Exploration of dynamic fixed effects logit models from a traditional angle*

Yoshitsugu Kitazawa**

April 30, 2013

Abstract
This paper proposes the transformations for the dynamic fixed effects logit models. Firstly, the transformations construct the valid moment conditions (including the stationarity moment conditions) for the case without explanatory variable. Combining portions of the valid moment conditions gives just the first-order condition of the conditional MLE proposed by Chamberlain (1985). Next, the valid moment conditions are constructed by using the transformations for the case with strictly exogenous continuous explanatory variables, when the number of time periods is greater than or equal to four. This implies that for the dynamic fixed effects logit model with strictly exogenous continuous explanatory variables, the estimators can be constructed which are consistent and asymptotically normal and whose convergence rates equal the inverse of the square root of the cross-sectional sample size. In addition, the small sample properties of the GMM estimators using these moment conditions are investigated by using Monte Carlo experiments.

Keywords: dynamic fixed effects logit models; moment conditions; stationarity; strictly exogenous continuous explanatory variables; root-N consistent estimators; Monte Carlo experiments

JEL classification: C23; C25

1. Introduction

Incorporating dynamics into the binary choice models is one of the issues which attract the interest of econometricians, where the logit specification is often used and the micro datasets are often dealt with (e.g. the analyses on the household brand choice and the female labor force participation, etc.). The dynamics allows for the persistence of an event in past, the logit specification is simple and tractable in terms of the structure, and nowadays the micro datasets are much more accessible than before. In many cases, the micro datasets available have the panel structure where the number of individuals is large but the number of time periods is small.

---

* Discussion Paper Series, Faculty of Economics, Kyushu Sangyo University Discussion Paper, April 2013, No. 60 Errata information: http://www.ip.kyusan-u.ac.jp/J/kitazawa/ERRATA/errata_dfelm.html

** Faculty of Economics, Kyushu Sangyo University, 3-1 Matsukadai 2-chome, Higashi-ku, Fukuoka, 813-8503, Japan. E-mail: kitazawa@ip.kyusan-u.ac.jp

The problem unavoidable in dealing with panel data models is the treatment of the individual heterogeneity. In this paper, the exploration for the panel logit models incorporating dynamics is conducted for the case where the individual heterogeneity is treated as the fixed effect instead of the random effect, since the former treatment is more flexible than the latter one in terms of the model specification. Then, this paper newly proposes the fairly traditional approach solving the incidental parameters problem considered by Neyman and Scott (1948) (which pertains to the fixed effects models) for the dynamic fixed effects logit models. It is shown that this approach gives rise to the (asymptotically normal) root-N consistent estimators (in which the convergence rate equals the inverse of the square root of the cross-sectional sample size) for the dynamic fixed effects logit models. The two types of dynamic fixed effects logit models are explored: those without explanatory variable and with strictly exogenous continuous explanatory variables.

For the dynamic fixed effects logit model without explanatory variable (hereafter the simple dynamic fixed effects logit model), Chamberlain (1985) proposes the (asymptotically normal) root-N consistent estimator.[1] The conditional maximum likelihood estimator (hereafter CMLE) for the simple dynamic fixed effects logit model (which needs four or more time periods) is obtained after ruling out the fixed effects in the manner analogous to that used in obtaining the CMLE proposed by Chamberlain (1980) for the static fixed effects logit model.[2]

In contrast, for the dynamic fixed effects logit model with strictly exogenous continuous explanatory variables, it can be said that no root-N consistent estimator has been proposed until now, although some alternative approaches (solving or alleviating the incidental parameters problem) have been applicable to and/or proposed for this model: First, although the estimator proposed by Honoré and Kyriazidou (2000) as an extension of the CMLE proposed by Chamberlain (1985) for the simple dynamic fixed effects logit model is consistent and asymptotically normal with respect to the cross-sectional size with the number of time periods being fixed, it is not the root-N

[1] See also Hsiao (2003, pp. 211-216), Baltagi (2009, pp. 242-244), and Kyriazidou (2010), etc. on this issue.

[2] The genesis of the CMLE for the static fixed effects logit model is Rasch (1960, 1961). The first-order condition of the CMLE for the static fixed effects logit model is also derived by Bonhomme (2012) and Kitazawa (2012) in separate ways.

consistent estimator by reason of using the kernel weight.[3] Second, although the bias-corrected estimators proposed by Carro (2007), Bester and Hansen (2009), Fernández-Val (2009), Hahn and Kuersteiner (2011), and Yu et al. (2012) aim at obtaining as unbiased estimators as possible for the moderately large number of time periods, they are never the root-N consistent estimators.[4] Third, although the pseudo CMLE proposed by Bartolucci and Nigro (2012) is the root-N consistent estimator for the pseudo true values of the parameters of interest instead of the true values of the parameters of interest, which is created by using the approximation of the dynamic fixed effects logit model which is a modified version of the quadratic exponential model of Bartolucci and Nigro (2010), it is adamantly stated that this estimator falls into the category of the approximation estimator instead of the root-N consistent estimator for the true dynamic fixed effects logit model with strictly exogenous continuous explanatory variables.

Different from the methods proposed until now, in this paper, the incidental parameters problems in the dynamic fixed effects logit models are solved by eliminating the fixed effects after the models are transformed in order that the expressions including the fixed effects are separated out as the additive terms. Eliminating the fixed effects gives the valid moment conditions for constructing the root-N consistent estimators for the dynamic fixed effects logit models.

The valid moment conditions for the simple dynamic fixed effects logit model are derived in the following manner: First, the model is transformed into the simple linear panel data models with additive fixed effects. Next, the error-components structures holding between the logit model and the transformed linear panel data models give the valid moment conditions (including the stationarity moment conditions), by using the methodology analogous to that proposed by Ahn (1990) and Ahn and Schmidt (1995) for the simple ordinary dynamic panel data model. In addition, it is shown that the first-order condition of the CMLE proposed by Chamberlain (1985) can be rewritten as the combinations of some of these moment conditions.

Likewise, the derivation of the valid moment conditions for the dynamic fixed

[3] Chintagunta et al. (2001) apply the estimator proposed by Honoré and Kyriazidou (2000) to the household brand choice model on the yogurt purchases.

[4] In Fernández-Val (2009), the analysis on the female labor force participation is conducted by using various bias-corrected estimators.

effects logit model with strictly exogenous continuous explanatory variables is as follows: First, the model is transformed in order that the logit probabilities composed of the fixed effects and the explanatory variables are separated out as the additive terms. Next, the valid moment conditions, which need four or more time periods, are obtained by applying a variety of the hyperbolic tangent differencing transformation (hereafter HTD transformation) proposed by Kitazawa (2012) for the static fixed effects logit model to the transformed forms of the model. This implies that the root-N consistent estimators are dug up for the dynamic fixed effects logit model with strictly exogenous continuous explanatory variables.

The generalized method of moments estimator (hereafter GMM estimator) proposed by Hansen (1982) provides the dynamic fixed effects logit models with the root-N consistent estimators. For the simple dynamic fixed effects logit model, it is recognized that there are the root-N consistent estimators other than the CMLE proposed by Chamberlain (1985), while for the dynamic fixed effects logit model with strictly exogenous continuous explanatory variables, the presence of the root-N consistent estimators is manifested in the case of four or more time periods.

Now, a crack is opened into the sense of stagnation in which the recent researches on the dynamic fixed effects logit model with strictly exogenous continuous explanatory variables seem to be packed in the atmosphere of relinquishing the pursuit of the root-N consistent estimators. The crack is opened by dint of the extremely traditional reaction. It seems reasonable to say that Hahn's (2001) suggestion is no longer applicable to the case of four or more time periods, in which it is stated that the root-N consistent estimation is infeasible in more general specifications in the dynamic fixed effects logit model and accordingly the substantial improvement over the estimator proposed by Honoré and Kyriazidou (2000) is unlikely.[5] To the best of author's knowledge, it can be said that the first attainment is conducted in this paper, in which the root-N consistent estimators are contrived for the dynamic fixed effects logit model with strictly exogenous continuous explanatory variables.

Some Monte Carlo experiments not only investigate the small sample behaviors of the GMM estimators proposed in this paper for the dynamic fixed effects logit models

[5] It is thought that Hahn (2001) discusses the infeasibility of the root-N consistent estimator under a special initial condition for three time periods.

but also reinforce the corroboration of the root-N consistency of the GMM estimators.

The rest of the paper is as follows. In section 2, both dynamic fixed effects logit models are presented: those without explanatory variable and with the strictly exogenous continuous explanatory variable and then the root-N consistent estimators (i.e. the GMM estimators) are constructed using the valid moment conditions for both models. In section3, some Monte Carlo results are reported for the GMM estimators. Section 4 concludes.

2. Models and estimations

In this section, the root-N consistent estimators are constructed for the dynamic fixed effects logit models. Firstly, the process for constructing the valid moment conditions is exhibited for the model without explanatory variable, where the dynamic fixed effects logit model is transformed into the linear panel data models and then the methodology analogous to that proposed by Ahn (1990) and Ahn and Schmidt (1995) for the ordinary simple dynamic panel data model is applied to the error-components structures holding between the simple dynamic fixed effects logit model and the transformed linear panel data models. As a matter of course, the stationarity moment conditions are proposed for the simple dynamic fixed effects logit model. In addition, it is shown that the first-order condition of the conditional MLE proposed by Chamberlain (1985) can be assembled by using these moment conditions. Secondly, the extension of the above-mentioned transformations is applied to the model with the strictly exogenous continuous explanatory variable in order to construct the valid moment conditions by using the methodology analogous to that proposed by Kitazawa (2012) for the static fixed effects logit model. This will be very intriguing, because the root-N consistent estimation can be achieved for the dynamic fixed effects logit model with strictly exogenous continuous explanatory variables. Thirdly, one type of the root-N consistent estimators (i.e. the GMM estimator) is introduced, which uses the valid moment conditions mentioned above.

Throughout the paper, subscripts i and t denote the individual and the time period, respectively. It is assumed that the number of individuals $N \to \infty$, while the number of time periods T is fixed.

2.1. Simple model and transformations

The binary dependent variable y_{it} for the simple dynamic fixed effects logit model is specified as follows:

$$y_{it} = p(\eta_i, y_{i,t-1}) + v_{it}, \qquad \text{for } 2 \le t \le T, \tag{2.1.1}$$

$$\text{with} \qquad \mathrm{E}[v_{it} \mid \eta_i, y_{i1}, v_i^{t-1}] = 0, \tag{2.1.2}$$

where $p(\eta_i, y_{i,t-1}) = \exp(\eta_i + \gamma y_{i,t-1}) / (1 + \exp(\eta_i + \gamma y_{i,t-1}))$ with η_i and γ being the fixed effect and the parameter of interest respectively, v_{it} is the disturbance, y_{i1} is the initial value of the binary dependent variable, and $v_i^{t-1} = (v_{i1}, \ldots, v_{i,t-1})$ with v_{i1} being empty.[6]

The logit probability $p(\eta_i, y_{i,t-1})$ with which $y_{it} = 1$ can be also written in the following form:

$$p(\eta_i, y_{i,t-1}) = f(\eta_i) y_{i,t-1} + g(\eta_i), \tag{2.1.3}$$

$$\text{with} \qquad f(\eta_i) = h(\eta_i) - g(\eta_i), \tag{2.1.4}$$

$$g(\eta_i) = \exp(\eta_i) / (1 + \exp(\eta_i)), \tag{2.1.5}$$

$$h(\eta_i) = \exp(\eta_i + \gamma) / (1 + \exp(\eta_i + \gamma)). \tag{2.1.6}$$

The form (2.1.3) with (2.1.4) - (2.1.6) implies the logit specification of the linear AR(1) (autoregressive model of order 1) regression form considered by Al-Sadoon et al. (2012) for the dynamic binary choice panel data model with fixed effects.[7]

Based on the facts above, the simple dynamic fixed effects logit model is transformed into the following two types of simple panel data models with additive fixed effects:

$$y_{it} = \delta y_{i,t-1}(1 - y_{it}) y_{i,t+1} + g(\eta_i) + w_{it}, \qquad \text{for } 2 \le t \le T-1, \tag{2.1.7}$$

$$\text{with} \qquad \mathrm{E}[w_{it} \mid \eta_i, y_{i1}, v_i^{t-1}] = 0, \tag{2.1.8}$$

and

$$y_{it} = -\delta(1 - y_{i,t-1}) y_{it} (1 - y_{i,t+1}) + h(\eta_i) + \omega_{it}, \quad \text{for } 2 \le t \le T-1, \tag{2.1.9}$$

$$\text{with} \qquad \mathrm{E}[\omega_{it} \mid \eta_i, y_{i1}, v_i^{t-1}] = 0, \tag{2.1.10}$$

where $\delta = \exp(\gamma) - 1$ and then both of the forms (2.1.7) and (2.1.9) are linear with

[6] This type of description is used by Kitazawa (2012) for the static fixed effects logit model and by Blundell et al. (2002) for count panel data model.

[7] Al-Sadoon et al. (2012) propose the root-N consistent estimators for the exponential distribution specification of the linear AR(1) regression form for the dynamic binary choice panel data models. The origin of the regression form is Pesaran and Timermann (2009).

respect to δ, although δ is inevitably above -1. The former separates out $g(\eta_i)$ as the additive fixed effect, while the latter separates out $h(\eta_i)$. The newly defined disturbances w_{it} and ω_{it} satisfy the conditional moment restrictions (2.1.8) and (2.1.10), respectively. In this paper, the transformations (2.1.7) and (2.1.9) are referred to as the "g-form" and "h-form" respectively. The derivations of (2.1.7) with (2.1.8) and (2.1.9) with (2.1.10) are shown in Appendix A.

2.2. Mean and covariance restrictions and moment conditions

In this subsection, the moment conditions are constructed by utilizing the relationships between the disturbances in the original dynamic fixed effects logit specification and those in its transformations, based on the conditional moment restrictions in the transformations. The methodology used is analogous to that Ahn (1990) and Ahn and Schmidt (1995) for the ordinary simple dynamic panel data model.

Firstly, the moment conditions based on the g-form are derived. The conditional moment restrictions (2.1.8) give the following mean and covariance restrictions:

$$\mathrm{E}[w_{it}] = 0, \qquad \text{for } 2 \le t \le T-1, \tag{2.2.1}$$

$$\mathrm{E}[y_{i1} w_{it}] = 0, \qquad \text{for } 2 \le t \le T-1, \tag{2.2.2}$$

$$\mathrm{E}[v_{is} w_{it}] = 0, \qquad \text{for } 2 \le s \le t-1;\ 3 \le t \le T-1, \tag{2.2.3}$$

By using the relationships after replacing the unobservable variables w_{it} and v_{is} with the observable variables $u_{it} = y_{it} - \delta y_{i,t-1}(1 - y_{it}) y_{i,t+1}$ and y_{is} respectively, the following $T-3$ and $(T-2)(T-3)/2$ unconditional moment conditions for estimating γ consistently are obtained:

$$\mathrm{E}[\Delta u_{it}] = 0, \qquad \text{for } 3 \le t \le T-1, \tag{2.2.4}$$

$$\mathrm{E}[y_{is} \Delta u_{it}] = 0, \qquad \text{for } 1 \le s \le t-2;\ 3 \le t \le T-1, \tag{2.2.5}$$

where Δ is the first-differencing operator such that $\Delta u_{it} = u_{it} - u_{i,t-1}$.[8]

Next, the moment conditions based on the h-form are derived in the same manner as that for the g-form. The conditional moment restrictions (2.1.9) give the following mean and covariance restrictions:

$$\mathrm{E}[\omega_{it}] = 0, \qquad \text{for } 2 \le t \le T-1, \tag{2.2.6}$$

$$\mathrm{E}[y_{i1} \omega_{it}] = 0, \qquad \text{for } 2 \le t \le T-1, \tag{2.2.7}$$

[8] In this paper, the observable variable is defined as the variable constructed using data and parameters of interest, as is similar to that in Ahn (1990) and Ahn and Schmidt (1995).

$$\mathrm{E}[v_{is}\omega_{it}] = 0, \qquad \text{for } 2 \le s \le t-1;\ 3 \le t \le T-1, \tag{2.2.8}$$

By using the relationships after replacing the unobservable variables ω_{it} and v_{is} with the observable variables $\upsilon_{it} = y_{it} + \delta(1-y_{i,t-1})y_{it}(1-y_{i,t+1})$ and y_{is} respectively, the following $T-3$ and $(T-2)(T-3)/2$ unconditional moment conditions for estimating γ consistently are obtained:

$$\mathrm{E}[\Delta\upsilon_{it}] = 0, \qquad \text{for } 3 \le t \le T-1, \tag{2.2.9}$$

$$\mathrm{E}[y_{is}\Delta\upsilon_{it}] = 0, \qquad \text{for } 1 \le s \le t-2;\ 3 \le t \le T-1, \tag{2.2.10}$$

The derivations of the moment conditions (2.2.4) and (2.2.5) based on the g-form and the moment conditions (2.2.9) and (2.2.10) based on the h-form are shown in Appendix B.

The moment conditions (2.2.4) and (2.2.5) for the g-form and (2.2.9) and (2.2.10) for the h-form are linear with respect to δ, implying that the linear estimations for δ can be conducted by using these moment conditions.

It might be said that the moment conditions (2.2.5) and (2.2.10) correspond to the standard moment conditions in the ordinary dynamic panel data model, which are proposed by Holtz-Eakin et al. (1985) and Arellano and Bond (1991), while the moment conditions corresponding to the additional non-linear moment conditions proposed by Ahn (1990) and Ahn and Schmidt (1995) are the redundancies (see Appendix B).

In this paper, the moment conditions (2.2.5) and (2.2.10) are referred to as the standard moment conditions based on the g-form and h-form for the simple dynamic fixed effects logit model, respectively.

2.3. Stationarity in the simple dynamic fixed effects logit model

It is recognizable that the stationary state can be defined easily in the simple dynamic fixed effects logit model, paying notice to the form of the logit probability (2.1.3) with (2.1.4) - (2.1.6). When the initial condition of the dynamic fixed effects logit model (2.1.1) is written as

$$y_{i1} = g(\eta_i)/(1-f(\eta_i)) + v_{i1}, \tag{2.3.1}$$

$$\text{with} \qquad \mathrm{E}[v_{i1} \mid \eta_i] = 0, \tag{2.3.2}$$

the binary dependent variable y_{it} is stationary:

$$\mathrm{E}[y_{it} \mid \eta_i] = g(\eta_i)/(1-f(\eta_i)), \tag{2.3.3}$$

which is a probability.[9]

In this case, two types of the $T-3$ stationarity moment conditions are constructed: those based on the g-form are

$$\mathrm{E}[\Delta y_{i,t-1}\, u_{it}] = 0\,, \qquad \text{for } 3 \le t \le T-1, \tag{2.3.4}$$

while those based on the h-form are

$$\mathrm{E}[\Delta y_{i,t-1}\, \upsilon_{it}] = 0\,, \qquad \text{for } 3 \le t \le T-1. \tag{2.3.5}$$

The derivation of the moment conditions (2.3.4) and (2.3.5) is shown in Appendix C.

It might be said that the moment conditions (2.3.4) and (2.3.5) correspond to the stationarity moment conditions in the ordinary dynamic panel data model, which are proposed by Arellano and Bover (1995) and discussed in Ahn and Schmidt (1995) and Blundell and Bond (1998).

2.4. Relationship with the CMLE proposed by Chamberlain (1985)

Based on the sequential four time periods (i.e. $t-2$, $t-1$, t and $t+1$), the CMLE proposed by Chamberlain (1985) for the simple dynamic fixed effects logit model is obtained by maximizing the following objective function with respect to γ:

$$\sum_{i=1}^{N} \ell_{it}\,, \tag{2.4.1}$$

with

$$\ell_{it} = (\Delta y_{it})^2(\gamma y_{i,t-1}(y_{i,t-2} - y_{i,t+1}) - \ln(1+\exp(\gamma(y_{i,t-2} - y_{i,t+1}))))\,.\text{[10]} \tag{2.4.2}$$

The CMLE for γ, which needs four or more time periods as is seen from (2.4.1) with (2.4.2), is a root-N consistent estimator. The detail on this estimator is shown in Hsiao (2003, pp. 211-216). For this case of four time periods, it is corroborated that the CMLE is asymptotically efficient under the condition that $(\Delta y_{it})^2 = 1$ and $(y_{i,t-2} - y_{i,t+1})^2 = 1$ (see Appendix D).[11]

It is of interest that the first-order condition for (2.4.1) with (2.4.2) can be written

---

[9] It is assumed that v_{it} is not empty when the stationarity is imposed on the dependent variable y_{it}.

[10] The maximization problem is written referring to Hsiao (2003, pp. 211-216), Baltagi (2009, pp. 242-244), and Kyriazidou (2010).

[11] Alternatively, according to Wooldridge (2011), it can be also said that the CMLE is asymptotically efficient in the class of estimators putting no assumption between the initial conditions on the dependent variables and the fixed effects.

as the following plain sum of the moment conditions (2.2.4), (2.2.5) for $s = t-2$ multiplied by (-1), and (2.2.10) for $s = t-2$ multiplied by (-1):

$$\mathrm{E}[(1-y_{i,t-2})\Delta u_{it} - y_{i,t-2}\Delta \upsilon_{it}] = 0 . \quad (2.4.3)$$

Further, if the binary dependent variable y_{it} is stationary as is specified in previous subsection, the first-order condition (2.4.3) can be rewritten as the following plain sum of the stationarity moment conditions (2.3.4) and (2.3.5):

$$\mathrm{E}[y_{i,t-1}(\Delta u_{it} + \Delta \upsilon_{it})] = 0 . \quad (2.4.4)$$

These imply that the first-order condition of the CMLE proposed by Chamberlain (1985) for the simple dynamic fixed effects logit model can be written as the sums of the moment conditions based on the g-form and h-form.[12] The proof is shown in Appendix E.

2.5. Extension to the model with the strictly exogenous continuous explanatory variable

In this subsection, the model is extended to that accompanied by the strictly exogenous continuous explanatory variable. The discussion reaches the most interesting part, which is the region previously-untrodden by researchers to the best of author's knowledge. It will be seen that the root-N consistent estimators are present for the dynamic fixed effects logit model with strictly exogenous continuous explanatory variables. The valid moment conditions are presented, which construct the (asymptotically normal) root-N consistent estimators, such as the GMM estimator.

The binary dependent variable y_{it} for the dynamic fixed effects logit model with the strictly exogenous continuous explanatory variable x_{it} is specified as follows:

$$y_{it} = p(\eta_i, y_{i,t-1}, x_{it}) + v_{it}, \qquad \text{for } 2 \le t \le T , \quad (2.5.1)$$

$$\text{with} \qquad \mathrm{E}[v_{it} \mid \eta_i, y_{i1}, v_i^{t-1}, x_i^T] = 0 , \quad (2.5.2)$$

where $p(\eta_i, y_{i,t-1}, x_{it}) = \exp(\eta_i + \gamma y_{i,t-1} + \beta x_{it}) / (1 + \exp(\eta_i + \gamma y_{i,t-1} + \beta x_{it}))$ with η_i being the fixed effect and γ and β being the parameters of interest, y_{i1} is the initial value of the binary dependent variable, $v_i^{t-1} = (v_{i1}, \ldots, v_{i,t-1})$ with v_{i1} being empty, and $x_i^T = (x_{i1}, \ldots, x_{iT})$.

As is similar to the simple model discussed in previous subsections, the logit

[12] In the old versions of Buchinsky et al. (2010) and Bonhomme (2012), the first-order condition of the CMLE for the simple dynamic fixed effects logit model is derived under the setting considered in Hahn (2001) for three periods, by using the methods different from that proposed in this paper. The former is related to Johnson's (2004) results.

probability $p(\eta_i, y_{i,t-1}, x_{it})$ with which $y_{it}=1$ can be also written in the following form:

$$p(\eta_i, y_{i,t-1}, x_{it}) = f(\eta_i, x_{it}) y_{i,t-1} + g(\eta_i, x_{it}), \tag{2.5.3}$$

with $$f(\eta_i, x_{it}) = h(\eta_i, x_{it}) - g(\eta_i, x_{it}), \tag{2.5.4}$$

$$g(\eta_i, x_{it}) = \exp(\eta_i + \beta x_{it}) / (1 + \exp(\eta_i + \beta x_{it})), \tag{2.5.5}$$

$$h(\eta_i, x_{it}) = \exp(\eta_i + \gamma + \beta x_{it}) / (1 + \exp(\eta_i + \gamma + \beta x_{it})), \tag{2.5.6}$$

Different from the case for the simple dynamic fixed effects logit model, $g(\eta_i, x_{i,t+1})$ and $h(\eta_i, x_{i,t+1})$ are separated out in order that equation (2.5.1) is transformed into the following g-form and h-form for the dynamic fixed effects logit model with the strictly exogenous continuous explanatory variable, respectively:

$$U_{it} = g(\eta_i, x_{i,t+1}) + W_{it}, \qquad \text{for } 2 \le t \le T-1, \tag{2.5.7}$$

with $$\mathrm{E}[W_{it} \mid \eta_i, y_{i1}, v_i^{t-1}, x_i^T] = 0, \tag{2.5.8}$$

$$\begin{aligned} U_{it} = {} & y_{it} + (1-y_{it}) y_{i,t+1} - (1-y_{it}) y_{i,t+1} \exp(-\beta \Delta x_{i,t+1}) \\ & - \delta y_{i,t-1} (1-y_{it}) y_{i,t+1} \exp(-\beta \Delta x_{i,t+1}), \end{aligned} \tag{2.5.9}$$

and

$$\Upsilon_{it} = h(\eta_i, x_{i,t+1}) + \Omega_{it}, \qquad \text{for } 2 \le t \le T-1, \tag{2.5.10}$$

with $$\mathrm{E}[\Omega_{it} \mid \eta_i, y_{i1}, v_i^{t-1}, x_i^T] = 0, \tag{2.5.11}$$

$$\begin{aligned} \Upsilon_{it} = {} & y_{it} y_{i,t+1} + y_{it} (1-y_{i,t+1}) \exp(\beta \Delta x_{i,t+1}) \\ & + \delta (1-y_{i,t-1}) y_{it} (1-y_{i,t+1}) \exp(\beta \Delta x_{i,t+1}), \end{aligned} \tag{2.5.12}$$

where $\delta = \exp(\gamma) - 1$. The newly defined disturbances W_{it} and Ω_{it} satisfy the conditional moment restrictions (2.5.8) and (2.5.11), respectively. The derivations of the g-form (2.5.7) with (2.5.8) and (2.5.9) and h-form (2.5.10) with (2.5.11) and (2.5.12) are shown in Appendix F and G, respectively.

The valid moment conditions are obtained by eliminating the fixed effect η_i from the g-form and h-form. By utilizing the relationship between the hyperbolic tangent function and the logit probability (i.e. a variety of the HTD transformation), which is presented by Kitazawa (2012) with the aim of obtaining the valid moment conditions for the static fixed effects logit model, the conditional moment conditions for estimating γ and β consistently are obtained on the basis of the g-form and h-form as follows:

$$\mathrm{E}[\hbar U_{it} \mid \eta_i, y_{i1}, v_i^{t-2}, x_i^T] = 0, \qquad \text{for } 3 \le t \le T-1, \tag{2.5.13}$$

with

$$\hbar U_{it} = U_{it} - y_{i,t-1} - \tanh((1/2)(-\gamma y_{i,t-2} + \beta(\Delta x_{it} + \Delta x_{i,t+1})))(U_{it} + y_{i,t-1} - 2U_{it}\, y_{i,t-1}), \tag{2.5.14}$$

and

$$\mathrm{E}[\hbar \Upsilon_{it} \mid \eta_i, y_{i1}, v_i^{t-2}, x_i^T] = 0, \qquad \text{for } 3 \le t \le T-1, \tag{2.5.15}$$

with

$$\hbar \Upsilon_{it} = \Upsilon_{it} - y_{i,t-1} - \tanh((1/2)(\gamma(1 - y_{i,t-2}) + \beta(\Delta x_{it} + \Delta x_{i,t+1})))(\Upsilon_{it} + y_{i,t-1} - 2\Upsilon_{it}\, y_{i,t-1}). \tag{2.5.16}$$

The derivations of the moment conditions (2.5.13) with (2.5.14) (which are based on the g-form) and the moment conditions (2.5.15) with (2.5.16) (which are based on the h-form) are shown in Appendix H.

What the moment conditions (2.5.13) with (2.5.14) and the moment conditions (2.5.15) with (2.5.16) make clear at once is that the root-N consistent estimators can be constructed for the dynamic fixed effects logit model with strictly exogenous continuous explanatory variables, when the number of time periods is greater than or equal to four.[13]

2.6. Root-N consistent estimators using the valid moment conditions

This subsection reviews one type of the root-N consistent estimators using the valid moment conditions proposed in previous subsections for the dynamic fixed effects logit models. The GMM estimator proposed by Hansen (1982) is obtained by minimizing the quadratic form comprised of the sample analogues of the moment conditions and a weighting matrix.

Any set of the valid unconditional moment conditions for the dynamic fixed effects logit models can be collectively written in the following $m \times 1$ vector form:

[13] For the dynamic fixed effects logit model with strictly exogenous continuous explanatory variables, author furthermore expects the presence of the root-N consistent CMLE which will be written as the combination of the moment conditions based on the g-form and h-form (i.e. (2.5.13) with (2.5.14) and (2.5.15) with (2.5.16)) as is the case for the model without explanatory variable.

$$\mathrm{E}[\varphi_i(\theta)] = 0\,, \qquad (2.6.1)$$

where m is the number of the moment conditions and $\theta = \gamma$ for the simple dynamic fixed logit model (i.e. (2.1.1) with (2.1.2)), while $\theta = (\gamma\ \beta)'$ for the dynamic fixed effects logit model with the strictly exogenous continuous explanatory variable (i.e. (2.5.1) with (2.5.2)).

The optimal GMM estimator $\hat{\theta}_{\mathrm{GMM}}$, which is based on the moment conditions (2.6.1), is obtained by minimizing the following quadratic form with respective to θ:

$$\bar{\varphi}(\theta)'\left(\bar{\Theta}(\hat{\theta}_1)\right)^{-1}\bar{\varphi}(\theta)\,, \qquad (2.6.2)$$

with $$\bar{\varphi}(\theta) = (1/N)\sum_{i=1}^{N}\varphi_i(\theta)\,, \qquad (2.6.3)$$

$$\bar{\Theta}(\hat{\theta}_1) = (1/N)\sum_{i=1}^{N}\varphi_i(\hat{\theta}_1)\varphi_i(\hat{\theta}_1)'\,, \qquad (2.6.4)$$

where $\hat{\theta}_1$ is any consistent estimator for θ. It is well-known that the following relationship holds for the optimal GMM estimator:

$$N^{1/2}(\hat{\theta}_{\mathrm{GMM}} - \theta_0) \xrightarrow{d} \mathrm{N}(0,\ (D(\theta_0)'(\Theta(\theta_0))^{-1}D(\theta_0))^{-1})\,, \qquad (2.6.5)$$

where $D(\theta_0) = (\partial\,\mathrm{E}[\varphi_i(\theta)]/\partial\theta')\,|_{\theta=\theta_0}$ and $\Theta(\theta_0) = \mathrm{E}[\varphi_i(\theta_0)\varphi_i(\theta_0)']$, with θ_0 being the true value of θ. The relationship (2.6.5) is a representation of the (asymptotically normal) root-N consistent estimator.

For the simple dynamic fixed effects logit model (i.e. (2.1.1) with (2.1.2)), the two types of GMM estimators are constructed on the basis of the g-form: the GMM(g-STD) estimator using the moment conditions (2.2.4) and (2.2.5) and the GMM(g-SYS) estimator using (2.2.4), (2.2.5) and (2.3.4), while those are constructed on the basis of the h-form as well: the GMM(h-STD) estimator using the moment conditions (2.2.9) and (2.2.10) and the GMM(h-SYS) estimator using (2.2.9), (2.2.10) and (2.3.5). In addition, the GMM(FOC-o) and GMM(FOC-s) estimators are defined, which uses the moment conditions (2.4.3) and (2.4.4) for $3 \le t \le T-1$, respectively.

The GMM(g-STD) and GMM(h-STD) estimators and the GMM(g-SYS) and GMM(h-SYS) estimators correspond to the GMM estimator using the standard moment conditions only and that using both of the standard and stationarity moment conditions, in the framework of the ordinary dynamic panel data model, respectively. The moment conditions used in the GMM(FOC-o) estimator are the first-order conditions of the

CMLE for the simple dynamic fixed effects logit model without assuming the stationary dependent variable, while the moment conditions used in the GMM(FOC-s) estimator are those assuming the stationary dependent variable.

It should be noted that since all the moment conditions used in the GMM estimators defined above are linear with respect to δ and $\gamma = \ln(\delta + 1)$, the parameter of interest γ can be estimated without using the non-linear optimization.

For the dynamic fixed effects logit model with the strictly exogenous explanatory continuous variable (i.e. (2.5.1) with (2.5.2)), the GMM estimators are constructed by using the valid unconditional moment conditions generated from the valid conditional moment conditions (2.5.13) with (2.5.14) based on the g-form and (2.5.15) with (2.5.16) based on the h-form. Accordingly, it is claimed that the discovery of the root-N consistent estimators is conducted for the dynamic fixed effects logit model with strictly exogenous continuous explanatory variables.

3. Monte Carlo

In this section, the finite sample behaviors of the root-N consistent estimators for the dynamic fixed effects logit models with no explanatory variable and with the strictly exogenous continuous explanatory variable are investigated by using some Monte Carlo experiments. In the data generating processes (hereafter, DGP) for both models, the dependent variables are designed to be stationary. The experiments are implemented by using the econometric software TSP version 5.1 (see Hall and Cummins, 2009). Another objective of the Monte Carlo experiments is to score the insurance goal which puts on a firm basis the presence of the root-N consistent estimators for the dynamic fixed effects logit model with strictly exogenous continuous explanatory variables.

3.1. Model without explanatory variable

The DGP is as follows:

$$y_{it} = \begin{cases} 1 & \text{if } p(\eta_i, y_{i,t-1}) > \zeta_{it} \\ 0 & \text{otherwise} \end{cases},$$

$$y_{i1} = \begin{cases} 1 & \text{if } q(\eta_i) > \zeta_{i1} \\ 0 & \text{otherwise} \end{cases},$$

$$p(\eta_i, y_{i,t-1}) = \exp(\eta_i + \gamma y_{i,t-1}) / (1 + \exp(\eta_i + \gamma y_{i,t-1})),$$

$$q(\eta_i) = \frac{1}{1 + (1 + \exp(\eta_i)) / (\exp(\eta_i)(1 + \exp(\eta_i + \gamma)))},$$

$$\zeta_{it} \sim \mathrm{U}(0,1); \;\; \eta_i \sim \mathrm{N}(0, \sigma_\eta^2).$$

In the DGP, values are set to the parameters γ and σ_η^2. The experiments are carried out with cross-sectional sizes $N = 1000$, 5000 and 10000, numbers of time periods $T = 4$ and 8, and number of replications $R_N = 10000$.

Table 1 and 2 are the illustrations of the Monte Carlo experiments on the root-N consistent estimators mentioned in previous section for the simple dynamic fixed effects logit model when $T = 4$ and 8, respectively.

The size alleviations of bias and rmse (root mean squared error) for the all GMM estimators are found as N increases, which are the reflection of the root-N consistency of the GMM estimators.

The downward biases for the GMM(g-STD), GMM(h-STD), GMM(g-SYS), and GMM(h-SYS) estimators are discernible, especially for the high values of the persistence parameter γ and the variance σ_η^2 which generates the fixed effects. These GMM estimators are presumably afflicted with the weak instruments problem studied by Bound et al. (1995) and Staiger and Stock (1997), which results from the usage of the standard moment conditions (2.2.5) and (2.2.10) employing the lagged dependent variables dated $t-2$ and before as the instruments for the g-form and h-form dated t, respectively.

For the more persistent γ and the smaller N (i.e. $\gamma = 2.5$ and $N = 1000$), the GMM(g-STD), GMM(h-STD), GMM(g-SYS), and GMM(h-SYS) estimators when $T = 8$ behave worse than those when $T = 4$. This is probably due to the excessive use of the standard moment conditions when $T = 8$, where the weak instruments problem is salient.

The weak instruments problem is frequently seen in analyses using the ordinary dynamic panel data model and the count panel data model (see Blundell and Bond, 1998, and Blundell et al., 2002, etc.).

As is similar to the Monte Carlo experiments carried out by Blundell and Bond (1998) for the ordinary dynamic panel data model, it can be said that the additional usage of the stationarity moment conditions improves the small sample performances of the GMM estimators, especially for the high value of the persistence parameter γ, as long as comparing the results of the GMM(g-STD) and GMM(h-STD) estimators with those of the GMM(g-SYS) and GMM(h-SYS) estimators respectively. However, the dramatic improvement in terms of bias and rmse for the high value of the persistence parameter γ is conducted by using the GMM(FOC-o) estimator which uses the first-order conditions of the CMLE written as the plain sums of fractions of the moment conditions used mainly in the GMM(g-STD) and GMM(h-STD) estimators.

Although it cannot be said that the GMM(FOC-s) estimator, which uses the first-order conditions of the CMLE written using the plain sums of the stationarity moment conditions, behaves well for the low value of the persistent parameter γ and the large value of the variance σ_η^2 which is associated with the dispersed fixed effects, it behaves well for the high value of the persistence parameter γ, which is comparable to the GMM(FOC-o) estimator in terms of bias and rmse.

3.2. Model with the strictly exogenous continuous explanatory variable

The DGP is as follows:

$$y_{it} = \begin{cases} 1 & \text{if } p(\eta_i, y_{i,t-1}, x_{it}) > \zeta_{it} \\ 0 & \text{otherwise} \end{cases},$$

$$y_{i1} = \begin{cases} 1 & \text{if } q(\eta_i, x_{i1}) > \zeta_{i1} \\ 0 & \text{otherwise} \end{cases},$$

$$p(\eta_i, y_{i,t-1}, x_{it}) = \exp(\eta_i + \gamma y_{i,t-1} + \beta x_{it}) / (1 + \exp(\eta_i + \gamma y_{i,t-1} + \beta x_{it})),$$

$$q(\eta_i, x_{i1}) = \frac{1}{1 + (1 + \exp(\eta_i + \beta x_{i1})) / (\exp(\eta_i + \beta x_{i1})(1 + \exp(\eta_i + \gamma + \beta x_{i1})))},$$

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

$$x_{i1} = (\tau / (1 - \rho))\eta_i + (1 / (1 - \rho^2)^{1/2})\varepsilon_{i1},$$

$$\zeta_{it} \sim \mathrm{U}(0,1);\ \ \eta_i \sim \mathrm{N}(0,\ \sigma_\eta^2);\ \ \varepsilon_{it} \sim \mathrm{N}(0,\sigma_\varepsilon^2).$$

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 cross-sectional sizes $N = 1000$, 5000 and 10000, numbers of time periods $T = 4$ and 8, and number of replications $R_N = 10000$.

The small sample properties are investigated for the root-*N* consistent GMM estimators proposed in previous section for the dynamic fixed effects logit model with the strictly exogenous continuous explanatory variable (i.e. (2.5.1) with (2.5.2)). In the Monte Carlo experiments carried out in this subsection, the GMM estimator based on the g-form, which is referred to as the GMM(g-HTD), uses the following $T-3$, $(T-2)(T-3)/2$, and $3(T-3)$ unconditional moment conditions constructed from the conditional moment conditions (2.5.13) with (2.5.14):

$$\mathrm{E}[\hbar U_{it}] = 0, \qquad \text{for } 3 \le t \le T-1, \qquad (3.2.1)$$

$$\mathrm{E}[y_{is}\, \hbar U_{it}] = 0, \qquad \text{for } 1 \le s \le t-2;\ 3 \le t \le T-1, \qquad (3.2.2)$$

$$\mathrm{E}[\Delta x_{is}\, \hbar U_{it}] = 0, \qquad \text{for } t-1 \le s \le t+1;\ 3 \le t \le T-1, \qquad (3.2.3)$$

while that based on the h-form, which is referred to as the GMM(h-HTD), uses the following $T-3$, $(T-2)(T-3)/2$, and $3(T-3)$ unconditional moment conditions constructed from the conditional moment conditions (2.5.15) with (2.5.16):

$$\mathrm{E}[\hbar \Upsilon_{it}] = 0, \qquad \text{for } 3 \le t \le T-1, \qquad (3.2.4)$$

$$\mathrm{E}[y_{is}\, \hbar \Upsilon_{it}] = 0, \qquad \text{for } 1 \le s \le t-2;\ 3 \le t \le T-1, \qquad (3.2.5)$$

$$\mathrm{E}[\Delta x_{is}\, \hbar \Upsilon_{it}] = 0, \qquad \text{for } t-1 \le s \le t+1;\ 3 \le t \le T-1, \qquad (3.2.6)$$

By using the moment conditions (3.2.1) - (3.2.3) and (3.2.4) - (3.2.6), the root-*N* consistent estimations of the parameters of interest (i.e. γ and β) can be conducted for the model (2.5.1) with (2.5.2).[14] The moment conditions (3.2.2) and (3.2.5) are often referred to as the sequential moment conditions with respect to the dependent variable, which correspond to the standard moment conditions proposed for the simple dynamic

---

[14] When $\beta x_{it} + TD_t$ is used instead of βx_{it} in the model (2.5.1) with (2.5.2), where TD_t is time dummy, $\beta \Delta x_{it}$ and $\beta \Delta x_{i,t+1}$ in $\hbar U_{it}$ and $\hbar \Upsilon_{it}$ are replaced by $\beta \Delta x_{it} + \Delta TD_t$ and $\beta \Delta x_{i,t+1} + \Delta TD_{t+1}$, respectively. In this case, the root-*N* consistent estimations of the first-differenced time dummies ΔTD_t (for $3 \le t \le T$) will be possible jointly with those of γ and β, by using the moment conditions (3.2.1) - (3.2.3) and (3.2.4) - (3.2.6). It is thought that the possibility of the root-*N* consistent estimators is also shown for the dynamic fixed effects logit model with time dummies in addition to the strictly exogenous continuous explanatory variables.

fixed effects logit model in previous section and frequently used in the context of the ordinary dynamic panel data (see Holtz-Eakin et al., 1988, and Arellano and Bond, 1991, etc.).

Table 3 and 4 report the results of the Monte Carlo experiments on the root-N consistent estimators (i.e. the GMM(g-HTD) and GMM(h-HTD) estimators) for the dynamic fixed effects logit model with the strictly exogenous continuous explanatory variable when $T=4$ and 8, respectively.

It can be said that the size alleviations of bias and rmse for the GMM(g-HTD) and GMM(h-HTD) estimators back up the presence of the root-N consistent estimators for the dynamic fixed effects logit model with strictly exogenous continuous explanatory variables. The size alleviation of rmse is pronounced for the case where $T=8$. It is considered that this is due to the increase of sample size in substance.

Roughly speaking, what is true for the model without explanatory variable in previous subsection is true for the model with the strictly exogenous continuous explanatory variable. The larger downward biases for the GMM(g-HTD) and GMM(h-HTD) estimators of the persistence parameter γ are recognizable when the data of the dependent and explanatory variables are more persistent. It is conceivable that the GMM(g-HTD) and GMM(h-HTD) estimators are afflicted with the weak instruments problem.[15]

As the data of the dependent and explanatory variables are persistent, the small sample performances of the GMM(g-HTD) and GMM(h-HTD) estimators for the coefficient β on the explanatory variable also deteriorate. The sizes of bias and rmse with respect to β are small, compared to those with respect to γ. The sizes of bias are especially small. These are similar to the simulation results conducted by Kitazawa (2012) for the static fixed effects logit model.

4. Conclusion

In this paper, the transformations and valid moment conditions were advocated for the

[15] The sizes of bias for the GMM(g-HTD) and GMM(h-HTD) estimators for the persistence parameter γ are larger when $T=8$ than when $T=4$. It is conceivable that this is due to the excess usage of the weak instruments, as is the case with the interpretation in previous subsection.

dynamic fixed effects logit models without explanatory variable and with strictly exogenous continuous explanatory variables. For the model without explanatory variable, the valid moment conditions are constructed based on the error-components structures after the model is transformed into the simple linear panel data models with additive fixed effects, while for the model with strictly exogenous continuous explanatory variables, those are constructed by applying a variety of the HTD transformation after the model is transformed in order that the logit probabilities composed of the fixed effects and the explanatory variables are separated out as the additive terms. The valid moment conditions for the model without explanatory variable include the stationarity moment conditions and two of whose combinations are just the first-order condition of the CMLE proposed by Chamberlain (1985). The high point of the paper is that if the number of time periods of panel data is four or more, the GMM estimators, which are the root-N consistent estimators, can be constructed using the valid moment conditions, for the dynamic fixed effects logit model with strictly exogenous continuous explanatory variables as well as that without explanatory variable. The exploration of the dynamic fixed effects logit models from a traditional angle brought in the fruitful results. As might be expected by not a few researchers, it was the traditional approach that conduced to constructing the root-N consistent estimators for the dynamic fixed effects logit model with strictly exogenous continuous explanatory variables. Traditional, All Too Traditional!

Appendix A.

Plugging $\tanh((\eta_i + \gamma y_{i,t-1})/2) = 2p(\eta_i, y_{i,t-1}) - 1$, $\tanh(\eta_i/2) = 2g(\eta_i) - 1$, and $\tanh(\gamma y_{i,t-1}/2) = \tanh(\gamma/2) y_{i,t-1}$ (due to the fact that y_{it} is binary) into the formula with respect to the hyperbolic function:

$$\tanh((\eta_i + \gamma y_{i,t-1})/2) = \frac{\tanh(\eta_i/2) + \tanh((\gamma y_{i,t-1})/2)}{1 + \tanh(\eta_i/2)\tanh((\gamma y_{i,t-1})/2)}, \tag{A.1}$$

the following relationship is obtained:

$$\begin{aligned} &p(\eta_i, y_{i,t-1}) \\ &= 2(1 - p(\eta_i, y_{i,t-1}))g(\eta_i)\tanh(\gamma/2)y_{i,t-1} + p(\eta_i, y_{i,t-1})\tanh(\gamma/2)y_{i,t-1} \\ &\quad + g(\eta_i)\tanh(\gamma/2)(1 - y_{i,t-1}) + g(\eta_i)(1 - \tanh(\gamma/2)) . \end{aligned} \tag{A.2}$$

Applying $\mathrm{E}[y_{it} \mid \eta_i, y_{i1}, v_i^{t-1}] = p(\eta_i, y_{i,t-1})$,

$$\mathrm{E}[(1-y_{i,t-1})y_{it} \mid \eta_i, y_{i1}, v_i^{t-1}] = g(\eta_i)(1-y_{i,t-1}) \tag{A.3}$$

(which is obtained by utilizing the property that $(1-y_{it})y_{it} = 0$) and

$$\begin{aligned} &\mathrm{E}[(1-y_{it})g(\eta_i) \mid \eta_i, y_{i1}, v_i^{t-1}] \\ &= \mathrm{E}[\mathrm{E}[(1-y_{it})y_{i,t+1} \mid \eta_i, y_{i1}, v_i^{t}] \mid \eta_i, y_{i1}, v_i^{t-1}] = \mathrm{E}[(1-y_{it})y_{i,t+1} \mid \eta_i, y_{i1}, v_i^{t-1}] \end{aligned} \tag{A.4}$$

to (A.2), and then dividing both sides of (A.2) by $1-\tanh(\gamma/2)$, equations (2.1.7) with (2.1.8) are obtained.

Next, the following relationship is obtained:

$$\begin{aligned} &p(\eta_i, y_{i,t-1}) \\ &= 2p(\eta_i, y_{i,t-1})h(\eta_i)\tanh(\gamma/2)(1-y_{i,t-1}) - p(\eta_i, y_{i,t-1})\tanh(\gamma/2)(1-y_{i,t-1}) \\ &\quad + h(\eta_i)\tanh(\gamma/2)y_{i,t-1} + h(\eta_i)(1-\tanh(\gamma/2)) . \end{aligned} \tag{A.5}$$

by replacing $\tanh(\eta_i/2) = 2g(\eta_i) - 1$ and $\tanh(\gamma y_{i,t-1}/2) = \tanh(\gamma/2)y_{i,t-1}$ in (A.1) with $\tanh((\eta_i+\gamma)/2) = 2h(\eta_i) - 1$ and $\tanh((\gamma y_{i,t-1} - \gamma)/2) = -\tanh(\gamma/2)(1-y_{i,t-1})$, respectively. Applying $\mathrm{E}[y_{it} \mid \eta_i, y_{i1}, v_i^{t-1}] = p(\eta_i, y_{i,t-1})$,

$$\mathrm{E}[y_{i,t-1}y_{it} \mid \eta_i, y_{i1}, v_i^{t-1}] = h(\eta_i)y_{i,t-1} \tag{A.6}$$

(which is obtained by utilizing the property that $y_{it}^2 = y_{it}$) and

$$\mathrm{E}[y_{it}h(\eta_i) \mid \eta_i, y_{i1}, v_i^{t-1}] = \mathrm{E}[y_{it}y_{i,t+1} \mid \eta_i, y_{i1}, v_i^{t-1}] \tag{A.7}$$

to (A.5), and then dividing both sides of (A.5) by $1-\tanh(\gamma/2)$, equations (2.1.9) with (2.1.10) are obtained.

Appendix. B.

Replacing the unobservable variables w_{it} and v_{is} in (2.2.1) - (2.2.3) with the observable variables u_{it} and y_{is} respectively gives the following equations:

$$\mathrm{E}[u_{it}] = \mathrm{E}[g(\eta_i)], \qquad \text{for } 2 \le t \le T-1, \tag{B.1}$$

$$\mathrm{E}[y_{is}u_{it}] = \mathrm{E}[g(\eta_i)y_{is}], \qquad \text{for } 1 \le s \le t-1;\ 2 \le t \le T-1. \tag{B.2}$$

The valid moment conditions are constructed based on the compact relationships holding among (B.1) and (B.2).

First, subtracting $\mathrm{E}[u_{i,t-1}]$ from $\mathrm{E}[u_{it}]$ gives the moment conditions (2.2.4). Next, subtracting $\mathrm{E}[y_{is}u_{i,t-1}]$ from $\mathrm{E}[y_{is}u_{it}]$ (for $1 \le s \le t-2$) gives the moment conditions (2.2.5), while subtracting $\mathrm{E}[y_{i,t-2}u_{it}]$ from $\mathrm{E}[y_{i,t-1}u_{it}]$ gives

$$\mathrm{E}[\Delta y_{i,t-1}u_{it}] = \mathrm{E}[g(\eta_i)\Delta y_{i,t-1}], \tag{B.3}$$

where the unobservable variable $g(\eta_i)$ remains to be eliminated. To solve this problem, taking first-difference of (A.3) in Appendix A and then applying law of total expectation to the first-difference gives

$$\mathrm{E}[g(\eta_i)\Delta y_{i,t-1}] = -\mathrm{E}[(1-y_{i,t-1})y_{it} - (1-y_{i,t-2})y_{i,t-1}]. \qquad \text{(B.4)}$$

Then, plugging (B.4) into (B.3) gives the following $T-3$ moment conditions:

$$\mathrm{E}[\Delta y_{i,t-1}u_{it} + (1-y_{i,t-1})y_{it} - (1-y_{i,t-2})y_{i,t-1}] = 0, \qquad \text{for } 3 \le t \le T-1, \quad \text{(B.5)}$$

whose heft-hand sides are equivalent to the subtractions of (2.2.5) for $s = t-2$ from (2.2.4), taking notice of the fact that $(1-y_{it})y_{it} = 0$. Since these equivalences hold without using the expectation operator, the moment conditions (B.5) are superfluous.

The same logic is applied to (2.2.6) - (2.2.8) to give the moment conditions (2.2.9) and (2.2.10). The first-difference of (A.6) in Appendix A is used instead of (A.3) in order to obtain the following $T-3$ moment conditions:

$$\mathrm{E}[\Delta y_{i,t-1}\upsilon_{it} - y_{i,t-1}y_{it} + y_{i,t-2}y_{i,t-1}] = 0, \qquad \text{for } 3 \le t \le T-1, \quad \text{(B.6)}$$

whose left-hand-sides are equivalent to (2.2.10) multiplied by (-1), taking notice of the fact that $(1-y_{it})y_{it} = 0$. Based on the logic similar to that in previous paragraph, the moment conditions (B.6) are superfluous.

Finally, the moment conditions (B.5) and (B.6), which correspond to the additional non-linear moment conditions proposed by Ahn (1990) and Ahn and Schmidt (1995) in the framework of the ordinary dynamic panel data model, can be written by using the moment conditions (2.2.4) and (2.2.5) based on the g-form and the moment conditions (2.2.10) based on the h-form, respectively.

Appendix. C.

Since $\mathrm{E}[g(\eta_i)\Delta y_{i,t-1}] = 0$ according to (2.3.3), (B.3) in Appendix B reduces to the moment conditions (2.3.4). By the same token, the moment conditions (2.3.5) are obtained by paying attention to $\mathrm{E}[h(\eta_i)\Delta y_{i,t-1}] = 0$.

The moment conditions (2.3.4) and (2.3.5) are regarded as the replacements of the moment conditions (B.5) and (B.6) for the case of the stationary dependent variable, respectively. They cannot be regarded as being superfluous.

Appendix D.

The following relationships hold:

$$\mathrm{E}[\partial^2\ell_{it}/\partial\gamma^2] = -(1/4)\,\mathrm{sech}^2(\gamma/2)\,\mathrm{E}[(\Delta y_{it})^2(y_{i,t+1}-y_{i,t-2})^2], \tag{D.1}$$

$$\begin{aligned}\mathrm{E}[(\partial\ell/\partial\gamma)^2] &= (1/4)(1-\tanh^2(\gamma/2))\,\mathrm{E}[(\Delta y_{it})^2(y_{i,t+1}-y_{i,t-2})^2] \\ &\quad -(1/2)\tanh(\gamma/2)\,\mathrm{E}[\partial\ell_{it}/\partial\gamma],\end{aligned} \tag{D.2}$$

where $\mathrm{E}[\partial\ell_{it}/\partial\gamma] = (1/2)\,\mathrm{E}[(\Delta y_{it})(y_{i,t+1}-y_{i,t-2}) - \tanh(\gamma/2)(\Delta y_{it})^2(y_{i,t+1}-y_{i,t-2})^2]$.

Taking notice of the facts that $\mathrm{sech}^2(\gamma/2) = 1-\tanh^2(\gamma/2)$ and $E[\partial\ell_{it}/\partial\gamma]=0$, (D.1) multiplied by (-1) is equivalent to (D.2). This equivalence is conceptually the same as that firstly pointed out by Lee (2002, pp. 84-87) and compactly rewritten by Kitazawa (2012) in the framework of the GMM, on the CMLE for the static fixed effects logit model.

Appendix E.

A tedious calculation proves that the first-order condition with respect to γ for (2.4.1) with (2.4.2), which is multiplied by $\exp(\gamma)+1$, is the empirical counterpart of the following moment condition:

$$\mathrm{E}[A(y_{i,t-2},\Delta y_{it},y_{i,t+1}) + \delta B(y_{i,t-2},y_{i,t-1},y_{it},y_{i,t+1})] = 0, \tag{E.1}$$

where $A(y_{i,t-2},\Delta y_{it},y_{i,t+1}) = \Delta y_{it}y_{i,t+1} - y_{i,t-2}\Delta y_{it}$ and

$$\begin{aligned}B(y_{i,t-2},y_{i,t-1},y_{it},y_{i,t+1}) &= -2y_{i,t-2}y_{i,t-1}y_{it}y_{i,t+1} - y_{i,t-2}y_{it} - y_{i,t-1}y_{i,t+1} \\ &\quad + y_{i,t-2}y_{i,t-1}y_{it} + y_{i,t-2}y_{i,t-1}y_{i,t+1} + y_{i,t-2}y_{it}y_{i,t+1} + y_{i,t-1}y_{it}y_{i,t+1}.\end{aligned}$$

In addition, another tedious calculation (where the facts with respect to the binary variable (i.e. $y_{it}^2 = y_{it}$ and $(1-y_{it})y_{it}=0$) are of assistance) proves that the moment condition (2.4.3) reduces to

$$\mathrm{E}[C(y_{i,t-2},\Delta y_{it}) + \delta B(y_{i,t-2},y_{i,t-1},y_{it},y_{i,t+1})] = 0, \tag{E.2}$$

where $C(y_{i,t-2},\Delta y_{it}) = \Delta y_{it} - 2y_{i,t-2}\Delta y_{it}$.

Further, subtracting (2.1.9) from (2.1.7) gives

$$0 = -\delta\,\Delta y_{it}\Delta y_{i,t+1} - f(\eta_i) + \varpi_{it}, \tag{E.3}$$

where $\varpi_{it} = w_{it} - \omega_{it}$. Applying law of total expectation to the first difference of (E.3) gives

$$\mathrm{E}[\Delta y_{it}y_{i,t+1}] = \mathrm{E}[(1-y_{i,t-2})\Delta y_{it}], \tag{E.4}$$

where the fact that $\Delta y_{it}(y_{it}+y_{i,t-1}) = \Delta y_{it}$ is of assistance and further it should be noted that $y_{it} = g(\eta_i) + v_{it}$ if $\delta = 0$. Accordingly,

$$\mathrm{E}[A(y_{i,t-2}, \Delta y_{it}, y_{i,t+1})] = \mathrm{E}[C(y_{i,t-2}, \Delta y_{it})], \tag{E.5}$$

which indicates the equivalence between (E.1) and (E.2), implying that the first-order condition of the CMLE proposed by Chamberlain (1985) can be written as the moment condition (2.4.3).

Further, it is proved that the moment condition (2.4.3) reduces to the moment condition (2.4.4) if the dependent variable is stationary, by taking notice of the fact that the moment condition (2.4.3) is the plain sum of the moment conditions (B.5) and (B.6) in Appendix B and further by paying attention to the fact that the moment conditions (B.5) and (B.6) are respectively replaced by the moment conditions (2.3.4) and (2.3.5) for the case of the stationary dependent variable.

Appendix F.

The following four lemmas are needed in order to derive equation (2.5.7) with (2.5.8) and (2.5.9):

Lemma F.A: Equation (2.5.3) with (2.5.4) - (2.5.6) can be written as

$$\begin{aligned} p(\eta_i, y_{i,t-1}, x_{it}) &= 2(1 - p(\eta_i, y_{i,t-1}, x_{it}))g(\eta_i, x_{it})\tanh(\gamma/2)y_{i,t-1} \\ &+ p(\eta_i, y_{i,t-1}, x_{it})\tanh(\gamma/2)y_{i,t-1} \\ &+ g(\eta_i, x_{it})\tanh(\gamma/2)(1 - y_{i,t-1}) + g(\eta_i, x_{it})(1 - \tanh(\gamma/2)) . \end{aligned} \tag{F.A.1}$$

Proof: Plugging $\tanh((\eta_i + \gamma y_{i,t-1} + \beta x_{it})/2) = 2p(\eta_i, y_{i,t-1}, x_{it}) - 1$, $\tanh((\eta_i + \beta x_{it})/2) = 2g(\eta_i, x_{it}) - 1$, and $\tanh(\gamma y_{i,t-1}/2) = \tanh(\gamma/2)y_{i,t-1}$ into the formula with respect to the hyperbolic function:

$$\tanh((\eta_i + \gamma y_{i,t-1} + \beta x_{it})/2) = \frac{\tanh((\eta_i + \beta x_{it})/2) + \tanh((\gamma y_{i,t-1})/2)}{1 + \tanh((\eta_i + \beta x_{it})/2)\tanh((\gamma y_{i,t-1})/2)}, \tag{F.A.2}$$

equation (F.A.1) is obtained. Q.E.D.

Lemma F.B: The following relationship holds with respect to $g(\eta_i, x_{i,t+1})$:

$$g(\eta_i, x_{i,t+1}) / g(\eta_i, x_{it}) = (1 - \exp(\beta \Delta x_{i,t+1}))g(\eta_i, x_{i,t+1}) + \exp(\beta \Delta x_{i,t+1}) . \tag{F.B.1}$$

Proof: The derivation is simple, taking notice of the fact that $\exp(\beta \Delta x_{i,t+1})\exp(\eta_i + \beta x_{it}) = \exp(\eta_i + \beta x_{i,t+1})$. Q.E.D.

Lemma F.C: The following relationship holds with respect to $g(\eta_i, x_{i,t+1})$:

$$\begin{aligned}
&\exp(\beta\Delta x_{i,t+1})(1-g(\eta_i,x_{i,t+1}))(1-y_{i,t-1}) \\
&\quad =\exp(\beta\Delta x_{i,t+1})\,\mathrm{E}[(1-y_{i,t-1})(1-y_{it})\,|\,\eta_i,y_{i1},v_i^{t-1},x_i^T] \\
&\quad +(1-\exp(\beta\Delta x_{i,t+1}))\,\mathrm{E}[(1-y_{i,t-1})(1-y_{it})y_{i,t+1}\,|\,\eta_i,y_{i1},v_i^{t-1},x_i^T]\,.
\end{aligned} \tag{F.C.1}$$

Proof: The relationship similar to (A.3) used in Proof of Appendix A holds with respect to $g(\eta_i,x_{it})$:

$$\mathrm{E}[y_{it}\,|\,\eta_i,y_{i1},v_i^{t-1},x_i^T](1-y_{i,t-1})=g(\eta_i,x_{it})(1-y_{i,t-1})\,. \tag{F.C.2}$$

Subtracting (F.C.2) from $(1-y_{i,t-1})$ gives

$$\mathrm{E}[(1-y_{it})\,|\,\eta_i,y_{i1},v_i^{t-1},x_i^T](1-y_{i,t-1})=(1-g(\eta_i,x_{it}))(1-y_{i,t-1})\,. \tag{F.C.3}$$

Taking notice of the relationship that

$g(\eta_i,x_{i,t+1})/g(\eta_i,x_{it})=\exp(\beta\Delta x_{i,t+1})((1-g(\eta_i,x_{i,t+1}))/(1-g(\eta_i,x_{it})))$,

multiplying both sides of (F.C.3) by $g(\eta_i,x_{i,t+1})/g(\eta_i,x_{it})$ gives

$$\begin{aligned}
&(g(\eta_i,x_{i,t+1})/g(\eta_i,x_{it}))\,\mathrm{E}[(1-y_{it})\,|\,\eta_i,y_{i1},v_i^{t-1},x_i^T](1-y_{i,t-1}) \\
&\qquad =\exp(\beta\Delta x_{i,t+1})(1-g(\eta_i,x_{i,t+1}))(1-y_{i,t-1})\,.
\end{aligned} \tag{F.C.4}$$

Further, by utilizing (F.C.2) one period later, it follows that

$$\begin{aligned}
g(\eta_i,x_{i,t+1})\,\mathrm{E}[(1-y_{it})\,|\,\eta_i,y_{i1},v_i^{t-1},x_i^T] &=\mathrm{E}[g(\eta_i,x_{i,t+1})(1-y_{it})\,|\,\eta_i,y_{i1},v_i^{t-1},x_i^T] \\
&=\mathrm{E}[\mathrm{E}[(1-y_{it})y_{i,t+1}\,|\,\eta_i,y_{i1},v_i^{t},x_i^T]\,|\,\eta_i,y_{i1},v_i^{t-1},x_i^T] \\
&=\mathrm{E}[(1-y_{it})y_{i,t+1}\,|\,\eta_i,y_{i1},v_i^{t-1},x_i^T]\,.
\end{aligned} \tag{F.C.5}$$

Plugging (F.B.1) into (F.C.4) and then plugging (F.C.5) into (F.C.4), (F.C.1) is obtained.

Q.E.D.

Lemma F.D: The following relationship holds with respect to $g(\eta_i,x_{i,t+1})$:

$$(1-p(\eta_i,y_{i,t-1},x_{it}))g(\eta_i,x_{i,t+1})=\mathrm{E}[(1-y_{it})y_{i,t+1}\,|\,\eta_i,y_{i1},v_i^{t-1},x_i^T]\,. \tag{F.D.1}$$

Proof: Utilizing both relationships (F.C.5) used in Proof of Lemma F.C and

$$(1-p(\eta_i,y_{i,t-1},x_{it}))g(\eta_i,x_{i,t+1})=\mathrm{E}[(1-y_{it})\,|\,\eta_i,y_{i1},v_i^{t-1},x_i^T]g(\eta_i,x_{i,t+1})\,, \tag{F.D.2}$$

(F.D.1) is obtained. Q.E.D.

Using the lemmas above, the derivation of the g-forms (2.5.7) with (2.5.8) and (2.5.9) is conducted: First, equation (F.A.1) is transformed into

$$
\begin{aligned}
&(1-p(\eta_i, y_{i,t-1}, x_{it})) \\
&\quad = -2(1-p(\eta_i, y_{i,t-1}, x_{it}))g(\eta_i, x_{it})\tanh(\gamma/2)y_{i,t-1} \\
&\quad + (1-p(\eta_i, y_{i,t-1}, x_{it}))\tanh(\gamma/2)y_{i,t-1} \\
&\quad + (1-g(\eta_i, x_{it}))(\tanh(\gamma/2)(1-y_{i,t-1}) + (1-\tanh(\gamma/2))).
\end{aligned} \tag{F.1}
$$

Next, multiplying (F.1) by $g(\eta_i, x_{i,t+1})/g(\eta_i, x_{it})$ and then utilizing the relationship that $g(\eta_i, x_{i,t+1})/g(\eta_i, x_{it}) = \exp(\beta\Delta x_{i,t+1})((1-g(\eta_i, x_{i,t+1}))/(1-g(\eta_i, x_{it})))$ gives

$$
\begin{aligned}
&(g(\eta_i, x_{i,t+1})/g(\eta_i, x_{it}))(1-p(\eta_i, y_{i,t-1}, x_{it})) \\
&\quad = -2(1-p(\eta_i, y_{i,t-1}, x_{it}))g(\eta_i, x_{i,t+1})\tanh(\gamma/2)y_{i,t-1} \\
&\quad + (g(\eta_i, x_{i,t+1})/g(\eta_i, x_{it}))(1-p(\eta_i, y_{i,t-1}, x_{it}))\tanh(\gamma/2)y_{i,t-1} \\
&\quad + \exp(\beta\Delta x_{i,t+1})(1-g(\eta_i, x_{i,t+1}))(\tanh(\gamma/2)(1-y_{i,t-1}) + (1-\tanh(\gamma/2))).
\end{aligned} \tag{F.2}
$$

The g-forms (2.5.7) with (2.5.8) and (2.5.9) are obtained by dividing (F.2) by $\exp(\beta\Delta x_{i,t+1})(1-\tanh(\gamma/2))$, after plugging (F.B.1) into (F.2) and then plugging (F.C.1) and (F.D.1) into (F.2).

Appendix G.

The following four lemmas are needed in order to derive equation (2.5.10) with (2.5.11) and (2.5.12):

Lemma G.A: Equation (2.5.3) with (2.5.4) - (2.5.6) can be written as

$$
\begin{aligned}
&p(\eta_i, y_{i,t-1}, x_{it}) = 2p(\eta_i, y_{i,t-1}, x_{it})h(\eta_i, x_{it})\tanh(\gamma/2)(1-y_{i,t-1}) \\
&\quad - p(\eta_i, y_{i,t-1}, x_{it})\tanh(\gamma/2)(1-y_{i,t-1}) \\
&\quad + h(\eta_i, x_{it})\tanh(\gamma/2)y_{i,t-1} + h(\eta_i, x_{it})(1-\tanh(\gamma/2)).
\end{aligned} \tag{G.A.1}
$$

Proof: Plugging $\tanh((\eta_i + \gamma y_{i,t-1} + \beta x_{it})/2) = 2p(\eta_i, y_{i,t-1}, x_{it}) - 1$,
$\tanh((\eta_i + \gamma + \beta x_{it})/2) = 2h(\eta_i, x_{it}) - 1$,
and $\tanh((\gamma y_{i,t-1} - \gamma)/2) = -\tanh(\gamma/2)(1-y_{i,t-1})$ into the formula with respect to the hyperbolic function:

$$
\tanh((\eta_i + \gamma y_{i,t-1} + \beta x_{it})/2) = \frac{\tanh((\eta_i + \gamma + \beta x_{it})/2) + \tanh((\gamma y_{i,t-1} - \gamma)/2)}{1 + \tanh((\eta_i + \gamma + \beta x_{it})/2)\tanh((\gamma y_{i,t-1} - \gamma)/2)}, \tag{G.A.2}
$$

equation (G.A.1) is obtained. Q.E.D.

Lemma G.B: The following relationship holds with respect to $h(\eta_i, x_{i,t+1})$:

$$h(\eta_i, x_{i,t+1}) / h(\eta_i, x_{it}) = (1-\exp(\beta\Delta x_{i,t+1}))h(\eta_i, x_{i,t+1}) + \exp(\beta\Delta x_{i,t+1}) . \qquad \text{(G.B.1)}$$

Proof: The derivation is simple, taking notice of the fact that

$\exp(\beta\Delta x_{i,t+1})\exp(\eta_i + \gamma + \beta x_{it}) = \exp(\eta_i + \gamma + \beta x_{i,t+1})$. Q.E.D.

Lemma G.C: The following relationship holds with respect to $h(\eta_i, x_{i,t+1})$:

$$\begin{aligned} h(\eta_i, x_{i,t+1}) y_{i,t-1} &= \exp(\beta\Delta x_{i,t+1}) \mathrm{E}[y_{i,t-1} y_{it} \mid \eta_i, y_{i1}, v_i^{t-1}, x_i^T] \\ &\quad + (1-\exp(\beta\Delta x_{i,t+1})) \mathrm{E}[y_{i,t-1} y_{it} y_{i,t+1} \mid \eta_i, y_{i1}, v_i^{t-1}, x_i^T] . \end{aligned} \qquad \text{(G.C.1)}$$

Proof: The relationship similar to (A.6) used in Proof of Appendix A holds with respect to $h(\eta_i, x_{it})$:

$$\mathrm{E}[y_{it} \mid \eta_i, y_{i1}, v_i^{t-1}, x_i^T] y_{i,t-1} = h(\eta_i, x_{it}) y_{i,t-1} . \qquad \text{(G.C.2)}$$

Multiplying both sides of (G.C.2) by $h(\eta_i, x_{i,t+1}) / h(\eta_i, x_{it})$ gives

$$(h(\eta_i, x_{i,t+1}) / h(\eta_i, x_{it})) \mathrm{E}[y_{it} \mid \eta_i, y_{i1}, v_i^{t-1}, x_i^T] y_{i,t-1} = h(\eta_i, x_{i,t+1}) y_{i,t-1} . \qquad \text{(G.C.3)}$$

Further, by utilizing (G.C.2) one period later, it follows that

$$\begin{aligned} h(\eta_i, x_{i,t+1}) \mathrm{E}[y_{it} \mid \eta_i, y_{i1}, v_i^{t-1}, x_i^T] &= \mathrm{E}[h(\eta_i, x_{i,t+1}) y_{it} \mid \eta_i, y_{i1}, v_i^{t-1}, x_i^T] \\ &= \mathrm{E}[\mathrm{E}[y_{it} y_{i,t+1} \mid \eta_i, y_{i1}, v_i^{t}, x_i^T] \mid \eta_i, y_{i1}, v_i^{t-1}, x_i^T] \\ &= \mathrm{E}[y_{it} y_{i,t+1} \mid \eta_i, y_{i1}, v_i^{t-1}, x_i^T] . \end{aligned} \qquad \text{(G.C.4)}$$

Plugging (G.B.1) into (G.C.3) and then plugging (G.C.4) into (G.C.3), (G.C.1) is obtained.

Q.E.D.

Lemma G.D: The following relationship holds with respect to $h(\eta_i, x_{i,t+1})$:

$$p(\eta_i, y_{i,t-1}, x_{it}) h(\eta_i, x_{i,t+1}) = \mathrm{E}[y_{it} y_{i,t+1} \mid \eta_i, y_{i1}, v_i^{t-1}, x_i^T] . \qquad \text{(G.D.1)}$$

Proof: Utilizing both relationships (G.C.4) used in Proof of Lemma G.C and

$$p(\eta_i, y_{i,t-1}, x_{it}) h(\eta_i, x_{i,t+1}) = \mathrm{E}[y_{it} \mid \eta_i, y_{i1}, v_i^{t-1}, x_i^T] h(\eta_i, x_{i,t+1}) , \qquad \text{(G.D.2)}$$

(G.D.1) is obtained. Q.E.D.

Using the lemmas above, the derivation of the h-forms (2.5.10) with (2.5.11) and (2.5.12) is conducted: First, multiplying (G.A.1) by $h(\eta_i, x_{i,t+1}) / h(\eta_i, x_{it})$ gives

$$\begin{aligned} &(h(\eta_i, x_{i,t+1}) / h(\eta_i, x_{it})) p(\eta_i, y_{i,t-1}, x_{it}) \\ &\quad = 2 p(\eta_i, y_{i,t-1}, x_{it}) h(\eta_i, x_{i,t+1}) \tanh(\gamma/2)(1-y_{i,t-1}) \\ &\quad - (h(\eta_i, x_{i,t+1}) / h(\eta_i, x_{it})) p(\eta_i, y_{i,t-1}, x_{it}) \tanh(\gamma/2)(1-y_{i,t-1}) \\ &\quad + h(\eta_i, x_{i,t+1}) \tanh(\gamma/2) y_{i,t-1} + h(\eta_i, x_{i,t+1})(1-\tanh(\gamma/2)) . \end{aligned} \qquad \text{(G.1)}$$

The h-forms (2.5.10) with (2.5.11) and (2.5.12) are obtained by dividing (G.1) by $1-\tanh(\gamma/2)$, after plugging (G.B.1) into (G.1) and then plugging (G.C.1) and (G.D.1) into (G.1).

Appendix H.

The following formula holds for the hyperbolic tangent function:

$$\begin{aligned} g(\eta_i, x_{i,t+1}) - p(\eta_i, y_{i,t-2}, x_{i,t-1}) & \\ = \tanh((1/2)(-\gamma y_{i,t-2} + \beta(\Delta x_{i,t+1} + \Delta x_{it}))) & \\ \times (g(\eta_i, x_{i,t+1}) + p(\eta_i, y_{i,t-2}, x_{i,t-1}) - 2g(\eta_i, x_{i,t+1}) p(\eta_i, y_{i,t-2}, x_{i,t-1})). & \end{aligned} \tag{H.1}$$

Further, the following relationships are obtained from (2.5.1) with (2.5.2) and (2.5.7) with (2.5.8) and (2.5.9):

$$g(\eta_i, x_{i,t+1}) = \mathrm{E}[U_{it} \mid \eta_i, y_{i1}, v_i^{t-1}, x_i^T], \tag{H.2}$$

$$p(\eta_i, y_{i,t-2}, x_{i,t-1}) = \mathrm{E}[y_{i,t-1} \mid \eta_i, y_{i1}, v_i^{t-2}, x_i^T], \tag{H.3}$$

$$\begin{aligned} g(\eta_i, x_{i,t+1}) p(\eta_i, y_{i,t-2}, x_{i,t-1}) & \\ = \mathrm{E}[U_{it}\, y_{i,t-1} \mid \eta_i, y_{i1}, v_i^{t-1}, x_i^T] - v_{i,t-1} g(\eta_i, x_{i,t+1}), & \end{aligned} \tag{H.4}$$

Plugging (H.2) - (H.4) into (H.1) and then taking the expectation conditional on the information $(\eta_i, y_{i1}, v_i^{t-2}, x_i^T)$ for both sides of (H.1) gives the moment conditions (2.5.13) with (2.5.14).

Next, by replacing $g(\eta_i, x_{i,t+1})$, $-\gamma y_{i,t-2}$, and U_{it} in (H.1) - (H.4) with $h(\eta_i, x_{i,t+1})$, $\gamma(1-y_{i,t-2})$, and Υ_{it} respectively, the moment conditions (2.5.15) with (2.5.16) are obtained.

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.

Al-Sadoon, M.M., Li, T., Pesaran, M.H., 2012. An exponential class of dynamic binary choice panel data models with fixed effects. CESifo, Working Paper No. 4033.

Arellano, M., Bond, S., 1991. Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. Review of Economic Studies 58, 277-297.

Arellano, M., Bover, O., 1995. Another look at the instrumental variables estimation of error-component models. Journal of Econometrics 68, 29-51.

Baltagi, B.H., 2009. A companion to econometric analysis of panel data. John Wiley & Sons Ltd, Chichester.

Bartolucci F., Nigro, V., 2010. A dynamic model for binary panel data with unobserved heterogeneity admitting a $\sqrt{N}$ -consistent conditional estimator. Econometrica 79, 719–733.

Bartolucci F., Nigro, V., 2012. Pseudo conditional maximum likelihood estimation of the dynamic logit model for binary panel data. Journal of Econometrics 170, 102–116

Bester, A., Hansen, C., 2009. A penalty function approach to bias reduction in nonlinear panel models with fixed effects. Journal of Business and Economic Statistics 27, 131-148.

Blundell, R., Bond, S., 1998. Initial conditions and moment restrictions in dynamic panel data models. Journal of Econometrics 87, 115-143.

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

Bonhomme, S. 2012. Functional differencing. Econometrica 80, 1337-1385.

Bound, J., Jaeger, D.A., Baker, R., 1995. Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variables is weak. Journal of the American Statistical Association 90, 443–450

Buchinsky, M., Hahn, J., Kim, K., 2010. Semiparametric information bound of dynamic discrete choice models. Economics Letters 108, 109-112.

Carro, J.M., 2007. Estimating dynamic panel data discrete choice models with fixed effects. Journal of Econometrics 140, 503-528.

Chamberlain, G., 1980. Analysis of covariance with qualitative data. Review of Economic Studies 47, 225-238.

Chamberlain, G., 1985. Heterogeneity, omitted variable bias, and duration dependence. In: J.J. Heckman and B. Singer, Eds., Longitudinal Analysis of Labor Market Data, Cambridge University Press, Cambridge, 3–38.

Chintagunta, P., Kyriazidou, E., Perktold, J., 2001. Panel data analysis of household brand choices. Journal of Econometrics 103, 111-153.

Fernández-Val, I., 2009. Fixed effects estimation of structural parameters and marginal effects in panel probit models. Journal of Econometrics 150, 71-85.

Hahn, J., 2001. The information bound of a dynamic panel logit model with fixed effects. Econometric Theory 17, 913-932.

Hahn, J., Kuersteiner, G., 2011. Bias reduction for dynamic nonlinear panel models with fixed effects. Econometric Theory 27, 1152–1191.

Hall, B.H., Cummins, C., 2009. TSP 5.1 User's Guide. TSP International, Palo Alto, CA.

Hansen, L.P., 1982. Large sample properties of generalized method of moments estimators. Econometrica 50, 1029–1054.

Holtz-Eakin, D., Newey, W., Rosen, H.S., 1988. Estimating vector autoregressions with panel data. Econometrica 56, 1371–1395.

Honoré, B.E., Kyriazidou, E. 2000. Panel data discrete choice models with lagged dependent variables. Econometrica 68, 839-874.

Hsiao, C., 2003. Analysis of Panel Data, 2nd Edition. Cambridge University Press, Cambridge.

Johnson, E.G., 2004. Identification in discrete choice models with fixed effects. Working Paper, Bureau of Labor Statistics.

Kitazawa, Y., 2012. Hyperbolic transformation and average elasticity in the framework of the fixed effects logit model. Theoretical Economics Letters 2, 192-199.

Kyriazidou, E., 2010. Nonlinear panel data models. In: S. N. Durlauf and E. B. Blume, Eds., Microeconometrics, Palgrave and Macmillan, Basingstoke, 154-168.

Lee, M.J., 2002. Panel Data Econometrics. Academic Press, London.

Neyman, J., Scott, E.L., 1948. Consistent estimates based on partially consistent observations. Econometrica 16, 1-32.

Pesaran, M.H., Timermann, A., 2009. Testing dependence among serially correlated multicategory variables. Journal of the American Statistical Association 104, 325-337.

Rasch, G., 1960. Probabilistic Models for Some Intelligence and Attainment Tests. The

Danish Institute for Educational Research.

Rasch, G., 1961. On general laws and the meaning of measurement in psychology. Proceeding of the 4th Berkeley Symposium on Mathematical Statistics and Probability 4, 321-333.

Staiger, D., Stock, J.H., 1997. Instrumental variables regression with weak instruments. Econometrica 65, 557-586.

Wooldridge, J.M., 2011. Econometrics: panel data methods. In: R.A. Meyers, Eds., Complex Systems in Finance and Econometrics, Springer, New York, 215-237.

Yu, G., Gao, W., Shi, N.Z., 2012. A note on the estimation problem of dynamic binary panel data model with fixed effects. Pakistan Journal of Statistics 28, 271-278.

Table 1. Monte Carlo results for the simple dynamic fixed effects logit model, $T = 4$

	N = 1000		N = 5000		N = 10000	
	bias	rmse	bias	rmse	bias	rmse
Simulation (1a)						
GMM(g-STD) γ	**−0.034**	**0.232**	**−0.008**	**0.102**	**−0.005**	**0.072**
GMM(g-SYS) γ	−0.033	0.213	−0.007	0.092	−0.005	0.065
GMM(h-STD) γ	**−0.032**	**0.231**	**−0.008**	**0.102**	**−0.005**	**0.072**
GMM(h-SYS) γ	−0.027	0.214	−0.006	0.095	−0.004	0.066
GMM(FOC-o) γ	**−0.004**	**0.207**	**−0.002**	**0.091**	**−0.001**	**0.064**
GMM(FOC-s) γ	−0.029	0.306	−0.004	0.128	−0.004	0.089
Simulation (1b)						
GMM(g-STD) γ	**−0.048**	**0.278**	**−0.013**	**0.119**	**−0.005**	**0.084**
GMM(g-SYS) γ	−0.045	0.252	−0.011	0.107	−0.005	0.075
GMM(h-STD) γ	**−0.049**	**0.282**	**−0.013**	**0.120**	**−0.005**	**0.084**
GMM(h-SYS) γ	−0.041	0.254	−0.009	0.110	−0.004	0.078
GMM(FOC-o) γ	**−0.008**	**0.246**	**−0.004**	**0.106**	**−0.001**	**0.075**
GMM(FOC-s) γ	−0.045	0.515	−0.005	0.145	−0.004	0.104
Simulation (1c)						
GMM(g-STD) γ	**−0.394**	**0.841**	**−0.058**	**0.315**	**−0.021**	**0.206**
GMM(g-SYS) γ	−0.378	0.910	−0.043	0.289	−0.015	0.199
GMM(h-STD) γ	**−0.329**	**0.771**	**−0.045**	**0.295**	**−0.020**	**0.196**
GMM(h-SYS) γ	−0.274	0.695	−0.030	0.278	−0.013	0.188
GMM(FOC-o) γ	**0.086**	**0.532**	**0.017**	**0.213**	**0.011**	**0.149**
GMM(FOC-s) γ	0.053	0.602	0.012	0.238	0.009	0.166
Simulation (1d)						
GMM(g-STD) γ	**−0.414**	**0.865**	**−0.067**	**0.342**	**−0.025**	**0.211**
GMM(g-SYS) γ	−0.390	0.850	−0.042	0.295	−0.019	0.201
GMM(h-STD) γ	**−0.421**	**0.880**	**−0.068**	**0.327**	**−0.031**	**0.212**
GMM(h-SYS) γ	−0.327	0.751	−0.043	0.295	−0.020	0.197
GMM(FOC-o) γ	**0.083**	**0.535**	**0.018**	**0.219**	**0.009**	**0.150**
GMM(FOC-s) γ	0.053	0.596	0.013	0.241	0.006	0.167

Notes: 1) The parameter settings in the DGP are as follows: Simulation (1a): $\gamma = 0.5$; $\sigma_\eta^2 = 0.5$. Simulation (1b): $\gamma = 0.5$; $\sigma_\eta^2 = 1.5$. Simulation (1c): $\gamma = 2.5$; $\sigma_\eta^2 = 0.5$. Simulation (1d): $\gamma = 2.5$; $\sigma_\eta^2 = 1.5$. 2) Inappropriate replications (i.e. the replications where the linear estimates of δ are less than minus one, etc.) are eliminated in calculating the statistics. Their number is zero or extremely small for each GMM estimator in each parameter setting. 3) In each of the GMM estimations, the initial consistent estimate is obtained by using the inverse of cross-sectional average of the products between the instruments matrix as the non-optimal weighting matrix, where it should be noted that the components of the moment conditions used are decomposed into the products of the transformations and the instruments.

Table 2. Monte Carlo results for the simple dynamic fixed effects logit model, $T = 8$

	N = 1000		N = 5000		N = 10000	
	bias	rmse	bias	rmse	bias	rmse
Simulation (1a)						
GMM(g-STD) γ	**−0.066**	**0.107**	**−0.012**	**0.038**	**−0.006**	**0.026**
GMM(g-SYS) γ	−0.049	0.090	−0.009	0.034	−0.004	0.023
GMM(h-STD) γ	**−0.058**	**0.101**	**−0.011**	**0.038**	**−0.005**	**0.026**
GMM(h-SYS) γ	−0.050	0.089	−0.009	0.034	−0.005	0.023
GMM(FOC-o) γ	**−0.006**	**0.095**	**0.000**	**0.042**	**−0.001**	**0.030**
GMM(FOC-s) γ	−0.004	0.103	0.000	0.046	−0.001	0.033
Simulation (1b)						
GMM(g-STD) γ	**−0.090**	**0.134**	**−0.017**	**0.045**	**−0.008**	**0.031**
GMM(g-SYS) γ	−0.064	0.107	−0.012	0.038	−0.005	0.026
GMM(h-STD) γ	**−0.084**	**0.130**	**−0.016**	**0.045**	**−0.007**	**0.031**
GMM(h-SYS) γ	−0.068	0.109	−0.013	0.038	−0.006	0.026
GMM(FOC-o) γ	**−0.007**	**0.107**	**−0.001**	**0.048**	**0.000**	**0.034**
GMM(FOC-s) γ	−0.005	0.116	−0.001	0.052	0.000	0.037
Simulation (1c)						
GMM(g-STD) γ	**−1.118**	**1.243**	**−0.102**	**0.139**	**−0.037**	**0.072**
GMM(g-SYS) γ	−0.938	1.027	−0.090	0.128	−0.035	0.070
GMM(h-STD) γ	**−0.695**	**0.780**	**−0.085**	**0.125**	**−0.038**	**0.074**
GMM(h-SYS) γ	−0.566	0.639	−0.062	0.107	−0.030	0.069
GMM(FOC-o) γ	**−0.020**	**0.253**	**0.000**	**0.109**	**−0.002**	**0.078**
GMM(FOC-s) γ	−0.017	0.258	0.000	0.113	−0.002	0.081
Simulation (1d)						
GMM(g-STD) γ	**−1.176**	**1.310**	**−0.118**	**0.156**	**−0.042**	**0.077**
GMM(g-SYS) γ	−0.920	1.002	−0.096	0.135	−0.038	0.074
GMM(h-STD) γ	**−0.918**	**1.023**	**−0.108**	**0.146**	**−0.044**	**0.080**
GMM(h-SYS) γ	−0.662	0.745	−0.071	0.114	−0.032	0.071
GMM(FOC-o) γ	**−0.024**	**0.256**	**−0.002**	**0.111**	**−0.001**	**0.079**
GMM(FOC-s) γ	−0.022	0.260	−0.002	0.115	−0.001	0.082

Notes: See Notes in Table 1.

Table 3. Monte Carlo results for the dynamic fixed effects logit model with the strictly exogenous continuous explanatory variable, $T = 4$

		N = 1000		N = 5000		N = 10000	
		bias	rmse	bias	rmse	bias	rmse
Simulation (2a)							
GMM(g-HTD)	γ	**−0.047**	**0.261**	**−0.007**	**0.113**	**−0.004**	**0.079**
	β	0.006	0.143	0.002	0.062	0.000	0.044
GMM(h-HTD)	γ	**−0.042**	**0.260**	**−0.005**	**0.114**	**−0.003**	**0.079**
	β	0.003	0.142	0.002	0.063	0.000	0.044
Simulation (2b)							
GMM(g-HTD)	γ	**−0.068**	**0.339**	**−0.012**	**0.142**	**−0.004**	**0.100**
	β	0.015	0.198	0.003	0.086	0.001	0.060
GMM(h-HTD)	γ	**−0.059**	**0.340**	**−0.009**	**0.146**	**−0.003**	**0.102**
	β	0.006	0.192	0.002	0.086	0.001	0.061
Simulation (2c)							
GMM(g-HTD)	γ	**−0.146**	**0.556**	**−0.030**	**0.224**	**−0.010**	**0.156**
	β	0.035	0.316	0.009	0.138	0.003	0.096
GMM(h-HTD)	γ	**−0.131**	**0.549**	**−0.023**	**0.233**	**−0.010**	**0.160**
	β	0.022	0.307	0.007	0.138	0.003	0.097

Notes: 1) The parameter settings in the DGP are as follows: Simulation (2a): $\gamma = 0.5$; $\beta = 0.5$; $\rho = 0.5$; $\tau = 0.1$; $\sigma_\eta^2 = 0.5$; $\sigma_\varepsilon^2 = 0.5$. Simulation (2b): $\gamma = 0.8$; $\beta = 0.8$; $\rho = 0.7$; $\tau = 0.1$; $\sigma_\eta^2 = 0.5$; $\sigma_\varepsilon^2 = 0.5$. Simulation (2c): $\gamma = 1.1$; $\beta = 1.1$; $\rho = 0.9$; $\tau = 0.1$; $\sigma_\eta^2 = 0.5$; $\sigma_\varepsilon^2 = 0.5$. 2) Inappropriate replications (i.e. the non-convergence replications) are eliminated in calculating the statistics. Their number is zero or extremely small for each GMM estimator in each parameter setting. 3) In each of the GMM estimations, the initial consistent estimate is obtained by using the inverse of cross-sectional average of the products between the instruments matrix as the non-optimal weighting matrix, where it should be noted that the components of the moment conditions used are decomposed into the products of the transformations and the instruments. 4) The values of the Monte Carlo statistics are obtained using the true values of the parameters of interest as the starting values in the optimization for each replication. The values of the statistics obtained using the true values are almost the same as those obtained using two different types of the starting values.

Table 4. Monte Carlo results for the dynamic fixed effects logit model with the strictly exogenous continuous explanatory variable, $T = 8$

		N = 1000		N = 5000		N = 10000	
		bias	rmse	bias	rmse	bias	rmse
Simulation (2a)							
GMM(g-HTD)	γ	**−0.047**	**0.110**	**−0.010**	**0.044**	**−0.004**	**0.030**
	β	0.000	0.060	0.000	0.026	0.000	0.019
GMM(h-HTD)	γ	**−0.038**	**0.105**	**−0.008**	**0.044**	**−0.003**	**0.031**
	β	−0.005	0.060	−0.001	0.027	−0.001	0.019
Simulation (2b)							
GMM(g-HTD)	γ	**−0.079**	**0.152**	**−0.016**	**0.059**	**−0.007**	**0.040**
	β	0.007	0.082	0.002	0.036	0.001	0.026
GMM(h-HTD)	γ	**−0.067**	**0.146**	**−0.013**	**0.059**	**−0.006**	**0.041**
	β	−0.007	0.082	−0.001	0.037	0.000	0.026
Simulation (2c)							
GMM(g-HTD)	γ	**−0.198**	**0.294**	**−0.035**	**0.099**	**−0.016**	**0.067**
	β	0.025	0.128	0.010	0.059	0.005	0.041
GMM(h-HTD)	γ	**−0.179**	**0.279**	**−0.034**	**0.102**	**−0.017**	**0.069**
	β	0.006	0.125	0.006	0.059	0.004	0.042

Notes: See Notes in Table 3.