of the relationships holding among (2.2.5), (2.2.6) and (2.3.2) when the stationarity is assumed (*i.e.* the condensed full set when the stationarity and the equidispersion are assumed), while the moment conditions (2.2.8), (2.4.3) and (2.5.7) are the condensed full set of the relationships holding among (2.2.5), (2.2.6) and (2.4.2) when the stationarity is assumed (*i.e.* the condensed full set when the stationarity and the Negbin I-type model are assumed). The derivation of the moment conditions (2.5.4) - (2.5.7) is described in **Appendix D**.

## 2.6. GMM Estimator

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

$$E[g_i(\delta)] = 0, \quad (2.6.1)$$

where  $m$  is number of the moment conditions,  $\delta = [\gamma \beta']$ ,  $g_i(\delta)$  (which is the function of  $\delta$ ) is composed of the observable variables and  $\delta$  for the individual  $i$ .

Using the following empirical counterpart for (2.6.1):

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

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

$$\bar{g}(\delta)' W_N(\hat{\delta}_1) \bar{g}(\delta), \quad (2.6.3)$$

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

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

The efficient asymptotic variance for  $\delta$  is estimated by using

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

where  $D(\hat{\delta}) = \partial \bar{g}(\delta) / \partial \delta' |_{\delta=\hat{\delta}}$ . The GMM estimations for the LFM are explained in detail in [33,34].

Some GMM estimators are constructed for the LFM. The GMM estimators are classified into the three types.

First, the GMM estimators using the moment conditions linear with respect to  $\gamma$  (without taking account of the stationarity) are presented: the GMM(QD) estimator using only the quasi-differenced moment conditions (*i.e.* (2.2.8) and (2.2.10)), the GMM(QDP) estimator using the quasi-differenced moment conditions and the additional linear equidispersion moment conditions (*i.e.* (2.3.3)), and the GMM(QDN) estimator using the quasi-differenced moment conditions and the additional linear Negbin I-type moment conditions (*i.e.* (2.4.3)).

Second, the GMM estimators using the condensed full sets of the moment conditions under the provided assumptions (without taking account of the stationarity) are presented: the GMM(PR) estimator using the quasi-differenced moment conditions and the additional nonlinear moment conditions (*i.e.* (2.2.9)), the GMM(PRP) estimator using the quasi-differenced moment conditions, the additional linear equidispersion moment conditions and the additional nonlinear equidispersion moment conditions (*i.e.* (2.3.4)), and the GMM(PRN) estimator using the quasi-differenced moment conditions and the additional linear Negbin I-type moment conditions and the additional nonlinear Negbin I-type moment conditions (*i.e.* (2.4.4)).

Third, the GMM estimators incorporating the moment conditions concerning the stationarity are presented: the GMM(SA) estimator using the quasi-differenced moment conditions and the stationarity moment conditions (*i.e.* (2.5.4) and (2.5.5)), the GMM(SAP) estimator using the quasi-differenced moment conditions, the additional linear equidispersion moment conditions, the stationarity & equidispersion moment conditions (*i.e.* (2.5.6)) and the stationarity moment conditions with respect to  $x_{it}$  (*i.e.* (2.5.5)), and the GMM(SAN) estimator using the quasi-differenced moment conditions and the additional linear Negbin I-type moment conditions and the stationarity & Negbin I-type moment conditions (*i.e.* (2.5.7)) and the stationarity moment conditions with respect to  $x_{it}$ .

It should be noted that there can be a case where a manipulation is needed, when using the additional nonlinear moment conditions, the additional nonlinear equidispersion moment conditions, the additional nonlinear Negbin I-type moment conditions, the stationarity moment conditions, the stationarity & equidispersion moment conditions, and the stationarity & Negbin I-type moment conditions for the GMM estimations. If all values in  $x_{it}$  are positive (which are commonplace in the empirical analysis), the GMM estimates of  $\beta$  using these moment conditions seem to be in danger of going to infinity (see [3]). In this

case,  $x_{it}$  needs to be transformed in deviation from an appropriate value  $b$ , in order that  $x_{it}$  contains both positive and negative values evenly. The selection of  $b$  by [30] is the overall mean of  $x_{it}$  (*i.e.*  $b = (1/(NT)) \sum_{i=1}^N \sum_{t=1}^T x_{it}$ ). The GMM estimators subject to this transformation are the GMM(PR), GMM(PRP), GMM(PRN), GMM(SA), GMM(SAP), and GMM(SAN) estimators.

### 3. Monte Carlo

In this section, some small sample performances of the GMM estimators exhibited in previous section are investigated with Monte Carlo experiments. The experiments are implemented using the econometric software TSP version 4.5 (see [35]).

#### 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 being of the Negbin I-type.

The Poisson-type DGP is as follows:

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

$$y_{i,-T_G+1} \sim \text{Poisson}\left(\frac{\exp(\beta x_{i,-T_G+1} + \eta_i)}{1-\gamma}\right), \quad (3.1.2)$$

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

$$x_{i,-T_G+1} = \frac{\tau \eta_i}{1-\rho} + \frac{\varepsilon_{i,-T_G+1}}{\sqrt{1-\rho^2}}, \quad (3.1.4)$$

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

where  $t = -T_G + 1, \dots, -1, 0, 1, \dots, T$  with  $T_G$  being the number of pre-sample periods to be generated. In the DGP, values are set to the parameters  $\gamma$ ,  $\beta$ ,  $\rho$ ,  $\tau$ ,  $\sigma_\eta^2$  and  $\sigma_\varepsilon^2$ . The experiments are carried out with  $T_G = 50$ , 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  $N_R = 1000$ . This DGP setting is the same as that of [1], 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 and accordingly the dependent variables are in the stationary state if the absolute value of  $\gamma$  is less than one. The DGP is configured with the explanatory variables  $x_{it}$  being strictly exogenous<sup>10</sup>.

In the Negbin I-type DGP, (3.1.1) and (3.1.2) are re-

<sup>10</sup>[32] also conducts some Monte Carlo experiments in the similar setting.

placed by the following expressions respectively:

$$y_{it} \sim \text{Negbin}\left(\frac{\gamma y_{i,t-1} + \exp(\beta x_{it})}{\exp(\eta_i)}, \exp(\eta_i)\right), \quad (3.1.5)$$

$$y_{i,-TG+1} \sim \text{Negbin}\left(\frac{\exp(\beta x_{i,-TG+1})}{1-\gamma}, \exp(\eta_i)\right), \quad (3.1.6)$$

where the denotation  $X \sim \text{Negbin}(\alpha, \theta)$  implies that the count variable  $X$  is distributed as the negative binomial distribution whose probability function is

$$p(X) = \left(\frac{\Gamma(\alpha + X)}{\Gamma(\alpha)\Gamma(X+1)}\right) \left(\frac{1}{1+\theta}\right)^\alpha \left(\frac{\theta}{1+\theta}\right)^X$$

with  $\Gamma(\bullet)$  being the gamma function and  $\alpha$  and  $\theta$  being the parameters with  $\alpha \geq 0$  and  $\theta \geq 0$  respec-

tively.

### 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 PSM estimator. The Level and WG estimators are inconsistent in the DGP settings above. On the contrary, the PSM estimator is consistent if the long history is used in constructing the pre-sample means of the dependent variables. The details on these estimators are described in [1,10].

### 3.3. Results

For  $T = 4$ , Monte Carlo results for the Poisson-type DGP and the Negbin I-type DGP are shown in **Tables 1** and **2** respectively, while for  $T = 8$ , they are shown in **Tables 3** and **4** respectively.

**Table 1. Monte Carlo results for LFM, T = 4 (Poisson-type DGP).**

		N = 100		N = 500		N = 1000	
		bias	rmse	bias	rmse	bias	rmse
		<u>Sargan</u>	<u>df</u>	<u>Sargan</u>	<u>df</u>	<u>Sargan</u>	<u>df</u>
Level	$\gamma$	0.256	0.264	0.273	0.275	0.278	0.279
	$\beta$	0.545	0.656	0.549	0.571	0.557	0.573
WG	$\gamma$	<b>-0.452</b>	<b>0.463</b>	<b>-0.446</b>	<b>0.449</b>	<b>-0.446</b>	<b>0.447</b>
	$\beta$	<b>-0.260</b>	<b>0.272</b>	<b>-0.261</b>	<b>0.263</b>	<b>-0.263</b>	<b>0.264</b>
GMM(QD)	$\gamma$	-0.274	0.398	-0.104	0.161	-0.061	0.112
	$\beta$	-0.259	0.371	-0.124	0.219	-0.078	0.172
GMM(QDP)	$\gamma$	<u>4.42</u>	<u>4</u>	<u>4.58</u>	<u>4</u>	<u>4.58</u>	<u>4</u>
	$\beta$	<b>-0.054</b>	<b>0.155</b>	<b>-0.006</b>	<b>0.066</b>	<b>-0.001</b>	<b>0.045</b>
GMM(QDN)	$\gamma$	<b>-0.134</b>	<b>0.288</b>	<b>-0.028</b>	<b>0.148</b>	<b>-0.013</b>	<b>0.104</b>
	$\beta$	<u>8.72</u>	<u>6</u>	<u>7.91</u>	<u>6</u>	<u>7.56</u>	<u>6</u>
GMM(QD)	$\gamma$	0.262	0.273	0.234	0.237	0.227	0.228
	$\beta$	0.082	0.379	0.420	0.474	0.542	0.568
GMM(PR)	$\gamma$	<u>8.36</u>	<u>6</u>	<u>11.30</u>	<u>6</u>	<u>12.34</u>	<u>6</u>
	$\beta$	-0.090	0.214	-0.037	0.092	-0.016	0.067
GMM(PR)	$\gamma$	-0.159	0.282	-0.059	0.157	-0.026	0.134
	$\beta$	<u>5.62</u>	<u>5</u>	<u>5.66</u>	<u>5</u>	<u>5.66</u>	<u>5</u>
GMM(PR)	$\gamma$	<b>-0.024</b>	<b>0.158</b>	<b>-0.007</b>	<b>0.062</b>	<b>-0.002</b>	<b>0.043</b>
	$\beta$	<b>-0.166</b>	<b>0.281</b>	<b>-0.056</b>	<b>0.147</b>	<b>-0.029</b>	<b>0.107</b>
GMM(PR)	$\gamma$	<u>10.26</u>	<u>8</u>	<u>9.38</u>	<u>8</u>	<u>9.15</u>	<u>8</u>
	$\beta$	0.272	0.283	0.237	0.241	0.223	0.226
GMM(SA)	$\gamma$	-0.110	0.375	0.167	0.335	0.296	0.381
	$\beta$	<u>11.43</u>	<u>8</u>	<u>21.82</u>	<u>8</u>	<u>29.86</u>	<u>8</u>
GMM(SA)	$\gamma$	-0.023	0.139	-0.019	0.079	-0.010	0.059
	$\beta$	-0.053	0.212	-0.023	0.137	-0.012	0.104
GMM(SA)	$\gamma$	<u>9.70</u>	<u>9</u>	<u>9.46</u>	<u>9</u>	<u>9.46</u>	<u>9</u>
	$\beta$	<b>0.038</b>	<b>0.115</b>	<b>0.016</b>	<b>0.060</b>	<b>0.011</b>	<b>0.043</b>
GMM(SA)	$\gamma$	<b>-0.022</b>	<b>0.215</b>	<b>0.016</b>	<b>0.133</b>	<b>0.012</b>	<b>0.099</b>
	$\beta$	<u>14.06</u>	<u>12</u>	<u>12.87</u>	<u>12</u>	<u>12.91</u>	<u>12</u>
GMM(SAN)	$\gamma$	0.277	0.283	0.249	0.250	0.242	0.242
	$\beta$	0.036	0.263	0.350	0.391	0.471	0.490
PSM	$\gamma(4)$	<u>18.21</u>	<u>12</u>	<u>30.10</u>	<u>12</u>	<u>38.17</u>	<u>12</u>
	$\beta(4)$	0.132	0.156	0.157	0.162	0.163	0.167
PSM	$\gamma(8)$	0.191	0.296	0.205	0.225	0.211	0.229
	$\beta(8)$	<b>0.104</b>	<b>0.132</b>	<b>0.125</b>	<b>0.131</b>	<b>0.130</b>	<b>0.135</b>
PSM	$\gamma(25)$	<b>0.141</b>	<b>0.228</b>	<b>0.148</b>	<b>0.165</b>	<b>0.152</b>	<b>0.165</b>
	$\beta(25)$	0.046	0.091	0.061	0.072	0.066	0.073
PSM	$\gamma(50)$	0.058	0.139	0.062	0.083	0.065	0.078
	$\beta(50)$	<b>0.020</b>	<b>0.081</b>	<b>0.033</b>	<b>0.050</b>	<b>0.038</b>	<b>0.048</b>
PSM	$\gamma(50)$	<b>0.031</b>	<b>0.119</b>	<b>0.032</b>	<b>0.059</b>	<b>0.035</b>	<b>0.052</b>

See Notes of Tables.

**Table 2. Monte Carlo results for LFM, T = 4 (Negbin I-type DGP).**

		N = 100		N = 500		N = 1000	
		bias	rmse	bias	rmse	bias	rmse
		<u>Sargan</u>	<u>df</u>	<u>Sargan</u>	<u>df</u>	<u>Sargan</u>	<u>df</u>
Level	$\gamma$	0.133	0.165	0.154	0.162	0.158	0.163
	$\beta$	0.302	0.419	0.300	0.357	0.303	0.325
WG	$\gamma$	<b>-0.513</b>	<b>0.533</b>	<b>-0.500</b>	<b>0.506</b>	<b>-0.500</b>	<b>0.503</b>
	$\beta$	<b>-0.264</b>	<b>0.294</b>	<b>-0.275</b>	<b>0.280</b>	<b>-0.276</b>	<b>0.279</b>
GMM(QD)	$\gamma$	-0.350	0.521	-0.156	0.249	-0.098	0.165
	$\beta$	-0.327	0.496	-0.205	0.295	-0.132	0.242
		<u>4.27</u>	<u>4</u>	<u>4.45</u>	<u>4</u>	<u>4.36</u>	<u>4</u>
GMM(QDP)	$\gamma$	<b>-0.448</b>	<b>0.485</b>	<b>-0.449</b>	<b>0.460</b>	<b>-0.444</b>	<b>0.450</b>
	$\beta$	<b>-0.389</b>	<b>0.503</b>	<b>-0.437</b>	<b>0.468</b>	<b>-0.443</b>	<b>0.459</b>
		<u>7.13</u>	<u>6</u>	<u>11.09</u>	<u>6</u>	<u>15.57</u>	<u>6</u>
GMM(QDN)	$\gamma$	0.008	0.113	0.007	0.052	0.005	0.037
	$\beta$	-0.133	0.399	-0.054	0.193	-0.017	0.134
		<u>6.68</u>	<u>6</u>	<u>6.96</u>	<u>6</u>	<u>6.59</u>	<u>6</u>
GMM(PR)	$\gamma$	-0.124	0.269	-0.068	0.154	-0.043	0.120
	$\beta$	-0.219	0.408	-0.129	0.244	-0.082	0.202
		<u>5.31</u>	<u>5</u>	<u>5.48</u>	<u>5</u>	<u>5.26</u>	<u>5</u>
GMM(PRP)	$\gamma$	<b>-0.488</b>	<b>0.531</b>	<b>-0.547</b>	<b>0.565</b>	<b>-0.554</b>	<b>0.564</b>
	$\beta$	<b>-0.296</b>	<b>0.366</b>	<b>-0.337</b>	<b>0.351</b>	<b>-0.346</b>	<b>0.353</b>
		<u>13.04</u>	<u>8</u>	<u>36.36</u>	<u>8</u>	<u>59.54</u>	<u>8</u>
GMM(PRN)	$\gamma$	0.019	0.101	0.010	0.050	0.006	0.034
	$\beta$	-0.176	0.403	-0.080	0.202	-0.041	0.138
		<u>8.73</u>	<u>8</u>	<u>9.02</u>	<u>8</u>	<u>8.89</u>	<u>8</u>
GMM(SA)	$\gamma$	-0.089	0.195	-0.050	0.111	-0.037	0.085
	$\beta$	-0.125	0.287	-0.078	0.173	-0.046	0.139
		<u>9.59</u>	<u>9</u>	<u>9.83</u>	<u>9</u>	<u>9.19</u>	<u>9</u>
GMM(SAP)	$\gamma$	<b>-0.469</b>	<b>0.525</b>	<b>-0.594</b>	<b>0.612</b>	<b>-0.623</b>	<b>0.636</b>
	$\beta$	<b>-0.029</b>	<b>0.331</b>	<b>-0.203</b>	<b>0.260</b>	<b>-0.264</b>	<b>0.292</b>
		<u>17.06</u>	<u>12</u>	<u>43.04</u>	<u>12</u>	<u>71.02</u>	<u>12</u>
GMM(SAN)	$\gamma$	0.013	0.082	0.008	0.039	0.005	0.028
	$\beta$	-0.108	0.345	-0.039	0.159	-0.012	0.106
		<u>13.37</u>	<u>12</u>	<u>13.67</u>	<u>12</u>	<u>13.02</u>	<u>12</u>
PSM	$\gamma(4)$	0.057	0.121	0.085	0.103	0.092	0.104
	$\beta(4)$	0.144	0.276	0.153	0.337	0.156	0.206
	$\gamma(8)$	<b>0.041</b>	<b>0.113</b>	<b>0.068</b>	<b>0.090</b>	<b>0.076</b>	<b>0.089</b>
	$\beta(8)$	<b>0.109</b>	<b>0.236</b>	<b>0.113</b>	<b>0.177</b>	<b>0.120</b>	<b>0.156</b>
	$\gamma(25)$	0.005	0.106	0.032	0.067	0.040	0.061
	$\beta(25)$	0.053	0.197	0.053	0.111	0.059	0.090
	$\gamma(50)$	<b>-0.012</b>	<b>0.108</b>	<b>0.014</b>	<b>0.061</b>	<b>0.022</b>	<b>0.051</b>
	$\beta(50)$	<b>0.031</b>	<b>0.184</b>	<b>0.028</b>	<b>0.094</b>	<b>0.034</b>	<b>0.071</b>

See Notes of Tables.



**Table 3. Monte Carlo results for LFM, T = 8 (Poisson-type DGP).**

		N = 100		N = 500		N = 1000	
		bias	rmse	bias	rmse	bias	rmse
		<u>Sargan</u>	<u>df</u>	<u>Sargan</u>	<u>df</u>	<u>Sargan</u>	<u>df</u>
Level	$\gamma$	0.262	0.267	0.275	0.277	0.278	0.279
	$\beta$	0.537	0.586	0.550	0.565	0.559	0.568
<b>WG</b>	$\gamma$	<b>-0.189</b>	<b>0.198</b>	<b>-0.184</b>	<b>0.186</b>	<b>-0.185</b>	<b>0.186</b>
	$\beta$	<b>-0.126</b>	<b>0.139</b>	<b>-0.127</b>	<b>0.130</b>	<b>-0.127</b>	<b>0.129</b>
GMM(QD)	$\gamma$	-0.229	0.261	-0.076	0.096	-0.046	0.062
	$\beta$	-0.232	0.265	-0.105	0.131	-0.066	0.091
<b>GMM(QDP)</b>	$\gamma$	<b>-0.147</b>	<b>0.185</b>	<b>-0.019</b>	<b>0.044</b>	<b>-0.007</b>	<b>0.027</b>
	$\beta$	<b>-0.217</b>	<b>0.257</b>	<b>-0.057</b>	<b>0.093</b>	<b>-0.024</b>	<b>0.057</b>
GMM(QDN)	$\gamma$	0.235	0.241	0.225	0.226	0.221	0.221
	$\beta$	0.058	0.224	0.372	0.391	0.504	0.513
GMM(PR)	$\gamma$	-0.006	0.128	-0.029	0.054	-0.023	0.040
	$\beta$	-0.117	0.190	-0.064	0.096	-0.043	0.069
<b>GMM(PRP)</b>	$\gamma$	<b>-0.007</b>	<b>0.103</b>	<b>-0.006</b>	<b>0.038</b>	<b>-0.003</b>	<b>0.026</b>
	$\beta$	<b>-0.154</b>	<b>0.213</b>	<b>-0.066</b>	<b>0.097</b>	<b>-0.031</b>	<b>0.060</b>
GMM(PRN)	$\gamma$	0.253	0.257	0.228	0.229	0.211	0.212
	$\beta$	-0.154	0.268	0.061	0.203	0.206	0.246
GMM(SA)	$\gamma$	-0.012	0.079	-0.012	0.043	-0.009	0.031
	$\beta$	-0.070	0.134	-0.027	0.073	-0.017	0.053
<b>GMM(SAP)</b>	$\gamma$	<b>0.029</b>	<b>0.077</b>	<b>0.010</b>	<b>0.037</b>	<b>0.007</b>	<b>0.027</b>
	$\beta$	<b>-0.012</b>	<b>0.142</b>	<b>-0.003</b>	<b>0.074</b>	<b>0.004</b>	<b>0.058</b>
GMM(SAN)	$\gamma$	0.263	0.266	0.248	0.248	0.237	0.237
	$\beta$	-0.026	0.191	0.244	0.274	0.398	0.410
PSM	$\gamma(4)$	0.145	0.155	0.162	0.165	0.165	0.167
	$\beta(4)$	0.197	0.231	0.210	0.222	0.216	0.221
	$\gamma(8)$	<b>0.115</b>	<b>0.127</b>	<b>0.131</b>	<b>0.135</b>	<b>0.134</b>	<b>0.136</b>
	$\beta(8)$	<b>0.145</b>	<b>0.178</b>	<b>0.155</b>	<b>0.164</b>	<b>0.160</b>	<b>0.165</b>
	$\gamma(25)$	0.054	0.075	0.068	0.073	0.070	0.073
	$\beta(25)$	0.063	0.100	0.068	0.078	0.071	0.076
	$\gamma(50)$	<b>0.027</b>	<b>0.059</b>	<b>0.039</b>	<b>0.047</b>	<b>0.040</b>	<b>0.044</b>
	$\beta(50)$	<b>0.033</b>	<b>0.078</b>	<b>0.036</b>	<b>0.049</b>	<b>0.039</b>	<b>0.045</b>

See Notes of Tables.

Table 4. Monte Carlo results for LFM, T = 8 (Negbin I-type DGP).

		N = 100		N = 500		N = 1000	
		bias	rmse	bias	rmse	bias	rmse
		<u>Sargan</u>	<u>df</u>	<u>Sargan</u>	<u>df</u>	<u>Sargan</u>	<u>df</u>
Level	$\gamma$	0.143	0.157	0.159	0.163	0.159	0.161
	$\beta$	0.289	0.348	0.300	0.315	0.295	0.302
WG	$\gamma$	<b>-0.221</b>	<b>0.237</b>	<b>-0.215</b>	<b>0.220</b>	<b>-0.216</b>	<b>0.219</b>
	$\beta$	<b>-0.136</b>	<b>0.170</b>	<b>-0.140</b>	<b>0.148</b>	<b>-0.144</b>	<b>0.147</b>
GMM(QD)	$\gamma$	-0.280	0.333	-0.106	0.133	-0.070	0.092
	$\beta$	-0.298	0.360	-0.153	0.186	-0.106	0.135
		<u>18.67</u>	<u>16</u>	<u>18.88</u>	<u>16</u>	<u>18.00</u>	<u>16</u>
GMM(QDP)	$\gamma$	<b>-0.494</b>	<b>0.515</b>	<b>-0.434</b>	<b>0.441</b>	<b>-0.423</b>	<b>0.427</b>
	$\beta$	<b>-0.333</b>	<b>0.422</b>	<b>-0.287</b>	<b>0.331</b>	<b>-0.298</b>	<b>0.322</b>
		<u>28.57</u>	<u>22</u>	<u>58.02</u>	<u>22</u>	<u>90.16</u>	<u>22</u>
GMM(QDN)	$\gamma$	-0.016	0.077	-0.001	0.031	0.000	0.021
	$\beta$	-0.197	0.302	-0.071	0.123	-0.038	0.081
		<u>24.79</u>	<u>22</u>	<u>25.79</u>	<u>22</u>	<u>25.67</u>	<u>22</u>
GMM(PR)	$\gamma$	-0.006	0.171	-0.034	0.084	-0.035	0.062
	$\beta$	-0.180	0.310	-0.107	0.157	-0.083	0.119
		<u>22.66</u>	<u>21</u>	<u>23.27</u>	<u>21</u>	<u>23.23</u>	<u>21</u>
GMM(PRP)	$\gamma$	<b>-0.422</b>	<b>0.465</b>	<b>-0.505</b>	<b>0.514</b>	<b>-0.532</b>	<b>0.536</b>
	$\beta$	<b>-0.305</b>	<b>0.351</b>	<b>-0.336</b>	<b>0.343</b>	<b>-0.347</b>	<b>0.350</b>
		<u>36.33</u>	<u>28</u>	<u>81.18</u>	<u>28</u>	<u>128.67</u>	<u>28</u>
GMM(PRN)	$\gamma$	0.005	0.070	0.004	0.032	0.001	0.021
	$\beta$	-0.197	0.326	-0.107	0.171	-0.067	0.110
		<u>30.96</u>	<u>28</u>	<u>34.41</u>	<u>28</u>	<u>35.01</u>	<u>28</u>
GMM(SA)	$\gamma$	-0.068	0.133	-0.041	0.071	-0.032	0.052
	$\beta$	-0.174	0.242	-0.087	0.127	-0.059	0.092
		<u>31.11</u>	<u>29</u>	<u>31.68</u>	<u>29</u>	<u>31.23</u>	<u>29</u>
GMM(SAP)	$\gamma$	<b>-0.444</b>	<b>0.476</b>	<b>-0.593</b>	<b>0.603</b>	<b>-0.638</b>	<b>0.644</b>
	$\beta$	<b>0.059</b>	<b>0.273</b>	<b>-0.069</b>	<b>0.164</b>	<b>-0.149</b>	<b>0.182</b>
		<u>44.30</u>	<u>36</u>	<u>114.44</u>	<u>36</u>	<u>188.60</u>	<u>36</u>
GMM(SAN)	$\gamma$	0.005	0.058	0.004	0.027	0.002	0.018
	$\beta$	-0.153	0.311	-0.073	0.142	-0.038	0.086
		<u>39.12</u>	<u>36</u>	<u>41.62</u>	<u>36</u>	<u>41.75</u>	<u>36</u>
PSM	$\gamma(4)$	0.078	0.106	0.098	0.106	0.099	0.103
	$\beta(4)$	0.153	0.265	0.156	0.173	0.154	0.164
	$\gamma(8)$	<b>0.061</b>	<b>0.094</b>	<b>0.080</b>	<b>0.090</b>	<b>0.082</b>	<b>0.087</b>
	$\beta(8)$	<b>0.115</b>	<b>0.189</b>	<b>0.121</b>	<b>0.138</b>	<b>0.120</b>	<b>0.130</b>
	$\gamma(25)$	0.025	0.076	0.042	0.059	0.044	0.053
	$\beta(25)$	0.054	0.134	0.059	0.083	0.059	0.072
	$\gamma(50)$	<b>0.006</b>	<b>0.073</b>	<b>0.023</b>	<b>0.047</b>	<b>0.025</b>	<b>0.038</b>
	$\beta(50)$	<b>0.027</b>	<b>0.118</b>	<b>0.033</b>	<b>0.064</b>	<b>0.032</b>	<b>0.051</b>

See Notes of Tables.

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. These results are almost the same as those obtained by [1].

It can be seen that the GMM estimators incorporating the additional moment conditions behave better than the GMM(QD) estimator using only the conventional moment conditions based on the quasi-differencing transformation, as long as the additional moment conditions are valid.

It is considered that the GMM(QD) estimator suffers from the weak instruments problem pointed out by [36,37], due to exclusively using the quasi-differenced moment conditions applying the lagged levels of dependent and explanatory variables (which are regarded as the weak instruments) to the quasi-differencing transformations (see, e.g., [1]). The weak instruments problem also occurs in the ordinary dynamic panel data model (see, e.g., [29]).

The small sample property of the distribution-free GMM estimator incorporating the additional nonlinear moment conditions (*i.e.* GMM(PR) estimator) is more preferable than that of the conventional GMM(QD) estimator. It can be said that the joint usage of the quasi-differenced moment conditions and the additional nonlinear moment conditions improves the small sample property of the GMM estimator, compared to the single usage of the quasi-differenced moment conditions. This result is similar to that for the ordinary dynamic panel data model (see, e.g., [38]).

For the Poisson-type DGP (see **Table 1** and **3**), the small sample properties of the GMM estimators incorporating the moment conditions associated with equidispersion (*i.e.* GMM(QDP), GMM(PRP), and GMM(SAP) estimators) improve as the cross-sectional size  $N$  increases from 100, 500 to 1000, while those of the GMM estimators incorporating the moment conditions associated with Negbin I-type model (*i.e.* GMM(QDN), GMM(PRN), and GMM(SAN) estimators) deteriorate, where the augmentations of the Monte Carlo means of Sargan test statistics (which are the reflection of the inconsistency) are recognizable for larger  $N$ . For the Negbin I-type DGP (see **Tables 2** and **4**), the countertrend is found. It is shown that each of both GMM estimators incorporating the moment conditions associated with equidispersion and incorporating those associated with the Negbin I-type model can discern each other's underlying specification<sup>11</sup>.

<sup>11</sup>Although [39] proposes the moment condition based on the variance-covariance restrictions under the assumption with the explanatory variables being strictly exogenous and dynamics being excluded, it cannot distinguish between the equidispersion and the Negbin I-type model. However, it is shown that the moment conditions based on the variance-covariance restrictions proposed in this paper can distinguish between them, under the assumption weaker than [39].

The performances of the GMM estimators incorporating the moment conditions concerning the stationarity (*i.e.* GMM(SA), GMM(SAP), GMM(SAN) estimators) are fairly favorable (especially in terms of bias), compared to that of the conventional GMM(QD) estimator, as long as the moment conditions used are valid. These results are similar to that for the ordinary dynamic panel data model, in which the additional usage of the stationarity moment conditions improves the small sample performance of the GMM estimator (see [29], etc.).

#### 4. Conclusion

In this paper, some additional moment conditions other than the conventional quasi-differenced moment conditions were newly proposed for the LFM with explanatory variables being predetermined: the additional nonlinear moment conditions, the additional linear equidispersion moment conditions, the additional nonlinear equidispersion moment conditions, the additional linear Negbin I-type moment conditions, the additional nonlinear Negbin I-type moment conditions, the stationarity moment conditions, the stationarity & equidispersion moment conditions, and the stationarity & Negbin I-type moment conditions. In the limited Monte Carlo experiments, it was shown that the GMM estimators perform better when incorporating the additional moment conditions than when using the conventional moment conditions only, as long as the additional moment conditions are valid.

#### 5. Acknowledgements

This paper is a collection of the oral presentations conducted in the 2008 International Symposium on Econometric Theory and Applications, Seoul National University (May 2008), the 2011 Asian Meeting of The Econometric Society, Korea University (August 2011), and the 18th International Panel Data Conference, Banque de France (July 2012). The presentations were also conducted in the domestic conferences in Japan: the meetings of The Japanese Economic Association (Kinki University, September 2008; Kwansai Gakuin University, September 2010), the meeting of The Japan Association for Applied Economics (Kobe University, November 2009), and the meeting of Kansai Keiryō Keizaigaku Kenkyukai (Kyoto University, January 2010) and in the seminars in Korea and Japan: Korea University (March 2008) and Kyushu University (June 2009). Author expresses the gratitude to the participants and the commentators: Kosuke Oya, Atsushi Fujii, and Atsushi Yoshida.

#### REFERENCES

- [1] R. Blundell, R. Griffith and F. Windmeijer, "Individual Effects and Dynamics in Count Data Models," *Journal of Econometrics*, Vol. 108, No. 1, 2002, pp. 113-131.

- [http://dx.doi.org/10.1016/S0304-4076\(01\)00108-7](http://dx.doi.org/10.1016/S0304-4076(01)00108-7)
- [2] G. Chamberlain, "Comment: Sequential Moment Restrictions in Panel Data," *Journal of Business and Economic Statistics*, Vol. 10, No. 1, 1992, pp. 20-26. <http://dx.doi.org/10.1080/07350015.1992.10509881>
- [3] J. M. Wooldridge, "Multiplicative Panel Data Models without the Strict Exogeneity Assumption," *Econometric Theory*, Vol. 13, No. 5, 1997, pp. 667-678. <http://dx.doi.org/10.1017/S0266466600006125>
- [4] J. A. Hausman, B. H. Hall and Z. Griliches, "Econometric Models for Count Data with an Application to the Patent-R&D Relationship," *Econometrica*, Vol. 52, No. 4, 1984, pp. 909-938. <http://dx.doi.org/10.2307/1911191>
- [5] B. H. Hall, Z. Griliches and J. A. Hausman, "Patents and R and D: Is There a Lag?" *International Economic Review*, Vol. 27, No. 2, 1986, pp. 265-283. <http://dx.doi.org/10.2307/2526504>
- [6] L. P. Hansen, "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, Vol. 50, No. 4, 1982, pp. 1029-1054. <http://dx.doi.org/10.2307/1912775>
- [7] B. Crépon and E. Duguet, "Estimating the Innovation Function from Patent Numbers: GMM on Count Panel Data," *Journal of Applied Econometrics*, Vol. 12, No. 3, 1997, pp. 243-263. [http://dx.doi.org/10.1002/\(SICI\)1099-1255\(199705\)12:3<243::AID-JAE444>3.0.CO;2-4](http://dx.doi.org/10.1002/(SICI)1099-1255(199705)12:3<243::AID-JAE444>3.0.CO;2-4)
- [8] J. G. Montalvo, "GMM Estimation of Count-Panel-Data Models with Fixed Effects and Predetermined Instruments," *Journal of Business and Economic Statistics*, Vol. 15, No. 1, 1997, pp. 82-89. <http://dx.doi.org/10.1080/07350015.1997.10524690>
- [9] J. Kim and G. Marschke, "Labor Mobility of Scientists, Technological Diffusion and the Firm's Patenting Decision," *The RAND Journal of Economics*, Vol. 36, No. 2, 2005, pp. 298-317.
- [10] R. Blundell, R. Griffith and J. Van Reenen, "Market Share, Market Value and Innovation in a Panel of British Manufacturing Firms," *Review of Economic Studies*, Vol. 66, No. 3, 1999, pp. 529-554. <http://dx.doi.org/10.1111/1467-937X.00097>
- [11] R. Blundell, R. Griffith and J. Van Reenen, "Dynamic Count Data Models of Technological Innovation," *Economic Journal*, Vol. 105, No. 429, 1995, pp. 333-344. <http://dx.doi.org/10.2307/2235494>
- [12] M. A. Al-Osh and A. A. Alzaid, "First-Order Integer-Valued Autoregressive (INAR(1)) Process," *Journal of Time Series Analysis*, Vol. 8, No. 3, 1987, pp. 261-275. <http://dx.doi.org/10.1111/j.1467-9892.1987.tb00438.x>
- [13] E. McKenzie, "Some ARMA Models for Dependent Sequences of Poisson Counts," *Advances in Applied Probability*, Vol. 20, No. 4, 1988, pp. 822-835. <http://dx.doi.org/10.2307/1427362>
- [14] A. A. Alzaid and M. A. Al-Osh, "An Integer-Valued  $p$ th-Order Autoregressive Structure (INAR( $p$ )) Process," *Journal of Applied Probability*, Vol. 27, No. 2, 1990, pp. 314-324. <http://dx.doi.org/10.2307/3214650>
- [15] D. Jin-Guan and L. Yuan, "The Integer-Valued Autoregressive (INAR( $p$ )) Model," *Journal of Time Series Analysis*, Vol. 12, No. 2, 1991, pp. 129-142. <http://dx.doi.org/10.1111/j.1467-9892.1991.tb00073.x>
- [16] M. Cincera, "Patents, R&D, and Technological Spillovers at the Firm Level: Some Evidence from Econometric Count Models for Panel Data," *Journal of Applied Econometrics*, Vol. 12, No. 3, 1997, pp. 265-280. [http://dx.doi.org/10.1002/\(SICI\)1099-1255\(199705\)12:3<265::AID-JAE439>3.0.CO;2-J](http://dx.doi.org/10.1002/(SICI)1099-1255(199705)12:3<265::AID-JAE439>3.0.CO;2-J)
- [17] R. M. Salomon and J. M. Shaver, "Learning by Exporting: New Insights from Examining Firm Innovation," *Journal of Economics and Management Strategy*, Vol. 14, No. 2, 2005, pp. 431-460. <http://dx.doi.org/10.1111/j.1530-9134.2005.00047.x>
- [18] Y. Uchida and P. Cook, "Innovation and Market Structure in the Manufacturing Sector: An Application of Linear Feedback Models," *Oxford Bulletin of Economics and Statistics*, Vol. 69, No. 4, 2007, pp. 557-580. <http://dx.doi.org/10.1111/j.1468-0084.2007.00450.x>
- [19] M. Abdelmoula and G. Bresson, "Spatial and Technological Spillovers in European Patenting Activities: A Dynamic Count Panel Data Model," *Annales d'Économie et de Statistique*, No. 87/88, 2008, pp. 167-194.
- [20] S. Gurmur and F. Pérez-Sebastián, "Patents, R&D and Lag Effects: Evidence from Flexible Methods for Count Panel Data on Manufacturing Firms," *Empirical Economics*, Vol. 35, No. 3, 2008, pp. 507-526. <http://dx.doi.org/10.1007/s00181-007-0176-8>
- [21] A. Lucena, "The Organizational Designs of R&D Activities and Their Performance Implications: Empirical Evidence for Spain," *Industry and Innovation*, Vol. 18, No. 2, 2011, pp. 151-176. <http://dx.doi.org/10.1080/13662716.2011.541103>
- [22] E. P. Gallié and D. Legros, "Firms' Human Capital, R&D and Innovation: A Study on French Firms," *Empirical Economics*, Vol. 43, No. 2, 2012, pp. 581-596. <http://dx.doi.org/10.1007/s00181-011-0506-8>
- [23] S. C. Ahn, "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, 1990.
- [24] S. C. Ahn and P. Schmidt, "Efficient Estimation of Models for Dynamic Panel Data," *Journal of Econometrics*, Vol. 68, No. 1, 1995, pp. 5-27. [http://dx.doi.org/10.1016/0304-4076\(94\)01641-C](http://dx.doi.org/10.1016/0304-4076(94)01641-C)
- [25] D. Holtz-Eakin, W. Newey and H. S. Rosen, "Estimating Vector Autoregressions with Panel Data," *Econometrica*, Vol. 56, No. 6, 1988, pp. 1371-1395. <http://dx.doi.org/10.2307/1913103>
- [26] M. Arellano and S. Bond, "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations," *Review of Economic Studies*, Vol. 58, No. 2, 1991, pp. 277-297. <http://dx.doi.org/10.2307/2297968>
- [27] R. Winkelmann, "Econometric Analysis of Count Data, 5th Edition," Springer, Berlin, Heidelberg, 2008.
- [28] M. Arellano and O. Bover, "Another Look at the Instrumental Variables Estimation of Error-Components Mod-

- els,” *Journal of Econometrics*, Vol. 68, No. 1, 1995, pp. 29-51. [http://dx.doi.org/10.1016/0304-4076\(94\)01642-D](http://dx.doi.org/10.1016/0304-4076(94)01642-D)
- [29] R. Blundell and S. Bond, “Initial Conditions and Moment Restrictions in Dynamic Panel Data Models,” *Journal of Econometrics*, Vol. 87, No. 1, 1998, pp. 115-143. [http://dx.doi.org/10.1016/S0304-4076\(98\)00009-8](http://dx.doi.org/10.1016/S0304-4076(98)00009-8)
- [30] F. Windmeijer, “Moment Conditions for Fixed Effects Count Data Models with Endogenous Regressors,” *Economics Letters*, Vol. 68, No. 1, 2000, pp. 21-24. [http://dx.doi.org/10.1016/S0165-1765\(00\)00228-7](http://dx.doi.org/10.1016/S0165-1765(00)00228-7)
- [31] S. C. Ahn and P. Schmidt, “Efficient Estimation of Dynamic Panel Data Models: Alternative Assumptions and Simplified Estimation,” *Journal of Econometrics*, Vol. 76, No. 1-2, 1997, pp. 309-321. [http://dx.doi.org/10.1016/0304-4076\(95\)01793-3](http://dx.doi.org/10.1016/0304-4076(95)01793-3)
- [32] V. Verdier, “Fixed Effects Estimation of Panel Data Models with Sequential Exogeneity,” Michigan State University, Mimeo, 2013 (Paper Presented at the 2013 Econometric Society Australasian Meeting, University of Sydney, Australia).  
As of September 2013,  
[http://econ.msu.edu/seminars/docs/FE\\_SequentialExogeneity.pdf](http://econ.msu.edu/seminars/docs/FE_SequentialExogeneity.pdf)
- [33] F. Windmeijer, “ExpEnd, A Gauss Programme for Non-Linear GMM Estimation of Exponential Models with Endogenous Regressors for Cross Section and Panel Data,” The Institute for Fiscal Studies, Department of Economics, UCL, Cemmap Working Paper, 2002, CWP 14/02.  
As of September 2013,  
<http://www.cemmap.ac.uk/wps/cwp0214.pdf>
- [34] F. Windmeijer, “GMM for Panel Count Data Models,” In: L. Mátyás and P. Sevestre, Eds., *The Econometrics of Panel Data. Fundamentals and Recent Developments in Theory and Practice*, 3rd Edition, Springer, Berlin, Heidelberg, 2008, pp. 603-624. [http://dx.doi.org/10.1007/978-3-540-75892-1\\_18](http://dx.doi.org/10.1007/978-3-540-75892-1_18)
- [35] B. H. Hall and C. Cummins, “TSP 5.0 User’s Guide,” TSP International, 2006.
- [36] J. Bound, D. A. Jaeger and R. M. Baker, “Problems with Instrumental Variables Estimation when the Correlation between the Instruments and the Endogenous Explanatory Variable is Weak,” *Journal of the American Statistical Association*, Vol. 90, No. 430, 1995, pp. 443-450. <http://dx.doi.org/10.1080/01621459.1995.10476536>
- [37] D. Staiger and J. H. Stock, “Instrumental Variables Regression with Weak Instruments,” *Econometrica*, Vol. 65, No. 3, 1997, pp. 557-586. <http://dx.doi.org/10.2307/2171753>
- [38] Y. Kitazawa, “Exponential Regression of Dynamic Panel Data Models,” *Economics Letters*, Vol. 73, No. 1, 2001, pp. 7-13. [http://dx.doi.org/10.1016/S0165-1765\(01\)00467-0](http://dx.doi.org/10.1016/S0165-1765(01)00467-0)
- [39] J. M. Wooldridge, “Distribution-Free Estimation of Some Nonlinear Panel Data Models,” *Journal of Econometrics*, Vol. 90, No. 1, 1999, pp. 77-97. [http://dx.doi.org/10.1016/S0304-4076\(98\)00033-5](http://dx.doi.org/10.1016/S0304-4076(98)00033-5)

## Appendix A

First, Equations (2.2.5) dated  $t-1$  and  $t$  give the following relationships:

$$E[y_{it}u_{i,t-1}] = E[y_{it}\phi_i\mu_{i,t-1}], \quad (\text{A.1})$$

$$E[y_{it}(\mu_{i,t-1}/\mu_{it})u_{it}] = E[y_{it}\phi_i\mu_{i,t-1}], \quad (\text{A.2})$$

while Equations (2.2.6) dated  $t-1$  and  $t$  for  $s < t-1$  give the following relationships:

$$E[u_{is}u_{i,t-1}] = E[\phi_i^2\mu_{is}\mu_{i,t-1} + v_{is}\phi_i\mu_{i,t-1}], \quad (\text{A.3})$$

$$E[u_{is}(\mu_{i,t-1}/\mu_{it})u_{it}] = E[\phi_i^2\mu_{is}\mu_{i,t-1} + v_{is}\phi_i\mu_{i,t-1}]. \quad (\text{A.4})$$

Accordingly, subtracting (A.1) from (A.2), subtracting (A.3) from (A.4), and then taking notice of (2.1.1), the moment conditions (2.2.8) are obtained.

Second, Equations (2.2.6) dated  $t$  for  $s = t-1$  and  $s = t-2$  give the following relationships:

$$E[(\mu_{i,t-2}/\mu_{i,t-1})u_{i,t-1}(1/\mu_{it})u_{it}] = E[\phi_i^2\mu_{i,t-2}], \quad (\text{A.5})$$

$$E[u_{i,t-2}(1/\mu_{it})u_{it}] = E[\phi_i^2\mu_{i,t-2}]. \quad (\text{A.6})$$

Accordingly, subtracting (A.6) from (A.5), the moment conditions (2.2.9) are obtained.

Third, Equations (2.2.7) dated  $t-1$  and  $t$  for  $s \leq t-1$  give the following relationships:

$$E[x_{is}u_{i,t-1}] = E[x_{is}\phi_i\mu_{i,t-1}], \quad (\text{A.7})$$

$$E[x_{is}(\mu_{i,t-1}/\mu_{it})u_{it}] = E[x_{is}\phi_i\mu_{i,t-1}]. \quad (\text{A.8})$$

Accordingly, subtracting (A.7) from (A.8), the moment conditions (2.2.10) are obtained.

## Appendix B

First, Equation (2.3.2) dated  $t-1$  and Equation (2.2.6) dated  $t$  for  $s = t-1$  give the following relationships:

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

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

Accordingly, subtracting (B.1) from (B.2) and then taking notice of (2.1.1) and (2.2.8), the moment conditions (2.3.3) are obtained.

Next, Equation (2.3.2) dated  $t$  and Equation (2.2.6) dated  $t$  for  $s = t-1$  give the following relationships:

$$E[(\mu_{i,t-1}/\mu_{it}^2)(u_{it}^2 - y_{it})] = E[\phi_i^2\mu_{i,t-1}], \quad (\text{B.3})$$

$$E[u_{i,t-1}(1/\mu_{it})u_{it}] = E[\phi_i^2\mu_{i,t-1}]. \quad (\text{B.4})$$

Accordingly, subtracting (B.4) from (B.3), the moment conditions (2.3.4) are obtained.

## Appendix C

First, Equation (2.4.2) dated  $t-1$  and Equation (2.2.6) dated  $t$  for  $s = t-1$  give the following relationships:

$$E\left[\left((u_{i,t-1}^2 - y_{i,t-1}) + (1/\mu_{i,t-1})(u_{i,t-1}^2 - y_{i,t-1}u_{i,t-1})\right)\right] \\ = E[\phi_i^2\mu_{i,t-1}(\mu_{i,t-1} + 1)], \quad (\text{C.1})$$

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

Accordingly, subtracting (C.1) from (C.2) and then taking notice of (2.1.1), (2.2.8) and the following moment conditions based on the transformation proposed by [3]:

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

for  $s = 1, \dots, t-2$ ;  $t = 3, \dots, T$ , which are also valid for the LFM with explanatory variables being predetermined (*i.e.* (2.1.1) and (2.1.2) with (2.1.3)), the moment conditions (2.4.3) are obtained.

Next, equation (2.4.2) dated  $t$  and equation (2.2.6) dated  $t$  for  $s = t-1$  give the following relationships:

$$E\left[\left(\mu_{i,t-1}/(\mu_{it}(\mu_{it} + 1))\right) \cdot \left((u_{it}^2 - y_{it}) + (1/\mu_{it})(u_{it}^2 - y_{it}u_{it})\right)\right] \\ = E[\phi_i^2\mu_{i,t-1}], \quad (\text{C.3})$$

$$E[u_{i,t-1}(1/\mu_{it})u_{it}] = E[\phi_i^2\mu_{i,t-1}]. \quad (\text{C.4})$$

Accordingly, subtracting (C.4) from (C.3), the moment conditions (2.4.4) are obtained.

## Appendix D

First, allowing for the assumptions on the stationarity (*i.e.* (2.5.1) and (2.5.2) with (2.5.3)), Equation (2.2.5) dated  $t$  and Equations (2.2.6) dated  $t$  for  $s = t-2$  and  $s = t-1$  give the following relationships:

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

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

$$E[u_{i,t-1}(1/\mu_{it})u_{it}] = E[\phi_i^2\varphi_i(\beta)]. \quad (\text{D.3})$$

Accordingly, using (D.1) - (D.3) and (2.1.1), the moment conditions (2.5.4) are obtained.

Second, allowing for the assumptions on the stationarity and utilizing the property of the moment generating function, Equations (2.2.7) dated  $t$  for  $s = t-1$  and  $s = t$  give the following relationships:

$$E[x_{i,t-1}(1/\mu_{it})u_{it}] = E[\phi_i\varphi_i'(0)], \quad (\text{D.4})$$

$$E[x_{it}(1/\mu_{it})u_{it}] = E[\phi_i \phi_i'(0)], \quad (\text{D.5})$$

where  $\phi_i'(0) = \partial \phi_i(k) / \partial k|_{k=0}$ . Accordingly, subtracting (D.4) from (D.5), the moment conditions (2.5.5) are obtained.

Third, allowing for the assumptions on the stationarity, Equation (2.3.2) dated  $t$  gives the following relationship:

$$E[(1/\mu_{it})(u_{it}^2 - y_{it})] = E[\phi_i^2 \phi_i(\beta)]. \quad (\text{D.6})$$

Accordingly, using (D.1), (D.3), (D.6), (2.1.1) and (2.5.4), the moment conditions (2.5.6) are obtained.

Fourth, allowing for the assumptions on the stationarity, Equation (2.4.2) dated  $t$  gives the following relationship:

$$\begin{aligned} & E\left[\left(1/(\mu_{it} + 1)\right)\left((u_{it}^2 - y_{it}) + (1/\mu_{it})(u_{it}^2 - y_{it}u_{it})\right)\right] \\ & = E[\phi_i^2 \phi_i(\beta)]. \end{aligned} \quad (\text{D.7})$$

Accordingly, using (D.1), (D.3), (D.7), (2.1.1) and (2.5.4), the moment conditions (2.5.7) are obtained.

## Notes of Tables

1) The setting of values of parameters in the DGP is as follows:  $\gamma = 0.5$ ;  $\beta = 0.5$ ;  $\rho = 0.5$ ;  $\tau = 0.1$ ;  $\sigma_\eta^2 = 0.5$ ;  $\sigma_\varepsilon^2 = 0.5$ . 2) The number of replications is 1000. 3) The instruments used for the GMM estimators are curtailed so that the past levels of dependent and explanatory variables dated  $t-3$  and before (*i.e.*  $y_{is}$  and  $x_{is}$  for  $t = 1, \dots, t-3$ ) are not used for the quasi-differencing transformation dated  $t$ . The curtailment is conducted for the reason of circumventing the exacerbation of the

small sample performance of the GMM estimator due to the excess usage of the weak instruments to be hereinafter described. 4) The instruments sets for GMM estimators include no time dummies. 5) The initial consistent estimates used for the GMM estimation are obtained in same manner as [33]. 6) The symbols ‘‘Sargan’’ and ‘‘df’’ denote the Monte Carlo mean of values of Sargan statistic for each GMM estimator and its degree of freedom, respectively. 7) As for the PSM estimators, the figures in the parentheses next to  $\gamma$  and  $\beta$  imply numbers of the pre-sample periods used for the estimations. 8) 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 fairly small. 9) The values of the Monte Carlo bias and rmse exhibited in the tables are those obtained using the true values of  $\gamma$  and  $\beta$  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 in almost all cases. 10) The individuals where the pre-sample means are zero are eliminated in each replication when estimating the parameters of interest using the PSM estimators. 11) The Monte Carlo means of proportions of zeros for the count dependent variables are about 22% in **Tables 1** and **3** where the DGP is of the Poisson-type, while about 32% in **Table 2** and **4** where the DGP is of the Negbin I-type.